

Mémoire de Master 2 :
Interface perceptive de contrôle dans un logiciel
de modélisation par synthèse physique : *Modalys*

Reinaldo DOS SANTOS

Etudiant en Master 2 Sciences et Technologie, spécialité MIS, parcours ATIAM
Université Pierre et Marie Curie, Jussieu, Paris VI

Laboratoire d'accueil : IRCAM, Paris
Equipe : Perception & Design Sonores
Responsables : Nicolas MISDARIIS et Olivier HOUIX

Mars / Juin 2006

Résumé

Parmi les nombreux systèmes de synthèse sonore existant actuellement, **Modalys** fait partie de ceux qui utilisent la synthèse par modèle physique, permettant ainsi un contrôle bas niveau de la synthèse, c'est-à-dire un contrôle des paramètres physiques de la source sonore. S'il ouvre de nouvelles perspectives au compositeur, ce type de synthèse n'est pas complètement intuitif. Pour pallier à cette lacune, nous tentons de définir une approche de contrôle de haut niveau, reflétant notre perception.

Partant de l'état actuel de recherche sur ce sujet, nous décidons de cibler spécifiquement l'interaction frottée, c'est-à-dire le frottement d'un archet sur une corde, à travers l'étude de l'influence des paramètres de la colophane de l'archet sur la perception des sons. Nous commençons par nous familiariser avec la modélisation sous **Modalys** de cette interaction, afin de générer un corpus de sons en faisant varier ces paramètres.

A partir du corpus de sons ainsi obtenu, nous étudions ses variations acoustiques en nous inspirant de notre contexte d'étude afin de trouver des descripteurs du timbre dont la variation est bien expliquée par celle des paramètres acoustiques cités précédemment.

L'étude de l'influence des différents paramètres permet de réduire le corpus à un nombre restreint de sons, permettant alors une expérience de perception du timbre. Ce test de perception permet de définir des dimensions perceptives, que l'on tente de corréler avec les descripteurs acoustiques du timbre, puis de définir sémantiquement grâce à ses descripteurs et aux verbalisations des sujets (par exemple : brillance, richesse spectrale).

Remerciements

Je remercie tout d'abord les responsables de l'équipe *Perception & Design Sonores*, Patrick Susini, Nicolas Misdariis et Olivier Houix de m'avoir accueilli pour ce stage et de m'avoir fait profiter de leur expérience et de leurs conseils.

Je remercie particulièrement mes responsables de stage Olivier Houix et Nicolas Misdariis, pour leur soutien et leur attention au quotidien.

Je remercie ensuite Nicholas Ellis et Joël Bensoam, pour leur aide sur Modalys.

Enfin je souhaite remercier toute l'équipe *Perception & Design Sonores*, qui avec sa bonne humeur a favorisé une excellente ambiance de travail et de partage d'idées pendant ces quelques mois.

Table des matières

Introduction	4
1 Etat de l'art	7
1.1 Stage de Mael Derio	7
1.2 Objectifs du stage	7
1.3 Lien frappé - frotté : <i>Rosin parameters</i>	8
1.3.1 Définition du graphe caractéristique de l'interface frottée	8
1.3.2 Présentation de l'interaction frappée comme un cas limite de l'interaction frottée	9
2 Présentation des outils	11
2.1 Création du son : Modalys	11
2.2 Descripteurs acoustiques	11
2.2.1 Analyse acoustique des sons : Matlab	11
2.2.2 Analyse statistique des descriptions acoustiques : R	13
2.2.3 Mise en place du test de dissemblance : PsiExp	14
3 Influences relatives des paramètres de colophane	15
3.1 Pas de continuité dans l'espace de l'interaction "contact"	15
3.2 Influence de r_1 et r_2 négligeable	16
3.2.1 Influence de r_1 quasi nulle	16
3.2.2 Influence de r_2 négligeable	16
3.3 Paramètres fixes	18
3.4 Paramètres variables	18
3.5 Comparaison des tendances de variation	18
3.6 Conclusion	19
4 Tests perceptifs à partir du corpus de sons	20
4.1 Expérience de dissemblance	20
4.1.1 Expérience	20
4.1.2 Matériel	21
4.1.3 Sujets	21
4.2 Etude des corrélations entre dissemblance et <i>Rosin parameters</i>	21

4.2.1	Synopsis d'une analyse multidimensionnelle	21
4.2.2	Choix du modèle spatial	22
4.2.3	Verbalisations	22
4.2.4	L'espace perceptif	23
4.3	Etude des corrélations	23
4.3.1	Dimension1	24
4.3.2	Dimension 2	26
4.4	Conclusion	28
Conclusion		28
Références		29
Annexe A : Descripteurs Acoustiques		30
Annexe B : Programme Modalys		34
Annexe C : Fonctionnement de Modalys		38
Annexe D : Consigne de l'expérience		43
Annexe E : Analyses statistiques		45

Introduction

L'enjeu de la synthèse sonore est considérable de nos jours, et la demande des compositeurs d'avoir accès à des sons de synthèse est particulièrement importante. La synthèse par modélisation physique présente l'avantage d'avoir un rendu proche de la réalité, mais également de générer des sons qui, tout en restant naturels, c'est-à-dire dont la source est un instrument naturel, dépassent les contraintes de la lutherie classique.

Modalys fait partie de ces logiciels qui utilisent la modélisation physique. Il permet de synthétiser des éléments de base, en distinguant les excitateurs (souffle, plectres, archets, maillets...) et les résonateurs (cordes, plaques, tubes...), et de les assembler pour créer des instruments plus complexes, puis de les faire interagir. Il est ainsi possible de régler les caractéristiques mécaniques de ces éléments, et **Modalys** synthétise le son à partir de modèles physiques utilisant l'analyse modale, c'est-à-dire calculant d'abord les modes de vibration propres de chaque élément puis le son obtenu par une combinaison des apports de chaque mode de vibration.

Ce type de synthèse présente une spécificité non négligeable par rapport aux autres types de synthèse : il permet de créer des sons dont on peut qualifier la source, selon ses caractéristiques mécaniques principalement. Par contre, il ne permet pas d'obtenir des sons qu'on pourrait qualifier en termes sémantiques sans nécessairement faire référence à une source produisant le son. Il peut donc être intéressant, pour un logiciel comme **Modalys**, de permettre, en plus de la description "bas niveau" consistant à gérer les paramètres mécaniques de la source (tels que la densité ou la masse volumique), une approche "haut niveau" où le contrôle se ferait de manière perceptive.

Nous allons pour notre part nous focaliser sur un système simple : la corde frottée par un archet. L'idée est alors de faire varier les caractéristiques mécaniques de l'archet (paramètres de colophane), afin de déterminer perceptivement les différences qui existent entre les sons obtenus. Le but est d'obtenir une relation mathématique entre des descripteurs reflétant la perception et nos caractéristiques mécaniques. Ce lien permettrait alors de créer une interface perceptive, qui contrôlerait les caractéristiques mécaniques de notre archet en vertu de notre lien mathématique, et d'obtenir le son voulu.

Nous commencerons donc par décrire le travail bibliographique et l'état de l'art, en particulier à travers le travail de Mael Derio sur la question de la plaque frappée. Puis nous présenterons les outils qui nous permettront de mener à bien cette étude. Puis nous présenterons l'influence respective des différents paramètres mécaniques étudiés, qui aboutiront à la création du corpus de sons que nous utiliserons alors pour mener nos tests perceptifs. C'est après l'analyse de ceux-ci que nous pourrions établir la relation entre les paramètres mécaniques et les dimensions perceptives.

Chapitre 1

Etat de l'art

1.1 Stage de Mael Derio

En 2005, Mael Derio [Der05], au cours de son stage de fin de Master 2, a créé une interface de contrôle perceptive pour les plaques frappées. Le protocole d'étude qu'il a employé est tout à fait similaire au nôtre. Son travail consistait en l'étude perceptive des paramètres mécaniques d'une plaque frappée à excitation constante. Ainsi, on ne modifiait pas la frappe, mais on se focalisait sur l'étude de l'influence perceptive des deux paramètres mécaniques d'amortissement, propres à la plaque : *const-loss* et *freq-loss*.

Il s'agissait tout d'abord de calculer les variations acoustiques de différents sons obtenus pour des valeurs différentes des paramètres mécaniques. Ce calcul se faisait à l'aide de descripteurs acoustiques. L'analyse statistique des résultats ainsi obtenus a permis de démontrer la cohérence de limiter l'étude aux deux paramètres d'amortissement. Il a également confirmé l'indépendance de ceux-ci entre eux.

Dès lors, il s'agissait de créer un corpus de 16 sons permettant de mettre en place une expérience de dissemblance sur une trentaine de sujets. L'analyse des résultats de cette expérience a permis de mettre en exergue un espace à trois dimensions. Il les a nommées "impulsivité", "résonance", et "brillance".

1.2 Objectifs du stage

L'objectif de ce stage est de transposer ce protocole sur l'étude perceptive des paramètres d'excitation à structure mécanique constante. En ce qui nous concerne, nous nous limiterons à la question de la corde frottée. Ainsi, nous chercherons à définir l'influence perceptive des paramètres mécaniques de l'archet.

Le but secondaire de ce stage est également de se familiariser avec les différents outils que nous utiliserons au cours de cette étude :

- **Modalys**, pour la création des sons
- les descripteurs acoustiques de **Matlab**
- **R**, pour l'analyse statistique des descriptions acoustiques
- **PsiExp** pour la mise en place du test de dissemblance

1.3 Lien frappé - frotté : *Rosin parameters*

1.3.1 Définition du graphe caractéristique de l'interface frottée

L'interaction frottée peut être intégralement définie de façon mécanique, à travers la relation non linéaire représentée sur le graphe suivant :

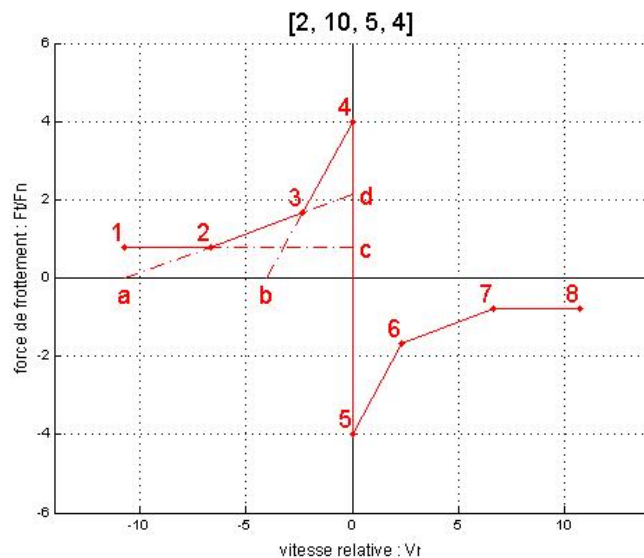


FIG. 1.1: exemple de graphe caractéristique de l'interaction frottée

En ordonnée, la force de frottement est définie comme le rapport de la force tangentielle sur la force normale, forces émises par l'archet sur la corde au point de contact. En abscisses, la vitesse relative représente la différence de vitesse tangentielle entre l'archet et la corde au point de contact.

Toutefois, sous **Modalys**, les caractéristiques non linéaires sont linéarisées en 3 segments. Considérons un segment de cette caractéristique, porté par la droite d'équation

$$(1.1) \quad v = \alpha f + b$$

La variation linéaire d'un élément vibrant étant donné par

$$(1.2) \quad v = v_h + Yf$$

avec v_h la vitesse en l'absence d'interaction. On obtient alors l'équation résolue par **Modalys** :

$$(1.3) \quad f = Zv_s$$

avec

$$(1.4) \quad v_s = \beta - v_h$$

$$(1.5) \quad Z = (Y - \alpha)^{-1}$$

Les paramètres contrôlés sont alors α et β , eux-mêmes fonctions des quatre paramètres

$$[r_0, r_1, r_2, r_3]$$

dits *Rosin parameters*, et qui définissent les caractéristiques de la colophane posée sur le crin de l'archet. Les relations entre ces paramètres et les points de contact de la courbe sont :

$$\begin{aligned} X_1 &= -r_0 \frac{r_1+6}{3} & Y_1 &= \frac{r_3}{r_2} \\ X_2 &= -r_0 \frac{r_1+6}{3} - \frac{r_1}{r_2} & Y_2 &= \frac{r_3}{r_2} \\ X_3 &= \frac{12r_0-4r_0r_1}{3*r_1-6} & Y_3 &= r_3 \frac{r_1}{3r_1-6} \\ X_4 &= 0 & Y_4 &= r_3 \end{aligned}$$

$$\begin{aligned} X_a &= -r_0 \frac{r_1+6}{3} & Y_a &= 0 \\ X_b &= -2r_0 & Y_b &= 0 \\ X_c &= 0 & Y_c &= \frac{r_3}{r_2} \\ X_d &= 0 & Y_d &= r_3 \frac{r_1+6}{3r_1} \end{aligned}$$

Par conséquent, on peut affirmer que l'interaction frottée est intégralement contrôlable à partir de ces *Rosin parameters*.

On remarquera également que ce graphe est symétrique par rapport à l'origine. Cette symétrie est cohérente, puisqu'elle ne fait que symboliser le sens de frottement de l'archet sur la corde (poussé/tiré).

1.3.2 Présentation de l'interaction frappée comme un cas limite de l'interaction frottée

De plus, on constate que sous **Modalys**, l'interaction frappée est également modélisée selon le même processus. En effet, l'interaction frappée présente le

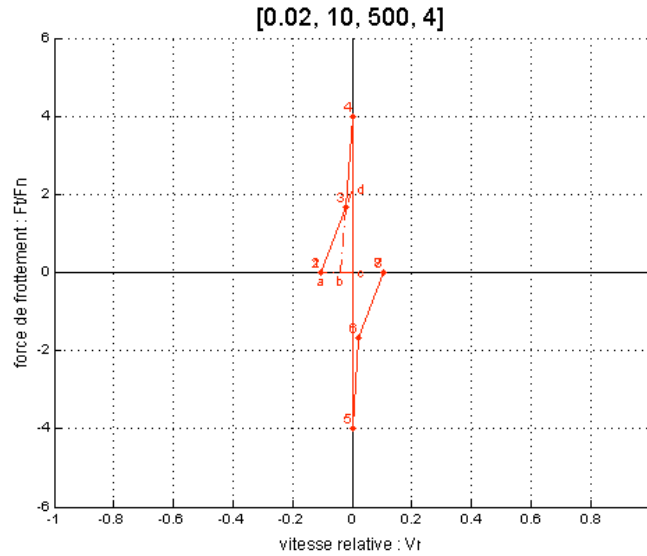


FIG. 1.2: interaction frappée, selon une modélisation “interaction frottée”

même graphe, pour lequel la force d’interaction “contact” est strictement positive pour une vitesse relative nulle, et nulle sinon.

$$\begin{cases} f_c > 0 & \text{si } V_r = 0 \\ f_c = 0 & \text{sinon} \end{cases}$$

Par conséquent, on peut donc définir l’interaction frappée comme un cas limite d’interaction frottée, avec les conditions suivantes :

$$(1.6) \quad r_0 \sim 0$$

$$(1.7) \quad r_2 \rightarrow \infty$$

Du constat précédent, on en vient naturellement à se demander si la similarité mécanique de la modélisation des deux interactions pourrait se retrouver au niveau perceptif. Cette question est d’autant plus intéressante à poser que cette modélisation dite “contact” tend à devenir le modèle d’interaction unique utilisé sous **Modalys**.

Autrement dit, *peut-on considérer qu’il y a une continuité dans l’espace perceptif engendré par l’interaction “contact”, regroupant les interactions frappée et frottée ?*

Chapitre 2

Présentation des outils

2.1 Création du son : Modalys

Modalys est un logiciel de synthèse sonore par modélisation physique. Il permet de définir séparément l'excitateur (ici, l'archet) et le résonateur (ici, la corde). C'est un logiciel qui possède déjà plusieurs interfaces, dont une interface de contrôle pouvant prendre la forme d'un programme en Lisp. C'est pour ce logiciel que nous cherchons à créer une interface de contrôle perceptive.

Pour de plus amples informations concernant ce logiciel, il faut se référer aux annexes B et C.

2.2 Descripteurs acoustiques

2.2.1 Analyse acoustique des sons : Matlab

Nous utilisons pour décrire nos sons d'un point de vue acoustique, deux bibliothèques sous Matlab créées à l'IRCAM et regroupant plusieurs descripteurs acoustiques. L'une a été créée par Geoffroy Peeters [Pee04] et l'autre par Jeremy Marozeau [Mar04] (fondée sur le travail de Slaney).

La bibliothèque de Peeters comprend les descripteurs psychoacoustiques classiques avec quelques variantes (centre de gravité spectral avec une amplitude en dBB par exemple). Celle de Marozeau rajoute des descripteurs plus récents et dont plusieurs prennent en compte le système auditif.

Nous avons regroupé les descripteurs par types dans un tableau (voir annexe A pour description plus complète).

Définition	Nom du descripteur
<i>Enveloppe temporelle du signal</i>	
Logarithme du temps de durée	lat
Centroïde temporel	cgt

Déviation standard temporelle	stdt
Durée effective	ed
Valeur maximum	maximum
ed*cgt	mix
Logarithme du temps d'attaque (Marozeau)	lat2
Impulsivité	impuls
Impulsivité 2	impuls 2
<i>Centroïdes spectraux</i>	
Centre de gravité spectral du spectre de puissance	CGS
Centre de gravité spectral du spectre de puissance (Amplitude en dBA)	CGSA
Centre de gravité spectral du spectre de puissance (Amplitude en dBB)	CGSB
Centre de gravité spectral du spectre de puissance (Amplitude en dBC)	CGSC
Centre de gravité spectral (Amplitude pondérée par la sonie de Moore et Glasberg, fréquence en hz)	schz
Centre de gravité spectral (Amplitude pondérée par la sonie de Moore et Glasberg, fréquence en erb)	scerb
Centre de gravité spectral (Amplitude pondérée par la sonie de Moore et Glasberg, fréquence en erb mais convertie en hz)	scerb
Centre de gravité spectral (Amplitude pondérée par la sonie de Zwicker et Fastl, fréquence en hz)	schznl
<i>Étalement spectral</i>	
Déviation standard spectrale du spectre de puissance	STD
Déviation standard spectrale du spectre de puissance (Amplitude en dBA)	STDA
Déviation standard spectrale du spectre de puissance (Amplitude en dBB)	STDB
Déviation standard spectrale du spectre de puissance (Amplitude en dBC)	STDC
Étalement spectral (Amplitude pondérée par la sonie de Moore et Glasberg, fréquence en erb)	spreaderb
Étalement spectral (Amplitude pondérée par la sonie de Moore et Glasberg, fréquence en erb convertie en hz)	spreadhz
Étalement spectral (Amplitude pondérée par la sonie de Zwicker et Fastl, fréquence en erb)	spreaderbnl
Étalement spectral (Amplitude pondérée par la sonie de Zwicker et Fastl, fréquence en erb convertie en hz)	spreadhznl
<i>Régularité spectrale</i>	
Régularité spectrale	richirr

Régularité spectrale (Prise en compte du système auditif)	richstderb
Régularité spectrale (Prise en compte du système auditif)	richdevherb
Forme du spectre	
Skewness du spectre de puissance	skew
Kurtosis du spectre de puissance	kurt
Slope du spectre de puissance	slope
Niveau sonore	
Niveau sonore	LdB
Niveau sonore (Amplitude en dBA)	LdBA
Niveau sonore (Amplitude en dBB)	LdBB
Niveau sonore (Amplitude en dBC)	LdBC

2.2.2 Analyse statistique des descriptions acoustiques : R

L'analyse par composantes principales (ACP) permet de visualiser quels sont les descripteurs acoustiques qui ont la plus forte variance.

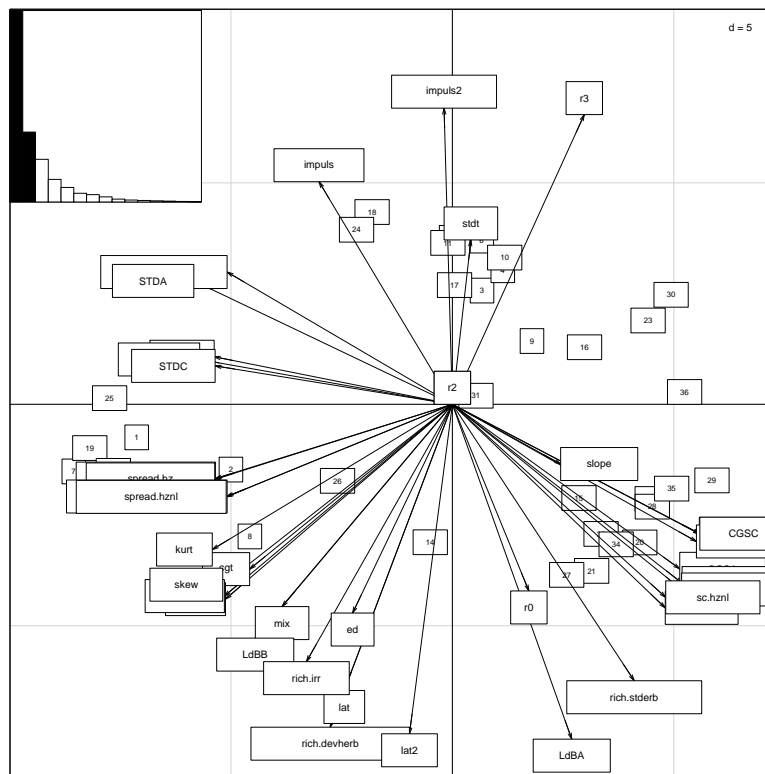


FIG. 2.1: exemple d'ACP

Pour lire une ACP, il faut savoir :

- que les axes représentent une combinaison de descripteurs dont la variance est la plus grande, le premier axe ayant une plus grande variance que le second
- que plus deux variables sont proches (d’angle petit), plus leur corrélation est grande
- que deux variables à angle droit (centré sur l’origine) ont une corrélation nulle
- qu’une variable proche de l’origine a une dispersion faible, voire inexistante si elle se situe au centre

2.2.3 Mise en place du test de dissemblance : **PsiExp**

PsiExp est un logiciel dont l’environnement est conçu pour la constitution et la réalisation d’expériences en psychoacoustique. Il offre des outils adaptés à la mise en place du protocole expérimental et des interfaces graphiques correspondantes et la disposition des données issues des tests.

PsiExp peut être utilisé pour différents types de tests d’écoute. Il permet par exemple l’analyse de la perception de sons et de bruits industriels (véhicules, machines, installations, salles, etc...). L’analyse réalisée avec **PsiExp** permet d’associer les caractéristiques objectives du son aux données perçues, afin par exemple de créer une interface qui contrôle directement les facteurs perceptifs.

Le logiciel permet de concevoir et de présenter sous la forme d’une interface spécifique un protocole d’expérience comprenant l’heure, le nom du sujet, les sons émis, les réponses fournies par le sujet et la valeur des paramètres mesurés susceptibles d’évoluer au cours de l’expérience. **PsiExp** permet ainsi de gérer des stimuli plus complexes que les sons simples utilisés traditionnellement en psychoacoustique. La modification des paramètres d’expérimentation est gérée automatiquement par **PsiExp**. Il peut générer différentes interfaces graphiques présentées au sujet pour l’expérimentation, lui permettant d’opérer des choix, des réglages de paramètres et de lancer l’écoute.

Chapitre 3

Influences relatives des paramètres de colophane

3.1 Pas de continuité dans l'espace de l'interaction "contact"

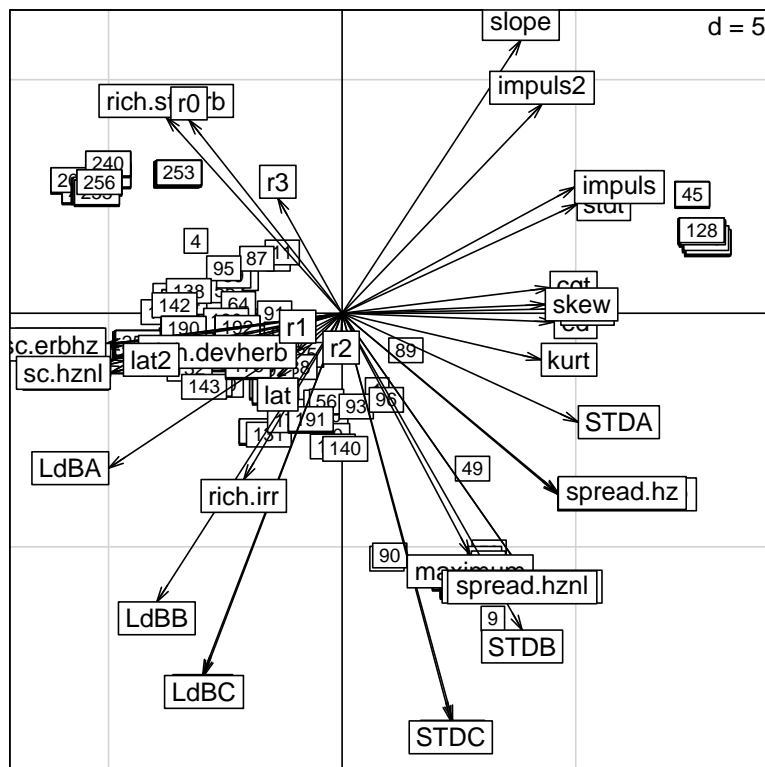


FIG. 3.1: exemple d'ACP de l'interaction "contact"

Selon l'analyse statistique faite de différents corpus composés de sons frappés

et frottés, il apparaît nettement sur la figure 3.1 que ces deux catégories ne permettent pas de définir un espace perceptif continu. En effet, la figure est clairement divisé en deux verticalement, selon le premier axe de variance, et les deux catégories ainsi définies recouvrent parfaitement les deux catégories de sons.

À l’écoute, cette catégorisation des deux composantes de l’interaction “contact” est cohérente. En effet, il est très clair que la perception auditive d’un son très court et sec et d’un son entretenu ne permet pas de les juger comme deux sons de mêmes caractéristiques. La dissimilitude réside tout particulièrement dans l’attaque du son, ainsi que le démontre le fait que le premier axe du graphe, de variance relativement élevé, est très bien corrélé avec les descripteurs acoustiques liés à l’impulsivité.

Ce résultat a pour conséquence de fausser complètement l’étude telle que nous l’avons définie, et donc de dépasser la problématique que nous avons énoncée. En effet, un jugement de dissemblance sur des sons que l’on peut catégoriser différemment pourrait biaiser notre étude. C’est pourquoi nous poursuivrons cette étude en nous limitant aux sons de type “interaction frottée”.

Nous chercherons donc à obtenir une relation mathématique entre des descripteurs reflétant la perception et nos *Rosin parameters*, dans une plage de variation de ceux-ci représentative de l’interaction frottée.

3.2 Influence de r_1 et r_2 négligeable

3.2.1 Influence de r_1 quasi nulle

Le corpus ici étudié est constitué de 256 sons. Nous faisons varier les quatre *Rosin parameters* sur quatre valeurs (voir Annexe B), les sons résultant étant la combinaison de ces valeurs. La figure 3.2 décrivant l’analyse acoustique de ce corpus est explicite : le paramètre r_1 est très proche de l’origine du graphe, ce qui signifie que son influence sur l’aspect perceptif du son est extrêmement faible. Par conséquent, nous pouvons décider de le considérer comme une constante du problème. Nous limiterons de cette façon le problème à trois variables : $[r_0, r_2, r_3]$.

3.2.2 Influence de r_2 négligeable

En ce qui concerne le second paramètre le plus faible, r_2 , il apparaît sur la figure 3.2 que son influence n’est pas inexistante, mais que malgré tout elle semble négligeable par rapport à celles de r_0 et r_3 . Pour s’en assurer, nous avons décidé d’étudier trois corpus différents, pour lesquels seuls les deux paramètres $[r_0, r_3]$ sont variables. Pour chacun des trois corpus, nous prenons une valeur différente de r_2 .

Cette expérience consiste à faire une analyse statistique de chacun des trois

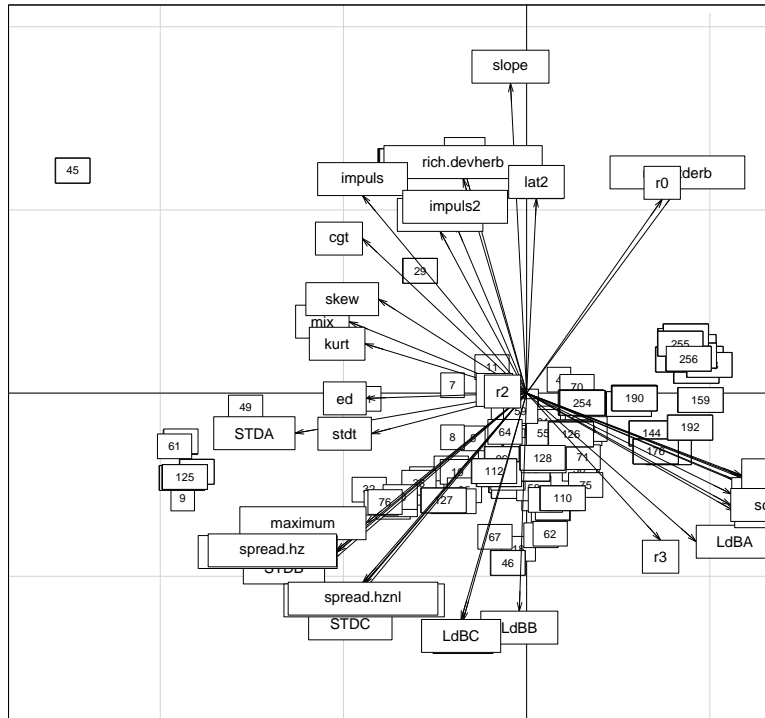


FIG. 3.2: ACP d'un corpus de 256 sons, avec les quatre paramètres variables

corpus, ceux-ci ne différant donc que par leur valeur de r_2 , puis d'étudier la corrélation entre les axes d'ACP des analyses ainsi faites. Si la corrélation est très forte, alors on pourra considérer que les axes ne varient pas (ou très peu) avec r_2 .

TAB. 3.1: étude de l'influence de r_2 sur la perception du son

	$r_2 = 5$	$r_2 = 500$	$r_2 = 50\ 000$
$r_2 = 5$	1 / 1	0.96 / 0.94	0.98 / 0.88
$r_2 = 500$	0.96 / 0.94	1 / 1	0.99 / 0.92
$r_2 = 50\ 000$	0.98 / 0.88	0.99 / 0.92	1 / 1

Le tableau ci dessus donne les corrélations entre les analyses, pour les deux axes. On constate que la corrélation est très forte pour les premiers axes, et plus faibles pour les seconds. Par ailleurs, on a également constaté que les seconds axes ont une variance nettement moins importante que les premiers : le graphe est "écrasé" selon le premier axe. Par conséquent, la corrélation est plus porteuse de sens pour le premier que pour le second.

On peut donc réduire l'étude à l'influence des paramètres $[r_0, r_3]$. On remarquera néanmoins que l'influence de r_1 et r_2 n'est pas nulle, mais qu'elle est simplement négligeable devant celle de r_0 et r_3 .

3.3 Paramètres fixes

Longueur de la corde (m)	length	0.5
Tension de la corde (N)	tension	150
Masse volumique de la corde (kg/m^3)	density	1000
Rayon de la corde (m)	radius	0.001
Module d'Young de la corde (N/m^2)	young	.001
<i>const-loss</i> de la corde	const-loss	1
<i>freq-loss</i> de la corde	freq-loss	1
Masse du "manche" de l'archet (kg)	small-mass	0.1
Masse de la "tête" de l'archet (kg)	large-mass	0.1
Raideur de l'archet (sens du frotté)	stiffness0	50 000
Raideur de l'archet (sens perpendiculaire au frotté)	stiffness1	50 000
<i>const-loss</i> de l'archet (sens du frotté)	const-loss0	0
<i>const-loss</i> de l'archet (sens perpendiculaire au frotté)	const-loss1	0
<i>freq-loss</i> de l'archet (sens du frotté)	freq-loss0	100
<i>freq-loss</i> de l'archet (sens perpendiculaire au frotté)	freq-loss1	100

3.4 Paramètres variables

r_0	2	34	88	135
r_3	4	27	61	74

Nous ne prenons que 4 valeurs par paramètre pour arriver à un total de 16 sons de 2 secondes chacun, une expérience de dissemblance ne pouvant durer trop longtemps (une heure par sujet au maximum). Si l'on compte une trentaine de secondes pour le jugement d'une paire, cela fait 120 paires, soit 16 sons.

3.5 Comparaison des tendances de variation

L'objectif de cette comparaison est double. Tout d'abord, il s'agit de confronter les deux paramètres $[r_0, r_3]$, afin de noter leur indépendance, ainsi que leur poids respectif. Elle a également pour objectif de définir à quelle catégorie de descripteur acoustique est corrélé chaque paramètre.

Concernant le paramètre r_0 , la figure 3.3 montre qu'il a l'air relativement bien corrélé avec *richstderb*, descripteur de la régularité spectrale, et *LdBA*, descripteur de niveau sonore.

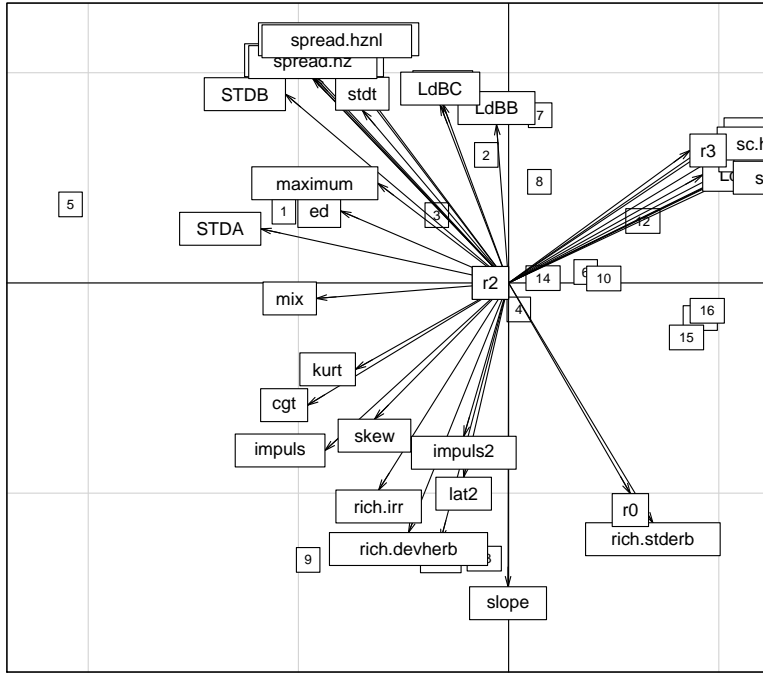


FIG. 3.3: ACP d'un corpus de 16 sons, avec r_0 et r_3 variables

Pour ce qui est du paramètre r_3 , il semble plutôt bien corrélé avec *impuls2*, descripteur d'impulsivité.

Enfin, le paramètre r_3 semble légèrement plus influent que r_0 . Malgré tout, les tendances de variation semblent du même ordre de grandeur.

3.6 Conclusion

Cette partie nous a essentiellement permis de cibler de façon plus précise le sujet de notre étude. En effet, nous avons pu démontrer qu'il était vain de tenter de trouver une continuité à l'espace perceptif engendré par l'interaction "contact". Nous nous sommes donc limités à l'étude de l'interaction frottée.

Nous avons également pu mettre en lumière que deux des paramètres de colophane (r_1 et r_2) étaient d'influence négligeable devant les deux autres (r_0 et r_1). Ce qui nous a permis de réduire notre corpus de sons en maintenant fixes r_1 et r_2 .

En réduisant ainsi le nombre de paramètres qui varient, nous prenons en compte notre contrainte d'un faible nombre de sons pour l'expérience de dissemblance et nous nous assurons d'avoir un résultat fiable. En effet chaque paramètre peut prendre un nombre plus important de valeurs (4 par paramètres dans notre cas au lieu de 2 si 4 paramètres varient), les corrélations entre les dimensions perceptives et les paramètres mécaniques étant de ce fait plus significatives.

Chapitre 4

Tests perceptifs à partir du corpus de sons

4.1 Expérience de dissemblance

Stimuli

Les sons sont numérotés de la façon suivante sur les figures :

Son n°	0000	0001	0002	0003	1000	1001	1002	1003
r_0	2	2	2	2	34	34	34	34
r_1	10	10	10	10	10	10	10	10
r_2	5	5	5	5	5	5	5	5
r_3	4	27	61	74	4	27	61	74
Son n°	2000	2001	2002	2003	3000	3001	3002	3003
r_0	88	88	88	88	135	135	135	135
r_1	10	10	10	10	10	10	10	10
r_2	5	5	5	5	5	5	5	5
r_3	4	27	61	74	4	27	61	74

Nous avons 16 sons, ce qui nous assure de ne pas faire une expérience trop longue où la concentration des sujets serait amoindrie.

4.1.1 Expérience

Il s'agit d'une expérience de dissemblance. Le but est que le sujet exprime sur une échelle continue la dissimilarité qu'il perçoit entre les deux sons d'une paire. Le test se déroule en trois temps :

1. on présente les 16 sons dans un ordre aléatoire au sujet.
2. une phase d'entraînement où 3 paires sont présentées de façon aléatoire et où le sujet est mis dans les mêmes conditions que pendant l'expérience

3. la phase d'expérience demande à chaque sujet de juger de la dissemblance des sons présentés par paires.

Les 120 paires sont présentées une à une dans un ordre aléatoire. Un silence de 0.4 secondes se situe entre les deux sons de chaque paire. Il est demandé au sujet de juger de la dissemblance entre les deux sons sur une échelle allant de [Très semblable] à [Très différent], puis de verbaliser brièvement les critères perçus comme différents. Il est possible de réécouter chaque paire autant de fois que le sujet le souhaite (voir consigne de l'expérience en annexe D).

L'expérience dure de 45 minutes à une heure selon les sujets.

4.1.2 Matériel

Afin d'isoler les sujets des bruits extérieurs, l'expérience a été effectuée dans une cabine insonorisée à double paroi (IAC) à l'IRCAM.

Dans la première cabine, l'écoute des sons s'est faite avec un casque Sennheiser HD 250 II, les niveaux ayant été réglés pour se situer entre 65 et 70 dB à l'aide d'un sonomètre B&K 2238 Mediator. L'ordinateur, un Mac G5-biprocasseur est équipé d'une carte audio interne : RME hdsp9652, de convertisseurs externes RME ADI-8 PRO dont la sortie est amplifiée par un amplificateur casque LAKE. Dans la seconde cabine, l'écoute des sons s'est faite avec un casque Sennheiser HD 250 II, les niveaux ayant été réglés pour se situer entre 65 et 70 dB à l'aide d'un sonomètre B&K 2238 Mediator. L'ordinateur, un Mac G5-biprocasseur est équipé d'une carte audio interne : RME hdsp9652, de convertisseurs externes RME ADI-8 PRO dont la sortie est amplifiée par un amplificateur casque Yamaha.

L'expérience est construite grâce au logiciel Psiexp qui gère l'interface utilisateur, la gestion audio se faisant grâce au logiciel Max/MSP.

4.1.3 Sujets

Un groupe de 19 sujets (14 hommes et 5 femmes) âgés de 18 à 41 ans (en majorité autour de 24 ans) a participé à l'expérience. Tous ont dit avoir une audition normale. Les sujets ont été rémunérés pour leur participation et recrutés grâce à la base de donnée propre à l'équipe et parmi mes connaissances.

4.2 Etude des corrélations entre dissemblance et *Rosin parameters*

4.2.1 Synopsis d'une analyse multidimensionnelle

Une analyse MDS tend à décrire les dissemblances entre les sons δ_{ij} obtenues grâce à l'expérience, par des distances entre des points dans une configuration X

de m dimensions. Un système de coordonnées est associé à cette configuration spatiale, m axes perpendiculaires liés à une origine O . Ces axes sont des dimensions au sens d'une métrique ayant la même unité. Chaque point est ainsi décrit par un unique m -tuple correspondant aux projections du point sur chacun des m axes. Ce m -tuple est donc un vecteur de coordonnées dans un espace cartésien. La distance d_{ij} entre les points dans cet espace peut être calculée, par exemple, grâce à une distance euclidienne.

L'analyse MDS associe les proximités δ_{ij} avec les distances correspondantes $d_{ij}(X)$ d'un espace MDS, X tel que :

$$(4.1) \quad \delta_{ij} \xrightarrow{f} d_{ij}(X)$$

avec f qui spécifie le modèle de MDS. Le but est de faire correspondre les deux entités δ_{ij} et $d_{ij}(X)$. Cette fonction f peut être de type exponentielle ou logarithmique comme dans le cas de la distance généralisée, telle que Shepard [She87] l'a définie. De manière générale, elle décrit une transformation linéaire à variation lente.

A chaque moment de l'analyse, une erreur est calculée entre le modèle $d_{ij}(X)$ et les proximités δ_{ij} . Cette erreur est le *stress* :

$$(4.2) \quad \sigma_r = \sum_{i < j} [d_u(i, j) - \delta(i, j)]^2$$

Le modèle est élaboré en minimisant cette erreur afin de calculer une géométrie optimale. Le lecteur pourra se référer aux livres de Borg et Groenen [BG97] et Schiffman et coll. [SRY81] pour les aspects calculatoires propres à la MDS.

4.2.2 Choix du modèle spatial

Nous avons utilisé le logiciel Xgvis [BS02] dans lequel nous avons choisi un modèle métrique de Kruskal Shepard. Nous avons un très bon stress (0.0751) pour un espace à 2 dimensions, c'est donc celui là que nous choisissons. Pour plus d'informations, se référer à l'Annexe E.

4.2.3 Verbalisations

A chaque sujet était demandé durant le test, sur quels critères reposait jugement de dissemblance entre les sons. Nous projetons ainsi de compléter notre analyse MDS par une analyse des verbalisations des sujets, suivant le protocole utilisé par A. Faure [Fau00].

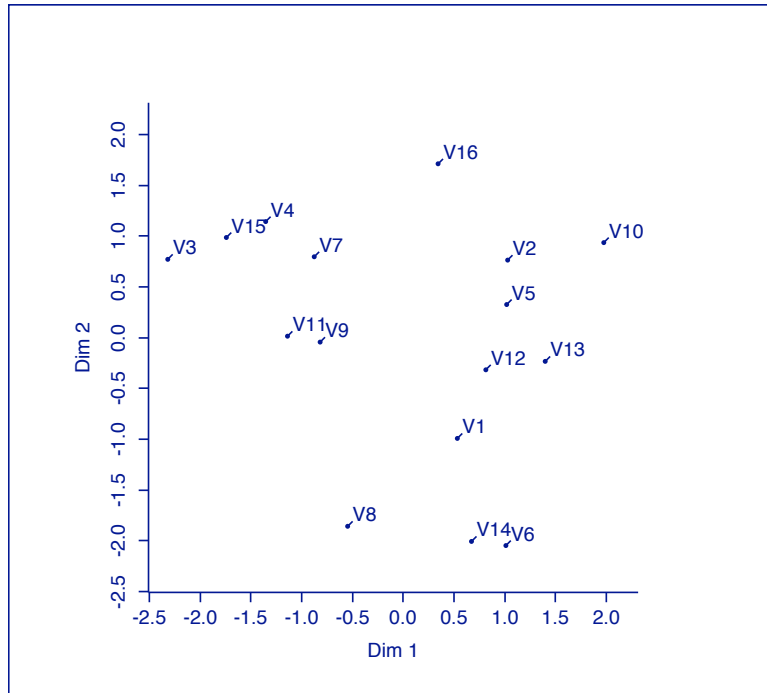


FIG. 4.1: Espace perceptif

4.2.4 L'espace perceptif

La figure 4.1 indique qu'il ne semble pas y avoir de relation directe entre les deux paramètres de colophane et les deux dimensions obtenues ci dessus. L'hypothèse la plus probable est que cette relation soit non linéaire. L'étude des corrélations, à défaut de nous donner cette relation, peut nous permettre *a minima* de confirmer cette hypothèse.

4.3 Etude des corrélations

Nous avons d'abord calculé des régressions multilinéaires sous **Matlab**, entre :

- les dimensions perceptives et les 35 descripteurs que nous avons introduits dans le chapitre précédent, calculés sur les 16 sons que nous avons utilisés dans l'expérience, de façon à pouvoir définir sur quels critères les jugements des sujets se sont reposés
- les dimensions perceptives et les paramètres mécaniques, c'est à dire r_0 et r_3 , de façon à les relier mathématiquement

Les corrélations que nous présentons sont celles définissant si les régressions multilinéaires sont une bonne approximation du paramètre qu'elles cherchent à approcher. La première colonne donne la régression, l'autre sa corrélation avec la dimension perceptive. Plus la corrélation s'approche de 1 plus l'approximation

est bonne à partir du moment où la régression revêt un caractère significatif.

4.3.1 Dimension1

Corrélations avec les paramètres acoustiques

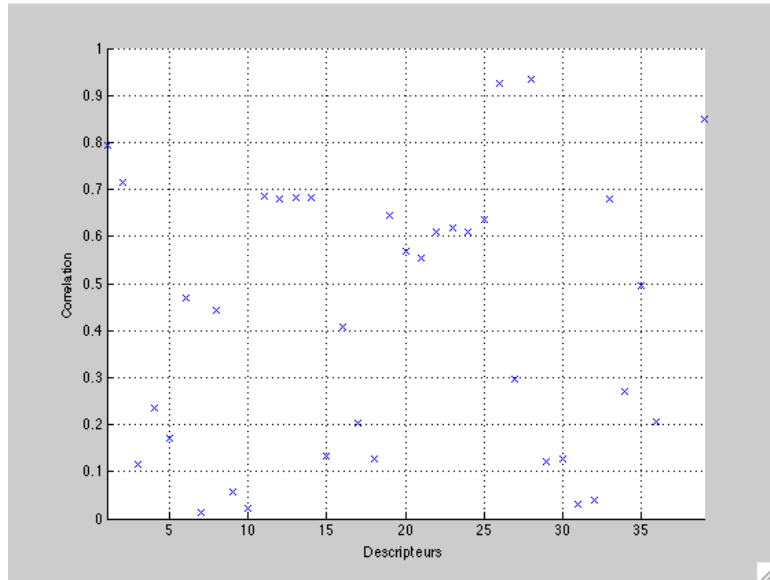


FIG. 4.2: Corrélations entre les descripteurs acoustiques et la dimension 1

On observe une bonne corrélation entre la première dimension perceptive et les descripteurs de régularité spectrale, en particulier *rich.irr*.

Cette proximité est confirmée par la régression linéaire :

$$\boxed{-5.8662 + 0.003 \times rich.devherb} \quad \boxed{0.866} \quad p.value < 0.001$$

(cf. Fig 4.3)

$$\boxed{-6.1239 + 0.9873 \times rich.irr} \quad \boxed{0.8442} \quad p.value < 0.001$$

On observe également une bonne corrélation avec le descripteur temporel *lat*, bien que la significativité soit plus faible :

$$\boxed{2.5812 + 1.9081 \times lat} \quad \boxed{0.8442} \quad p.value = 0.0002080$$

On remarque également une forte influence du CGS, mais il ne semble pas corrélé seul. On peut donc envisager des corrélations multiples. Calculons les différentes combinaisons :

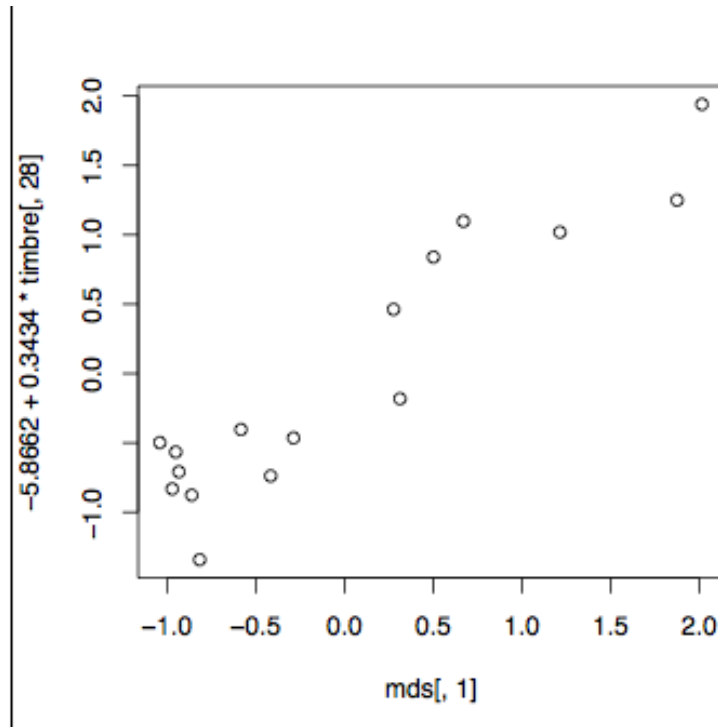


FIG. 4.3: Corrélation entre le descripteur *rich.devherb* et la dimension 1

$$\boxed{-3.7777 + 0.7133 \times lat + 0.7646 \times rich.irr} \quad \boxed{0.8848} \quad p.value < 0.001$$

$$\boxed{-3.5709240 + 0.8397274 \times rich.irr - 0.0011338 \times CGS} \quad \boxed{0.8823}$$

p.value < 0.001

$$\boxed{5.4562549 + 1.5826570 \times lat - 0.0022954 \times CGS} \quad \boxed{0.8564} \quad p.value < 0.001$$

(cf. Fig 4.4)

$$\boxed{-2.7931282 + 0.2888441 \times rich.devherb - 0.0014825 \times CGS} \quad \boxed{0.9536}$$

p.value < 0.001
(cf. Fig 4.5)

La dernière combinaison présente la meilleure corrélation avec une forte significativité. Il s'avère donc que la première dimension perceptible est bien corrélée à des descripteur spectraux (plus particulièrement la richesse spectrale).

Corrélations avec les paramètres mécaniques

$$\boxed{1.264214 - 0.030463 \times r_3} \quad \boxed{0.7011} \quad p.value < 0.001$$

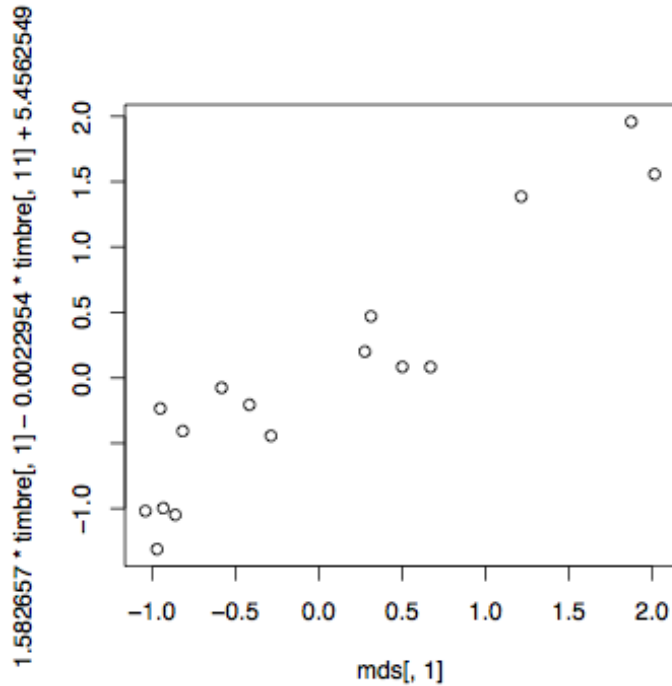


FIG. 4.4: Corrélation entre les descripteurs *lat* et *CGS* et la dimension 1

$0.998368 + 0.004106 \times r_0 - 0.030463 \times r_3$	0.7294	p.value < 0.001
--	--------	-----------------

Le taux de corrélation est relativement bon. Malgré tout, il reste envisageable de considérer qu'une régression non linéaire permettrait de mieux décrire le lien existant entre cette dimension et les deux paramètres de colophane.

r_0 n'étant pas significatif seul, on peut donc considérer cette dimension comme simplement reliée à r_3 . Ainsi, on constate bien que le paramètre r_3 semble nettement plus influent que le paramètre r_0 , comme il était également pressenti à l'écoute.

4.3.2 Dimension 2

Corrélations avec les paramètres acoustiques

On observe une corrélation moyenne entre la deuxième perceptive et le descripteur temporel *lat*, ainsi qu'avec la richesse spectrale :

$-2.5057 - 1.8523 \times lat$	0.5501	p.value < 0.001
-------------------------------	--------	-----------------

$4.35547 - 0.25496 \times rich.devherb$	0.4274	p.value = 0.003591
---	--------	--------------------

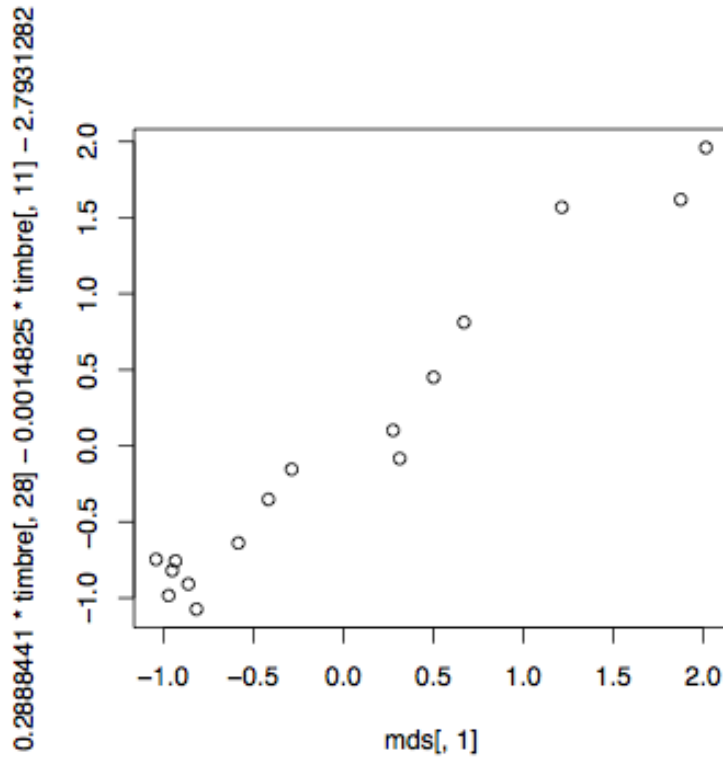


FIG. 4.5: Corrélacion entre les descripteurs *rich.devherb* et *CGS* et la dimension 1

Par conséquent, nous effectuons une régression multilinéaire avec ces deux descripteurs, afin de vérifier si leur combinaison permet une meilleure description :

$$\boxed{-0.85170 + -1.49040 \times lat - 0.06816 \times rich.devherb \mid 0.5283} \text{ p.value} = 0.002985$$

On constate que ce n'est pas le cas. Par conséquent, on peut relier cette dimension perceptible qu'au descripteur temporel *lat*. On retrouve ici les notions temporelles d'attaque, nettement sensibles à l'écoute.

Corrélations avec les paramètres mécaniques

Aucune régression multilinéaire liée aux deux paramètres de colophane ne donne de résultat significatif pour cette dimension. Il est donc nécessaire d'envisager une relation non-linéaire entre cette dimension et les paramètres r_0 et r_3 .

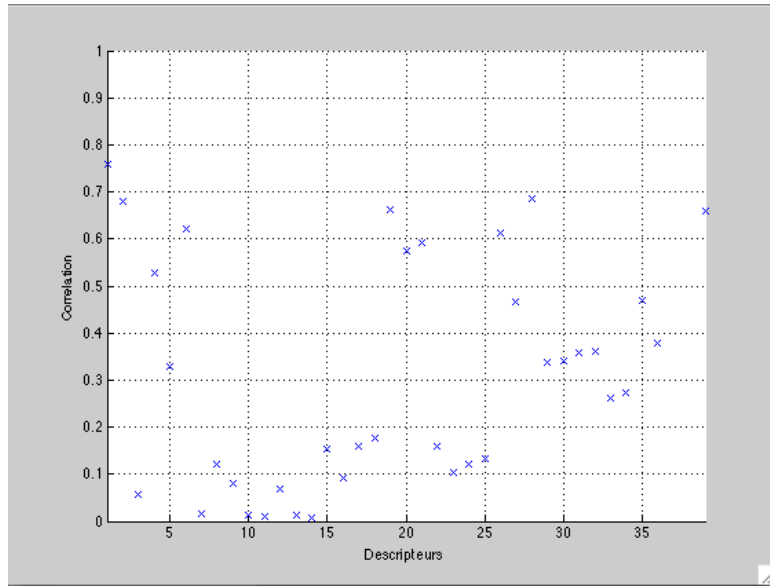


FIG. 4.6: Corrélacion entre les descripteurs acoustiques et la dimension 2

4.4 Conclusion

Tout d'abord, on notera qu'il s'agit ici de résultats intermédiaires. En effet, pour l'instant nous avons fait le test à 19 sujets, 30 permettront d'affiner nos résultats.

Nous avons trouvé deux dimensions à notre espace de sons. Les dimensions correspondent bien à ce qui était pressenti à l'écoute, c'est-à-dire une dimension reflétant une notion temporelle, d'attaque du son, et une autre reflétant une notion spectrale. On remarque que la dimension temporelle est beaucoup moins bien expliquée.

Malgré tout, au niveau mécanique, les corrélations entre les paramètres de colophane et ces dimensions, si elles sont correctes, ne permettent pas d'assurer une justesse suffisante à nos résultats. Il peut être envisagé de rechercher une relation non linéaire entre ceux-ci.

Enfin, cette démarche peut être orientée à l'aide des verbalisations demandées aux sujets. Une analyse sémantique de ces verbalisations pourra permettre de valider ou d'affiner les résultats ainsi obtenus.

Conclusion

Cette étude cherchait à permettre un contrôle perceptif de la synthèse dans Modalys, reposant sur la perception, pour l'interaction frottée. Pour cela, nous avons construit un corpus de sons représentatif de la variété sonore existant dans Modalys pour cette interaction, puis nous avons réalisé une expérience de perception mettant en relation des paramètres perceptifs et les paramètres mécaniques de la plaque. Nous avons pour cela analysé le fonctionnement du modèle physique sous-jacent au logiciel, puis nous avons fait une analyse des contributions acoustiques dues aux variations des paramètres de colophane de l'archet. Cela nous a permis de vérifier qu'il n'y a pas de continuité perceptive entre les interactions frappée et frottée puis de déterminer l'influence respective des quatre paramètres de colophane $[r_0, r_1, r_2, r_3]$. La comparaison des tendances de variation nous a permis également de restreindre notre étude à l'influence des paramètres r_0 et r_3 . Une fois le corpus de sons créé, nous avons alors choisi de faire une expérience de jugement de dissemblance afin de compléter notre connaissance de la perception des sons de corde frottée dans Modalys. A partir des résultats intermédiaires que nous avons, nous constatons que l'espace perceptif résultant de cette expérience possède deux dimensions, l'une à caractère spectral et l'autre à caractère temporel, et pour lesquelles nous avons des relations multilinéaires avec les paramètres r_0 et r_3 .

Néanmoins, la justesse des relations linéaires trouvées entre ces paramètres et les dimensions perceptives étant limitée, on peut envisager, à l'aide de régressions non linéaires, de trouver une relation, non linéaire cette fois-ci, qui décrirait plus précisément les liens entre les dimensions perceptives et les paramètres de colophane.

Nous pouvons maintenant envisager la poursuite de cette étude selon plusieurs axes. Il reste à analyser les verbalisations des sujets, afin de pouvoir définir sémantiquement les deux dimensions obtenues, et vérifier la cohérence entre les résultats obtenus par analyse sémiologique des verbalisations et ceux des régressions multilinéaires. Puis nous pourrions faire une première implémentation temps réel dans Max/MSP, afin de vérifier à l'écoute le bon fonctionnement de cette nouvelle interface perceptive de contrôle.

Le plus difficile restera alors à la charge du compositeur, qui devra imaginer la meilleure façon de créer des sons à l'aide de ce nouvel outil.

Bibliographie

- [Adr88] J. M. Adrien. PhD thesis, Ircam, 1988.
- [BG97] I Borg and P. Groenen. *Modern Multidimensional Scaling. Theory and Applications*. Springer-Verlag, New York, 1^{re} edition, 1997.
- [BS02] A. Buja and S.F. Swayne. Visualization methodology for multidimensional scaling. *Journal of Classification*, 19 :7–43, 2002.
- [Der05] M. Derio. Approche perceptive dans un logiciel de synthèse sonore par modélisation physique : Modalys. Master’s thesis, Université du Maine, Le Mans, 2005.
- [Fau00] A. Faure. *Des sons aux mots, comment parle-t-on du timbre musical ?* PhD thesis, Ecole des Hautes Etudes en Sciences Sociales, 2000.
- [Ma96] J. Morrisson and al. Modalys reference. Technical report, Ircam, 1996.
- [Ma97a] J. Morrisson and al. Modalys introduction. Technical report, Ircam, 1997.
- [Ma97b] J. Morrisson and al. Modalys tutorial. Technical report, Ircam, 1997.
- [Mar04] J. Marozeau. *L’effet de la fréquence fondamentale sur le timbre*. PhD thesis, Ircam, CNMAT, LMA, Université Pierre et Marie Curie - Paris VI, 2004.
- [Pee04] G. Peeters. Preliminary observations on material recovery from real impact sounds. Cuidado i.s.t. project report, Ircam, 2004.
- [She87] R.N. Shepard. Toward a universal law of generalization for psychological science. *Science*, 237 :1317–1323, 1987.
- [SRY81] S.S. Schiffman, M.L. Reynolds, and F.W. Young. *Introduction to Multidimensional Scaling. Theory, Methods and Applications*. Academic Press, Inc., 1981.

Annexe A : Descripteurs Acoustiques

Plus de précisions pourront être trouvées dans Peeters (2004) et Marozeau (2003, 2004).

Enveloppe temporelle du signal

Logarithme du temps de montée (lat)

Logarithme de la durée entre le moment où le signal commence et le moment où il atteint sa partie stable.

Centroïde temporel (cgt)

Il s'agit du centre temporel moyen de l'enveloppe d'énergie. On le calcule de la façon suivante si $e(t)$ est l'enveloppe d'énergie fonction du temps :

$$(4.3) \quad cgt = \frac{\sum_t e(t) \times t}{\sum_t e(t)}$$

Déviatoin standard temporelle (stdt)

Le principe est le même que pour le cgt, on calcule la déviatoin standard au lieu de calculer la moyenne

Durée effective (ed)

Il s'agit d'une mesure du temps pendant lequel le signal est perceptible significativement. Cette mesure est faite en calculant le temps durant lequel l'énergie du signal est au dessus d'une frontière de 40 % de son maximum

Valeur maximum (max)

Calcul de la valeur maximale du signal

mix

Mix entre la durée effective du signal et son centroïde temporelle.

Logarithme du temps d'attaque 2 (lat 2)

Logarithme de la durée entre le moment où elle dépasse 2 % du niveau maximum du signal et le moment où l'enveloppe temporelle atteint sa valeur maximale.

Impulsivité (impuls)

Durée pendant laquelle l'enveloppe temporelle de la puissance du signal est au dessus d'un seuil de 40% divisé par la durée durant laquelle cette enveloppe est supérieure au seuil de 10%.

Impulsivité 2 (impuls 2)

Durée pendant laquelle l'enveloppe temporelle de la puissance du signal est au dessus d'un seuil de 40%.

Centres de gravité spectraux

Centre de gravité spectral classique : (CGS, CGSA, CGSB, CGSC)

$$(4.4) \quad CGS = \frac{\sum_i f_i \times A_i}{\sum_i A_i}$$

où f_i représente la fréquence de la $i^{\text{ème}}$ raie d'un spectre, A_i l'amplitude de la $i^{\text{ème}}$ raie. Il existe des variantes dans la façon de calculer ce CGS, notamment dans la façon de calculer l'amplitude (respectivement dBA, dBB, dBC).

Centre de gravité spectral prenant en compte le système auditif (nommé centroïde spectral par Marozeau) (schz, scerb, scerbhz, schznl)

Le signal est d'abord filtré de façon à modéliser le filtrage effectué par l'oreille externe et moyenne. Le signal est ensuite découpé en bandes fréquentielles pour modéliser la sélectivité fréquentielle de l'oreille. On obtient ainsi le centroïde spectral instantané :

$$(4.5) \quad SC(t) = \frac{\sum_z z \times \psi_z(t)}{\sum_z \psi_z(t)}$$

où $\psi_z(t)$ représente la sonie partielle de la bande z à l'instant t . Enfin le centroïde spectral instantané est pondéré par la sonie instantanée (somme à travers toutes les bandes des sonies partielles) et moyenné afin d'obtenir un descripteur unique caractérisant le signal entier. Des variantes existent suivant la façon dont la sonie est calculée (modèle de Zwicker et Fastl ou de Moore et Glasberg).

Étalements spectraux

Déviations standard (STD, STDA, STDB, STDC)

$$(4.6) \quad STD = \sqrt{\frac{\sum_i (f_i - CGS)^2 \times A_i}{\sum_i A_i}}$$

Il existe des variantes sur le calcul de l'amplitude comme pour le CGS.

Déviations standard prenant en compte le système auditif (spreaderb, spreadhz, spreaderbnl, spreadhznl)

$$(4.7) \quad ES = \sqrt{\frac{\sum_z (z - SC)^2 \times \psi_z(t)}{\sum_z \psi_z(t)}}$$

Le principe est le même ici que pour le centroïde spectral.

Régularité spectrale

Ce type de descripteur permet de distinguer les sons ayant une enveloppe spectrale régulière des sons dont l'enveloppe spectrale est « accidentée ». La sensation associée est nommée « richesse ».

Régularité spectrale classique (« Déviation Spectrale » de Krimphoff)

$$(4.8) \quad Reg = \sum_k \left| 20 \log Ak - \frac{1}{3} (20 \log (Ak - 1) + 20 \log Ak + 20 \log (Ak + 1)) \right|$$

Régularité spectrale prenant en compte le système auditif

$$(4.9) \quad Reg2 = \sum_z \left| \psi_z - \frac{1}{3}(\psi_z - 1 + \psi_z + \psi_z + 1) \right|$$

Le principe est le même que pour le CGS prenant en compte le système auditif.

Forme du spectre (skew, kurt, slope)

Skewness

Le skewness décrit le degré d'asymétrie du spectre de puissance par rapport à sa valeur moyenne.

Kurtosis

Le kurtosis donne une mesure du caractère plus ou moins plat du spectre de puissance.

Slope

Le slope est une mesure de la décroissance de l'amplitude. Elle est calculée par régression linéaire de l'amplitude spectrale.

Niveaux sonores (LdB, LdBA, LdBB, LdBC)

Les niveaux sonores se calculent de la façon suivante :

$$(4.10) \quad LdB = 20 \log \frac{rms}{2 \times 10^{-5}}$$

où rms est la valeur efficace du spectre de puissance.

Différentes pondérations peuvent être ensuite appliquées pour prendre en compte la sensibilité de l'oreille, le signal sera alors obtenu en dBA, dBB, ou dBC.

Annexe B : Programme Modalys

```
(new)
```

```
;;;  
;;; Objets  
;;;
```

```
(setq mat (list  
           (list 2 34 88 135)  
           (list 10)  
           (list 5)  
           (list 4 27 61 74 )))
```

```
(loop  
  for i from 0 to 3  
  do
```

```
  (loop  
    for j from 0 to 0  
    do
```

```
    (loop  
      for k from 0 to 0  
      do
```

```
      (loop  
        for l from 0 to 3  
        do
```

```

(new)

(print (list i j k l))

(setf k0
      (nth i (nth 0 mat)))
(setf k1
      (nth j (nth 1 mat)))
(setf k2
      (nth k (nth 2 mat)))
(setf k3
      (nth l (nth 3 mat)))

(print (list k0 k1 k2 k3))

(setq my-string
      (make-object 'bi-string
                  (modes 40)
                  (length .5)
                  (tension 150)
                  (density 1000)
                  (radius .001)
                  (young .001)
                  (freq-loss 1)
                  (const-loss 1)))

(setq my-bow
      (make-object 'bi-two-mass
                  (small-mass .05)
                  (large-mass .05)
                  (stiffness0 50000)
                  (stiffness1 50000)
                  (freq-loss0 100)
                  (freq-loss1 100)
                  (const-loss0 0)
                  (const-loss1 0)))

;;;
;;; make bowed connection
;;;

```

```

(setq my-bow-h-bpt
      (make-access my-bow 1 'trans0))
(setq my-bow-v-bpt
      (make-access my-bow 1 'trans1))
(setq my-string-h-bpt
      (make-access my-string .1 'trans0))
(setq my-string-v-bpt
      (make-access my-string .1 'trans1))

(setq bc
      (make-connection 'bow
                       my-bow-v-bpt my-bow-h-bpt .01
                       my-string-v-bpt my-string-h-bpt 0
                       (list k0 k1 k2 k3)))

;;;
;;; push the bow
;;;

(setq my-bow-h-mov
      (make-access my-bow
                   (const 0) 'trans0))
(setq my-bow-v-mov
      (make-access my-bow
                   (const 0) 'trans1))

(make-connection 'speed my-bow-h-mov
                 (make-controller 'envelope 1
                                  (list (list 0.00 1)
                                        (list 0.10 4)
                                        (list 9.99 1))))

(make-connection 'position my-bow-v-mov
                 (make-controller 'envelope 1
                                  (list (list 0.0 0.010)
                                        (list 0.2 -0.001)
                                        (list 0.4 -0.001)
                                        (list 0.7 -0.001)
                                        (list 1.2 0.010)
                                        (list 9.9 0.010))))

;;;
;;; Ecoute
;;;

```

```

(make-point-output
  (make-access my-string
    (const .6) 'trans0))

(run 2)
(with-open-file
  (stream
    "Perception&Design Sonore-1304:Users:dossantos:Desktop:Modalys:Rosin.txt"
    :direction :output :if-does-not-exist :create
    :if-exists :append)
    (format stream "~& ~S ~S ~S ~S ~%"
      k0
      k1
      k2
      k3
      (code-char 13)
      (code-char 10))

    (save-aiff (format Nil
      "/Users/dossantos/Desktop/Modalys/Sons/Rosin~S~S~S~S.aif" i j k l)))

  (with-open-file
    (stream
      "Perception&Design Sonore-1304:Users:dossantos:Desktop:Matlab:rosin.txt"
      :direction :output :if-does-not-exist :create :if-exists :append)
      (format stream "~& ~S ~S ~S ~S ~%"
        i
        j
        k
        l
        (code-char 13)
        (code-char 10))))))

;;cat a.txt | tr '\r' '\n

```

Annexe C : Fonctionnement de Modalys

Ne seront abordés ici que les parts de **Modalys** qui concernent directement notre étude, à partir des travaux de M. Derio [Der05].

Contrôle

Nous allons expliquer ici comment fonctionne **Modalys** du point de vue de l'utilisateur, quels sont les contrôles à sa disposition ([Ma96], [Ma97a], [Ma97b]) sous la forme d'un programme écrit en LISP (voir annexe B).

Objets

L'idée de **Modalys** est d'assembler divers éléments pour construire des instruments plus complexes. Ces éléments, les objets dans le formalisme **Modalys**, peuvent être par exemple des cordes, des membranes, des tubes ou des plaques dans le cas qui nous intéresse. Un objet comme le "bi-two-mass" peut servir d'excitation quand il interagit avec les autres objets : en tant que médiateur, pour modéliser le pincement de corde avec un doigt, en tant qu'archet, etc...

Deux remarques :

- Les objets dans **Modalys** sont des modèles linéaires, c'est-à-dire qu'ils ne peuvent pas être cassés par exemple.
- Les objets ne peuvent vibrer que sur un ou deux axes prédéterminés.

Accès

Pour pouvoir définir une interaction entre objets ou écouter le son produit celle-ci, il faut définir des accès sur les objets. Ces accès définissent une position sur l'objet mais également une direction. Si l'on prend l'exemple d'une corde frottée, il faut définir un accès sur la corde qui donne la position de l'interaction, et une direction (normale ou transversale) pour définir dans ce cas si la corde est frappée ou plutôt frottée. Il faut définir le même genre d'accès sur l'archet. Enfin, si l'on veut écouter le son produit par l'événement, il faut placer un haut-parleur

sur un point de la corde qu'on a défini par un accès. Le haut-parleur mesure la vibration en ce point et le transforme en son. La direction est importante, puisqu'elle définit si on choisit la vibration normale ou transversale.

Connexions

Une fois que les accès sont choisis, on peut définir une action entre nos accès. Dans le cas de notre corde frottée, on prendra une connexion “bow”, qui décrit comment les objets vont rentrer en contact. L'objet excitateur, le “bi-two-mass” pouvant modéliser aussi bien un archet qu'un médiator, il est important de lui donner une action qui correspond, dans le cas qui nous occupe, à un frotté plutôt que par exemple à un pincement qui ne rendrait pas du tout le même son. Il ne suffit donc pas de faire rencontrer les objets, il faut également définir de quelle manière ils se rencontrent.

Controleurs

Quand l'action, la “connexion” est définie, nous pouvons la contrôler. Dans le cas de notre frottement de corde, on peut chercher à contrôler la position de l'archet au cours du temps et par exemple le faire entrer en contact plusieurs fois de suite avec la corde.

Temps différé, temps réel

Dans la version d'origine de **Modalys**, le contrôle de l'interaction entre les objets est défini au préalable, le son est ensuite synthétisé. Il existe aujourd'hui des versions de **Modalys** en temps réel.

Modèle physique

Le modèle physique de **Modalys** repose sur les concepts développés dans Adrien [Adr88]. La documentation n'étant pas encore disponible sur ce point, nous synthétisons à partir de la thèse ce que nous pensons pouvoir servir à la compréhension du fonctionnement d'une corde frottée dans **Modalys**. Nous nous reposons également pour la compréhension sur Drouin et coll. (1993). Nous ne présentons donc que la corde et le câble.

Corde

Sans amortissement

La corde au départ est modélisée par un ensemble de masses ressorts qui forment un chapelet. On obtient une équation de la forme :

$$(4.11) \quad [M] \frac{d^2}{dt^2}(y) + [K](y) = (Fext)$$

Pour obtenir les déformées modales du système discret (le chapelet constitué des masses ressorts), le modèle traduit les déformées modales de la corde continue dans la base des coordonnées généralisés associée au chapelet : un mode propre est défini par la donnée des N valeurs de ce mode aux points discrets du chapelet. On obtient ensuite les équations de N oscillateurs harmoniques indépendants comme nous l'avons décrit dans la partie sur les oscillateurs à degrés de liberté multiples.

On associe pour finir les fréquences harmoniques de la corde continue à la base des vecteurs propres du modèle discret. Il existe une erreur sur cette association, mais il semble que l'oreille tolère, d'après Adrien (1988), un certain degré d'approximation sur les qualités spatiales et que les propriétés fréquentielles soient les plus importantes du point de vue perceptif.

Avec amortissement

On introduit les efforts de frottement avec la matrice [B] :

$$(4.12) \quad [M] \frac{d^2}{dt^2}(y) + [B] \frac{d}{dt}(y) + [K](y) = (Fext)$$

D'après Adrien (1988), cette matrice peut rendre compte des efforts d'amortissement visqueux et fluides, ce qui conduit à exprimer [B] comme combinaison linéaire de [M] et [K]. Nous avons donc un amortissement proportionnel et ainsi que nous l'avons évoqué dans la partie sur les oscillateurs, l'introduction des éléments dissipatifs de cet ordre n'affecte pas la base des vecteurs propres (les déformées modales). On parle alors d'amortissement modal.

On a deux paramètres de contrôle b_1 et b_2 qui se rapportent respectivement à des pertes constantes et proportionnelles au carré de la fréquence de chaque mode. Les forces fluides se rapportant plus précisément à b_1 et les forces visqueuses à b_2 .

Câble

Pour donner une idée de l'étendue du modèle de câble, on peut dire que les comportements du câble pour les faibles ou grandes raideurs de flexion tendent vers les comportements des cordes idéales ou des barres non tendues.

Sans amortissement

Si l'on ajoute aux équations de la corde idéale les efforts transversaux résultants de la variation du moment de flexion le long du câble, on obtient le système d'équations régissant le mouvement d'une corde discrétisée pourvue d'une raideur de flexion, ou d'un câble.

Le système est modélisé simplement ici par une succession de barrettes avec des moments exercés dus à la raideur de flexion du câble.

$$(4.13) \quad [M] \frac{d^2}{dt^2}(y) + [[K] + [R]](y) = (Fext)$$

De la même façon que pour la corde idéale, les modes propres discrets utilisés sont les expressions discrétisées correspondant aux solutions du problème continu.

On remarque que le moment de flexion auquel est soumis la section droite séparant les barrettes i et $i+1$ est fonction du module d'Young notamment. Par suite $[R]$ dépend du module d'Young.

On utilise également dans ce modèle, pour les déformées modales, les expressions discrétisées correspondant aux solutions du problème continu. On utilise les déformées modales d'un câble avec des conditions aux limites usuelles : le câble scellé.

On remarque que les déformées modales dépendent du module d'Young et également de la force de rappel, donc de la raideur des ressorts.

On obtient ensuite encore une fois des équations indépendantes pour chaque oscillateur. Il est à noter que les relations d'orthogonalité nécessaires au découplage des équations ne sont pas évidentes, les déformées modales utilisées n'étant pas celle du problème discret. À chaque fois ces relations sont vérifiées, analytiquement pour la corde, de façon numérique pour le câble.

Avec amortissement

$$(4.14) \quad [M] \frac{d^2}{dt^2}(y) + [B] \frac{d}{dt}(y) + [[K] + [R]](y) = (Fext)$$

avec

$$(4.15) \quad [B] = b_1[M] + b_2([K] + [R])$$

Formellement on obtient pour l'amortissement :

$$(4.16) \quad b_1 + b_2 4\pi^2 f^2$$

Ainsi, d'après Adrien (1988), b_1 peut se rapporter à un amortissement fluide et b_2 à des pertes internes (puisqu'ils intéressent respectivement des phénomènes de dissipation constants et variants avec le carré de la fréquence).

Modèle d'archet : le bi-two-mass

Modèle unidimensionnel

Il s'agit d'un modèle d'excitation composé de deux masses, une large et une petite. Elles sont reliées par un ressort et un amortissement.

Modèle bidimensionnel

Nous avons toujours une masse large et une petite masse, mais cette fois ci la masse large a une forme en U qui englobe la petite masse. Elles sont reliées par deux ressorts et amortissements situés dans deux directions perpendiculaires.

Ce modèle à deux masses est utilisé pour simuler les marteaux, plectres, doigts et touche. Dans le cas du maillet, la petite masse joue le rôle de l'extrémité qui frotte et la masse large celui du manche de l'archet.

Annexe F : Consigne de l'expérience

Expérience sur la perception de sons de corde frottée

Nous nous intéressons à la façon dont les personnes perçoivent des sons produits par le frottement d'une corde avec des caractéristiques mécaniques différentes. Les sons proviennent du logiciel de synthèse sonore par modélisation physique Modalys. Le but de l'étude est de permettre un contrôle plus pertinent de ce type de synthèse par rapport à la perception.

- Pour vous donner une idée d'ensemble, nous vous présentons, au début de l'expérience, tous les sons dans un ordre quelconque. Pour écouter les sons cliquez sur le bouton [Écouter les sons]. Si vous souhaitez écouter une nouvelle fois les sons, cliquez sur [Encore]. Sinon cliquez sur [Terminer].
- L'expérience est un test de dissemblance dans lequel les échantillons sonores sont présentés par paire : le but du test n'est pas d'essayer d'identifier les sons mais plutôt d'utiliser votre impression globale, pour évaluer le degré de dissemblance entre les deux sons de chaque paire et positionner en conséquence le curseur sur l'échelle [Très similaire - Très différent]. Il vous est possible de réécouter la paire de sons autant de fois que vous le désirez avec le bouton de gauche [Rejouer].
- Essayez d'utiliser toute l'étendue de l'échelle de jugement au cours de l'expérience. Une fois que votre décision est prise sur la valeur de dissemblance entre les deux sons de la paire, vous validez votre résultat en cliquant sur le bouton de droite [Validez], et le bas de la fenêtre apparaît.
- Dans cette partie de la fenêtre, il vous est demandé de décrire verbalement les caractéristiques qui vous semblent faire la différence ou la similitude des deux sons présentés. La description doit être brève et synthétique. Inutile de faire des phrases, contentez-vous de noms ou d'adjectifs. Vous pouvez rejouer la paire de sons en cliquant sur [Rejouer]. Il est demandé de ne pas utiliser le symbole « ; » dans le champ de verbalisation.

- Une fois la verbalisation effectuée, cliquez sur la touche « return (retour chariot) » du clavier. Si vous ne souhaitez pas décrire verbalement une paire, vous devez quand même activer la fenêtre et cliquer sur cette touche « return ». Ceci étant fait, cliquez sur [Validez] et une nouvelle paire de sons se présente automatiquement. Il vous est alors impossible de revenir sur la paire précédente.
- Au préalable, une phase d'entraînement est proposée; elle s'effectue sur 3 paires choisies au hasard dans l'ensemble des échantillons de l'expérience. Profitez de ce moment pour, d'une part, vous familiariser avec la tâche et, d'autre part, commencer à dégager les critères perceptifs que vous utilisez naturellement pour effectuer votre jugement de dissemblance.

Le test s'effectue à votre convenance, en une seule fois ou bien avec des pauses (1 ou 2 maximum, pour ne pas rallonger démesurément la durée du test).

Annexe E : Analyses statistiques

Matrice de paramètres

L'analyse acoustique commence par le calcul des descripteurs acoustiques utiles sur un corpus de sons. Nous obtenons ainsi une matrice où chaque ligne correspond aux valeurs d'un descripteur acoustique, qui correspondent aux colonnes.

Calcul des corrélations

Le calcul de corrélation entre deux lignes de la matrice se fait de la manière suivante. On repère par un indice i la valeur du paramètre pour le son i ; soit n le nombre de sons. Soit m_x la moyenne des valeurs du paramètre X et m_y la moyenne des valeurs du paramètre Y.

La covariance s'obtient de la façon suivante :

$$(4.17) \quad cov(X, Y) = \frac{1}{n} \sum_i x_i y_i - m_x m_y$$

Le coefficient de corrélation r se calcule par :

$$(4.18) \quad r = \frac{cov(X, Y)}{\sqrt{V(X)V(Y)}}$$

avec V la variance.

On normalise par la racine du produit des variances, pour éviter que le résultat ne dépende de celles ci.

Le coefficient de corrélation prend des valeurs de -1 à 1. Une corrélation négative signifie que les paramètres corrélés varient en sens inverse. Si les valeurs sont faibles (en valeur absolue), les variables varient peu ensemble, si elles sont élevées (toujours en valeur absolue), les variables varient ensemble.

À partir de ce calcul, on forme une matrice des corrélations. Le nombre situé à la ligne i et la colonne j correspondra à la corrélation entre le paramètre situé

à la ligne i et celui situé à la ligne j de la matrice des paramètres.

Nous avons calculé une matrice des distances entre descripteurs à partir des corrélations. Dans un premier temps nous transformons les corrélations négatives en valeurs absolues. En soustrayant la matrice à 1, nous obtenons une matrice de dissemblance. Ainsi, plus deux descripteurs sont corrélés, plus la distance est faible.

Analyses par composantes principales

Le nombre de descripteurs acoustiques que nous avons (jusqu'à 35) ne permet pas de visualiser directement nos sons dans une base composée par les descripteurs, ce qui nous permettrait de détecter les principales variations acoustiques engendrées par la variation des paramètres mécaniques, et ainsi d'obtenir une carte de notre espace nous indiquant la corrélation et surtout la dispersion entre les paramètres.

Cependant, l'analyse par composantes principales (ACP) permet d'en visualiser les principales tendances, et ainsi de dégager les descripteurs acoustiques à la plus forte dispersion.

Dans cette méthode, chacun des paramètres de la matrice de départ est centré réduit, ceci afin de pouvoir comparer les dispersions de descripteurs qui n'ont pas forcément la même échelle. Nous pouvons en théorie choisir autant de composantes principales que de dimensions dans notre espace (c'est-à-dire 35 dans notre cas), mais chacune des premières composantes principales expliquant une importante proportion de la variance, souvent, 2 ou 3 suffisent. On projette ensuite les points de notre espace sur ces composantes principales.

Pour lire un ACP ou l'on visualise les descripteurs acoustiques comme c'est le cas dans notre étude, il suffit simplement de savoir :

- que les axes représentent une combinaison de descripteurs dont la dispersion est la plus grande, le premier axe ayant une plus grande dispersion que le second
- que plus deux variables sont proches sur le cercle des corrélations, plus leur corrélation est grande
- que deux variables sur le cercle des corrélations à angle droit (centré sur l'origine) ont une corrélation nulle
- qu'une variable loin du cercle des corrélations a une dispersion faible, voire inexistante si elle se situe au centre