MASTER - Année 2

Traitement du Son et de la Musique Avancé

Pierre Hanna

hanna@labri.fr

Année 1

- Méthodes d'analyse
 - représentation temporelle
 - représentation spectrale (FFT)
- Paramètres (amplitude, fréquence, timbre)
- Méthodes de synthèse
- Effets, Transformations

Année 2

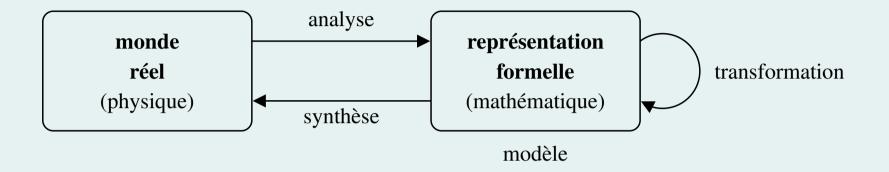
- Modèles analyse/synthèse
 - partie sinus
 - partie bruit
- Classification
 - Exemples
 - Projet Kaggle
- Synthèse (M. Robine)
- Traitement de la parole (J.L. Rouas)

A rélféchir:

- Stages (recherche, entreprise, ...)
- Projet de Dév 2nd semestre

Modélisation du son

Modèles



- Analyse : peu de contrainte de temps
- Synthèse : temps-réel

Modèles Physiques

Origine : physique des ondes sonores (Rayleigh fin 19ème)

- basés sur les modèles mathématiques d'acoustique
- simulation d'un instrument

Par rapport aux autres types de modélisations:

- avantages : qualité (défauts d'un instrument), contrôles réalistes, possibilité de définir des instruments non réels, . . .
- inconvénients : limites des instruments, temps de calcul (pas en temps réel), ...

Modèles Physiques

- Excitation/résonance
- Paradigme masse-ressort
 - ressort représente l'élasticité (corde, peau, ...)
 - déplacement longitudinal
 - déplacement transversal

$$y(x,t) = f_1(t + \frac{x}{v}) + f_2(t - \frac{x}{v})$$

Equations différentielles

$$Ky'' = \varepsilon$$

- $m{\varepsilon}$ K : tension de la corde, $m{\varepsilon}$ densité massique linéaire, y déplacement
- Etat intial
- Conditions aux limites (extrémités fixes/libres)

Cordes vibrantes, mais aussi membranes, ...

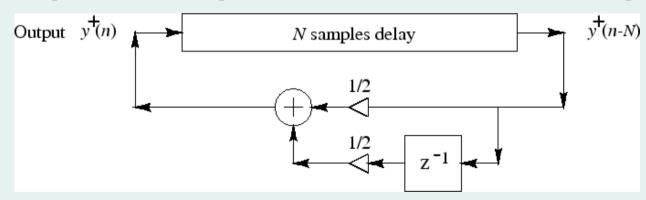
Cordes: Karplus-Strong

Algorithme de Karplus-Strong (CMJ 1983)

très efficace

$$H(z) = 0.5 + 0.5z^{-1}$$

- Basé sur la table d'onde, remplie au départ par des valeurs aléatoires
- Répétition de la table d'onde : hauteur perçue
- avec le temps, la table d'onde passe d'une onde aléatoire (bruit) à une onde quasi-périodique



- Son de corde original
- Son de corde 1
- Son de corde 2
- Son de corde 3
- Son de corde 4

Karplus-Strong pour les batteries

- Légère modification: ajout d'une probabilité p
- probabilité p d'avoir

$$y(n) = \frac{1}{2}(y(n-N) + y(n-N-1))$$

• probabilité 1 - p d'avoir

$$y(n) = -\frac{1}{2}(y(n-N) + y(n-N-1))$$

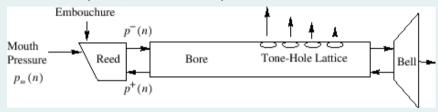
Résultats différents selon les valeurs de p

- \blacksquare Si p = 1: corde
- Si p = 0.5: perte de la périodicité et génère un son de percussion
- Si p = 0: coupe en partie les périodicités et génère un son de type harpe

Modèles Physiques

Autres exemples:

Waveguide Synthesis (89, Yamaha)



- Trajet de l'onde sonore le long d'un tuyau ou d'une corde
- Instruments à cordes et à vent
- Synthétiseurs commerciaux (Clavia Nord, Yamaha VL1, ...)
- Synthèse de la voix
 - Modèle source/filtre
 - Résonances : forme de l'enveloppe spectrale, notamment les formants

Modèles Physiques

Exemples de sons issus de la modélisation physique :

- Violoncelles
- Clarinettes/hautbois
- Cuivres
- Orgue

Modèles abstraits

Modèles abstraits : modèles de synthèse uniquement

- Paramètres mathématiques
- synthèse de sons complexes à partir d'*entités* simples
 - généralement des oscillateurs

Deux exemples les plus connus : synthèse FM et synthèse AM

Modèles spectraux

Son est représenté par une somme d'oscillateurs

Historique:

- Dès années 1903 pour films russes...
- Années 30: vocoder (voice coder) développé par Dudley (Bell Telephone Laboratories)
- Années 60: phase vocoder (Flanagan/Golden)
 Utilisation pour la musique (Risset, années 60)

Exemples de modélisations de voix:

- **9** 1939
- 1951
- ...
- Avec expressivité

Phase Vocoder par FFT

Vocoder : modification de la transformée de Fourier à court-terme

- Calculer la FFT (trames, chevauchement, fenêtrage)
- Pour chaque bin *k*
 - calcul de l'amplitude
 - calcul de la phase
- Calcul de la fréquence instantanée $F_m(k)$ au temps m (trames de N échantillons) :

$$F = \frac{1}{2\pi} \frac{d\phi}{dt}$$

$$F_m(k) = \frac{\theta_m(k) - \theta_{m-1}(k) + 2p\pi}{2\pi(t_m - t_{m-1})}$$

$$F_m(k) = \frac{\theta_m(k) - \theta_{m-1}(k)}{2\pi \frac{N}{F_c}}$$

Phase Vocoder par FFT

Utilisation:

- Compression
- Traitement parole
- Supression reverbération
- Réduction de bruit

Limites

- Bruit
- Attaques
- Modulations fréquence/amplitude (vibrato/tremolo)

Modèles sinusoïdaux

[McAulay, Quatieri (IEEE TASSP, 1986)] applications [Smith, Serra 1986]

Le signal audio *s* est donné dans le domaine temporel par:

$$s(t) = \sum_{p=1}^{P} a_p(t) \cos(\phi_p(t))$$

$$s(t) = \sum_{p=1}^{P} a_p(t) \cos(2\pi f_p(t)t + \phi_p(0))$$

où P est le nombre de partiels et

$$\phi_p(t) = \phi_p(0) + 2\pi \int_0^t f_p(u) du$$
 (1)

Les fonctions $f_p(t)$, $a_p(t)$ et $\phi_p(t)$ sont, respectivement, la fréquence, l'amplitude et la phase du p-ième partiel. Ces paramètres varient lentement dans le temps.

TP: partie sinus, spectre amplitude

A partir du code analyse_amp.c,

- Localiser/visualiser les pics du spectre d'amplitude correspondant à la partie sinus d'un son (par exemple piano_original.wav)
 - Pic (maximum local présent sur plusieurs trames consécutives)
 - Seuil sur amplitude
 - Différence aux voisins
 - Pic mais aussi les 2 voisins
 - Différence avec interpolation parabolique
- Tester sur plusieurs sons (mélanges sinus + bruit à synthétiser)
- Tester plusieurs approches

TP: partie sinus, spectre de phase

Localiser/visualiser les pics en utilisant le spectre de phase

- Observer le spectre de phase du son sinus440+bruit.wav
 - autour du pic de la sinusoïde
 - sur une région correspondant au bruit
- influence du zero-phase
- Proposer une méthode pour détecter les pics correspondant aux sinus
- Implémenter cette méthode et tester.

TP: partie sinus, analyse/synthèse

A partir du code modele.c,

- ReSynthétiser la partie sinus d'un son (par exemple piano_original.wav)
 - Ajouter un test pour décider si un bin est une sinus ou pas
 - mettre à 0 les parties du spectre non sinus
- Tester sur plusieurs sons
- Tester plusieurs approches

Modèles hybrides

Limites des modèles sinusoïdaux pour les sons complexes :

- Séparation du son en plusieurs composantes
 - Partiels
 - Transitoires
 - Bruit (résidu)
- Traitement indépendant de chaque composante
- Fusion des composantes pour la synthèse

Définition du bruit (1)

Langage courant:

- sons perturbateurs
- sons non désirés
- sons désagréables

Sinusoide de 1000Hz très désagréable, pourtant ce n'est pas un bruit.

Définition du bruit (2)

Acoustique, sciences:

- Spectre dense
- Signal défini statistiquement (2 signaux différents mais perception équivalente)

transitoires: traitement différent

Définition du bruit (3)

Modèles analyse/synthèse de sons:

- Tout ce qui n'est représenté par des sinusoïdes évoluant lentement dans le temps.
- Tout ce qui n'a pas été retenu par le vocodeur de phase.

Frontière son harmonique ou inharmonique complexe/bruit ? Critère pour décider si un spectre est dense ? Instrument seul/Multitude d'instruments

Intérêt du bruit (1)

Présent dans les sons d'instruments de musique:

- Sons denses (cymbales)
- Souffle pour les instruments à vent ou pour la parole
- Courte impulsions : coups de marteau (piano), cordes pincées (guitare, violon, . . .)
- Bandes de fréquences élargies (tampons du piano qui stoppent les vibrations de la corde)
- Multiples impulsions (chocs) comme pour les maracas.
- ⇒ Importance dans la synthèse de sons: un son reste *artificiel* si l'on ajoute pas de bruit

Intérêt du bruit (2)

Traitement de la parole

- Consonnes non voisées (s,ch,f,...)
- Voix chuchotées

Importance dans les expériences psychoacoustiques

• expériences de masquage: mélange de bruit avec des sons harmoniques

Modèles pour le bruit

Deux approches possibles:

- Approche temporelle: synthèse de bruits *classiques*/filtrés
- Approche spectrale: modèles de synthèse complétant les modèles de sons harmoniques

Approche temporelle

Etude de la distribution des échantillons

- Tendance à tendre vers une distribution gaussienne
- Deux sons bruités peuvent être très proches, tout en ayant des propriétés statistiques différentes
- Deux sons bruités peuvent être très différents, même si leurs propriétés sont ressemblantes

Seule approche envisageable :

- filtrer du bruit blanc
- pas vraiment de modèle...

Plan

Synthèse de Bruit : temporel



Synthèse temporelle

Description et algorithme de synthèse de bruits :

- Bruit blanc
- Bruit rose
- Bruit brownien
- Bruit ondulant

Bruit blanc (1)

white noise

Analogie avec lumière blanche (mélange de toutes les couleurs): contient **toutes** les fréquences.

- Spectre de fréquence continue
- Spectre d'amplitude moyenne égale

Modèle (n'existe pas naturellement) restreint à 20 – 20000Hz

Bruit blanc (2)

Propriétés auditives:

- Pas de hauteur perçue
- Pas de rythme perçu

Synthèse de bruit blanc

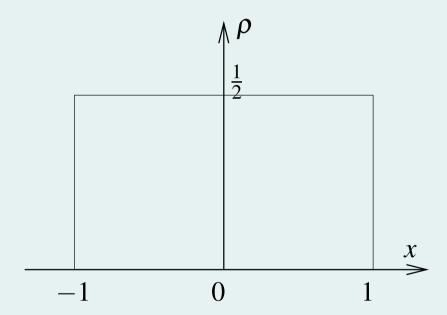
Méthodes reposent sur la définition: indépendance et même loi. Différences selon la loi choisie : perception (intensité), efficacité

- Uniforme (rectangulaire)
- Gauss (normale)

Bruit blanc: uniforme

$$\forall (x_1, x_2) \in]-1;1]^2, p(x=x_1) = p(x=x_2)$$

Densité de probabilité $\rho: \rho(x) = \frac{1}{2}$



Bruit blanc: uniforme

Algorithme:

Utilisation de la fonction de tirage aléatoire (rand du C)

Bruit blanc: gauss (1)

Appelée parfois normale (normal or gaussian)

Tendance naturelle des bruits à être gaussiens (théorème central limite)

$$Y = \sum_{k=1}^{N} X_k$$

Y suit une loi gaussienne si N est grand, quelle que soit la distibution des X_k .

Bruit blanc: gauss (2)

Suit une loi de Gauss:

$$\rho_N(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\frac{-x^2}{(2\sigma^2)}$$
 (2)

où $x \in \mathbb{R}$ et σ représente l'écart-type (largeur de la fonction de distribution).

Bruit blanc: gauss (5)

Amplitude RMS:

$$A_{RMS} = \sqrt{\langle x^2 \rangle} = \boldsymbol{\sigma}$$

Amplitude max: dépend de σ

Compromis:

- \bullet σ grand : probabilités de saturation
- \bullet σ petit : intensité perçue faible

 \implies normalisation

Bruit blanc: gauss- Algorithme (1)

Principe:

• addition d'un certain nombre de variables aléatoires x_i uniformement distribuées

Littérature indique N = 12 bonne approximation

Bruit coloré

Même analogie avec la couleur.

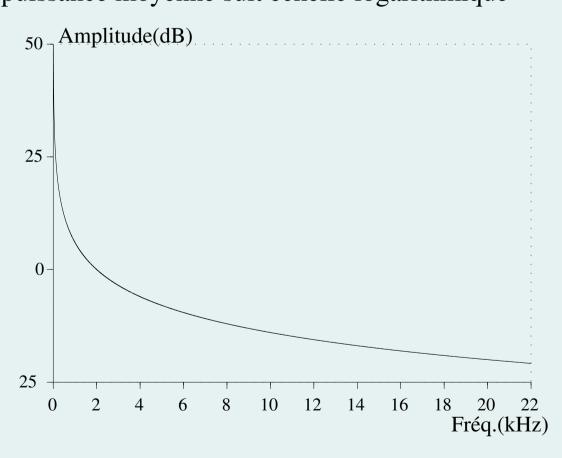
Forme du spectre permet d'associer une couleur

- Bruit rose
- Bruit brownien (rouge)

Autres ...

Bruit coloré

Bruits roses: distribution de puissance moyenne suit échelle logarithmique



Bruit rose: applications (2)

- Application en Musique: distribution des mélodies des morceaux de musique. La plupart (classique ou moderne) suivent une distribution rose.
- Composition aléatoire (Xenakis)

Bibliographie sur $\frac{1}{f}$ et sur $\frac{1}{f^2}$: http://linkage.rockefeller.edu/wli/1fnoise/

Bruit rose: algorithme de Voss (1)

Principe:

- Chaque échantillons est la somme de *n* variables aléatoires.
- Pour avoir l'échantillon, on ne retire qu'un certain nombre de ces variables.

Comme les hautes fréquences sont atténuées, deux valeurs successives vont être proches: peu de variables retirées.

Bruit rose: algorithme de Voss (2)

```
Data
            : N
           : tableau samples de taille 2^N
Result
début
     pour i = 0 à N - 1 faire
          oldbit[i] = 1;
     fin
     pour K = 0 à 2^N - 1 faire
          K = bit[0], bit[1], ..., bit[N-1];
          sum = 0;
          pour i = 0 à N - 1 faire
                \mathbf{si} \ bit[i] \neq oldbit[i] \ \mathbf{alors}
                     d[i] = rand(0.1);
                fin
                sum = sum + d[i];
                oldbit[i] = bit[i];
                samples[i] =sum;
          fin
     fin
fin
```

Bruit rose: algorithme de Voss (3)

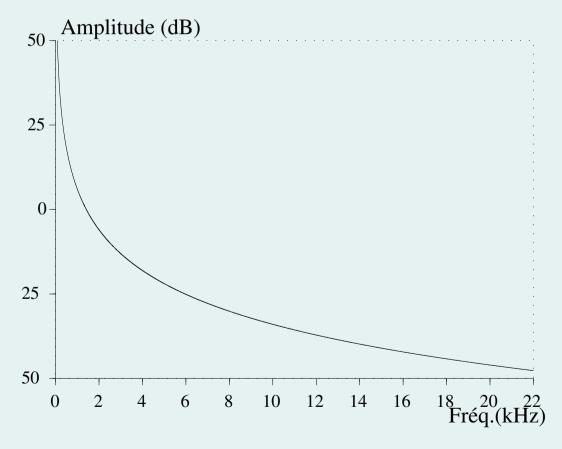
Exemple sur 3 bits:

K	0	1	2	3	4	5	6	7
bits	000	001	010	011	100	101	110	111
retirages	3	1	2	1	3	1	2	1

Bruit Brownien (1)

Aucun lien avec couleur marron, analogie avec mouvements browniens des molécules dans un fluide.

Autre nom: random walk ou l'homme saoul



Bruit Brownien (2)

$$S(f) = \frac{A}{f^2}$$

avec $f \ge f_{min}$

$$S(2f) = \frac{A}{4f^2} = \frac{1}{4}S(f)$$

ce qui signifie que, dès que f double de valeur, S(f) est divisé par 4.

Bruit Brownien: algorithme

Différences entre échantillons successifs = variables aléatoires uniformément distribuées.

Coloration d'un bruit (1)

Modification de la couleur d'un bruit

⇒ modification de l'enveloppe spectrale

Les modèles temporels sont basés sur les filtres

- ⇒ les bruits sont considérés comme des bruits blancs filtrés.
 - Linear Predictive Coding (analyse/synthèse de la parole).

Coloration d'un bruit (2)



Excitation: bruit blanc

Résonateur : filtres

Coloration d'un bruit (2)

Avantage:

- efficacité
- modélisation physique

Inconvénients:

- Lissage de l'enveloppe : compromis précision/complexité
- Intuitivité du modèle : suite de réels

TP : synthèse de bruit

A partir du code bruit.c,

- Synthétiser un bruit blanc
 - Distribution uniforme
 - Distribution normale
- Analyser le spectre d'amplitude obtenu
- Comparer
- Synthétiser un bruit brownien
- Comparer son spectre à celui du bruit blanc

Synthèse temporelle: conclusions

- Facilité et efficacité de synthétiser des bruits colorés de base.
- Plus complexe et moins intuitif pour couleurs moins simples.

Avantage: lien avec la physique

Problème: lien avec la perception

Plan

Synthèse de Bruit : spectral



Modèle bruit thermique (1)

$$x(t) = \sum_{n=0}^{N} [A_n \cos(\omega_n t) + B_n \sin(\omega_n t)]$$

où n est un entier et

- les pulsations ω_n sont également espacées
- les coefficients A_n et B_n sont des variables aléatoires.

Modèle bruit thermique (2)

Autre possibilité d'écriture:

$$x(t) = \sum_{n=0}^{N} C_n \cos(\omega_n t + \Phi_n)$$

où

- $\Phi_n = \arctan(\frac{B_n}{A_n})$ distribuées uniformément entre 0 et 2π
- $C_n = B_n^2 + A_n^2$ distribués selon la loi de Rayleigh, valeur plus probable σ

Plusieurs méthodes existantes, selon les applications

- Estimation directe depuis le signal
 - beaucoup de méthodes qui estiment les sinusoïdes
 - beaucoup de méthodes qui estiment le niveau de bruit
 - mais peu qui cherchent à estimer le bruit...
- Estimation à partir d'une estimation de la partie déterministe (sinusoïdale)

Erreur inévitable...

Décision binaire pour chaque pic du spectre

- Sinusoïde
- Bruit

Pic peut être un mélange sinus/bruit Autres approches basées sur l'analyse à long terme du spectre

Spectre des parties bruitées évolue rapidement au cours du temps Idée: Diminuer ces variations (ou la variance)

- Moyenne (Méthode de Welsh)
- Minimum
- Maximum

Etude théorique nécessaire

Amplitude des composants bruités seuls suivent une loi de Rayleigh:

$$p(M) = \frac{M}{\sigma^2} e^{\frac{-M^2}{2\sigma^2}}$$

où σ la valeur la plus probable

Remarque : probabilité non nulle d'avoir une amplitude très forte ou nulle.

Amplitude des composants sinusoïdaux bruités suivent une loi de Gauss:

$$\rho_N(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-x^2}{(2\sigma^2)}$$

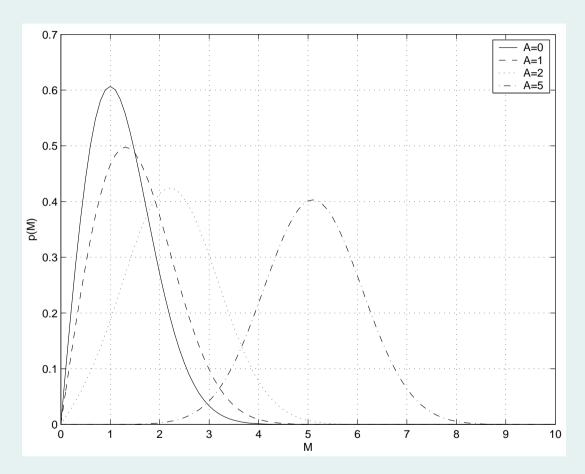
où $x \in \mathbb{R}$ et σ représente l'écart-type (largeur de la fonction de distribution). Loi pour les deux cas ?

Loi de Rice

$$p_{A,\sigma}(M) = \frac{M}{\sigma^2} e^{\frac{-\left(M^2 + A^2\right)}{2\sigma^2}} I_0\left(\frac{AM}{\sigma^2}\right)$$

où I_0 est la fonction de Bessel modifiée de la première espèce d'ordre 0:

$$I_0(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{x\cos(\phi)} d\phi$$



- A tend vers 0 : Rayleigh
- A plus grand que σ : Gauss (moyenne A et écart-type σ)

Modèle bruit thermique (3)

Théorie:

- Vérification de E(x[n]x[n+p]) = 0
- Répartition de l'énergie dans le spectre

Deux possibilités:

- 1. Amplitudes distribuées selon Rayleigh, phases uniformément choisies
- 2. Amplitudes fixes, phases uniformément choisies

La deuxième méthode est la plus utilisée.

Méthodes de synthèse spectrale (1)

• Détermination aléatoire des valeurs d'amplitude C_n Distribution Rayleigh:

$$\rho(x) = \frac{x}{\sigma^2} \exp{-\frac{y^2}{2\sigma^2}}$$

• Phases déterminées aléatoirement (loi uniforme) entre 0 et 2π

Quel que soit le nombre de composants sinusoïdaux du bruit (même 1), la distribution des échantillons suit une loi de Gauss.

Algorithme: distribution de Rayleigh

Algorithme à partir de valeurs aléatoires $0 \le x < 1$ uniformément distribuées et l'inverse de la fonction cumulative $\int_0^x \rho(z) dz$:

$$y_n = \sigma_n \sqrt{-2\ln(1-x)}$$

où σ_n^2 est la puissance moyenne du composant n.

Méthodes de synthèse spectrale (2)

- Valeurs d'amplitude C_n fixées
- Phases déterminées aléatoirement (loi uniforme) entre 0 et 2π

Avantage: chaque signal synthétisé a exactement le même spectre de puissance.

Méthodes de synthèse spectrale (3)

Les coeficients A_n et B_n ne sont plus distribués selon Gauss, et ne sont plus indépendants:

$$A_n = C_n \cos \Phi_n$$

$$B_n = C_n \sin \Phi_n$$

où C_n est une valeur fixe.

Phases uniformément distribuées $\Longrightarrow A_n$ et B_n suivent une loi décrite par la fonction de densité de probabilité d'une sinusoïde.

Théorème central limite: $N > 5 \Longrightarrow$ la distribution des échantillons est gaussienne.

Stationnarité (1)

Limites:

- Efficacité (nombre de sinusoïdes)
- Hypothèse de stationnarité (analyse)
- Transitions entre fenêtres successives (clics)

Pour résoudre ce problème: technique de recouvrement et addition.

Avantage: réduction du nombre d'oscillateurs.

- Etude de la distribution des amplitudes pour un bruit blanc sur plusieurs frames
- Distribution de Rice (Rayleigh/Gauss)
- Si Rayleigh : Bruit
- Si Gauss : Bruit + Sinus

La valeur de σ indique le niveau de bruit.

TP : synthèse de bruit

A partir du code bruit.c,

- Synthétiser un bruit blanc
 - Amplitudes uniformes, phases aléatoires
 - Amplitudes Rayleigh, phases aléatoires
- Comparer
- Visualiser et analyser les spectres d'amplitude obtenus

TP: analyse des distributions

A partir du code analyse_amp.c et du son sinus440+bruit.wav

- Analyser et afficher la ditribution des amplitudes pour une sinus
- et pour du bruit
- Comparer
- Tester pour d'autres sons avec différents niveaux de bruit

Plan

Retour aux modèles hybrides



Modèles Sinus+Bruit

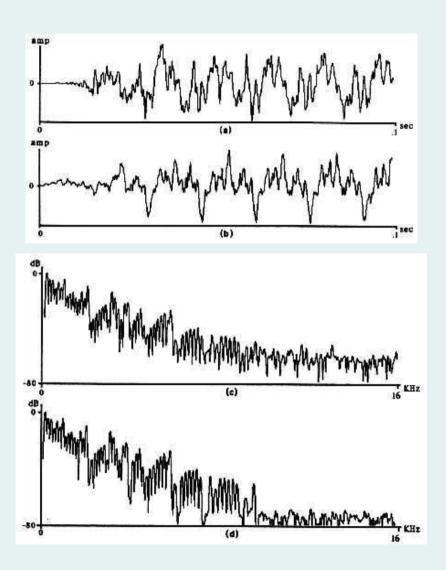
SMS Serra/Smith 1989

$$s(t) = \sum_{p=1}^{P} a_p(t) \cos(\phi_p(t)) + e(t)$$

- e(t): composant résiduel (ou bruit)
- Bruit blanc filtré
- Obtenu par soustraction spectrale:

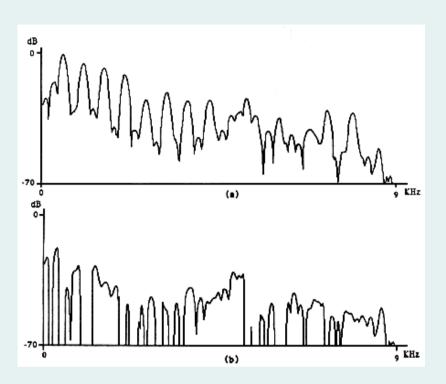
$$e(t) = s(t) - \sum_{p=1}^{P} a_p(t) \cos(\phi_p(t))$$

Soustraction



Besoin d'interpoler la phase...

Soustraction spectrale



$$|E_r(k)| = ||X_r(k)| - |D_r(k)||$$

- Soustraction des spectres d'amplitude
- Pas besoin d'avoir les informations de phase de la partie sinsuoïdale
 - inutile pour la soustraction spectrale
 - inutile pour la partie résiduelle

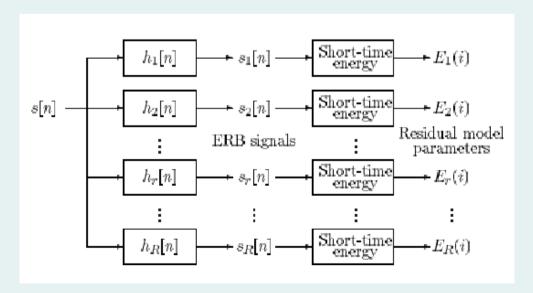
Modèles Sinus+Bruit

Partie bruit:

- Uniquement représentée par le densité spectrale de puissance (enveloppe spectrale)
- Phases considérées comme aléatoires

Autres modèles:

Energies dans des bandes ERB (Goodwin 96)



Somme de sinus de courtes durées (Moi 03)

Modèles Sinus+Bruit

Exemples:

- Piano
- Guitare
- Conga

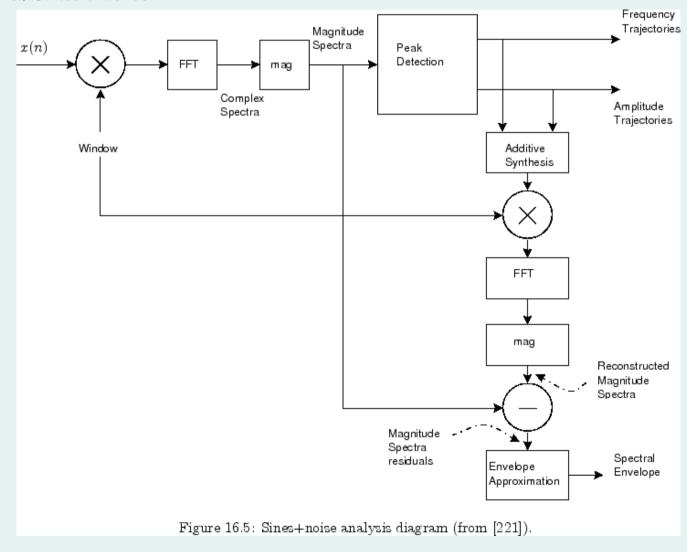
TP: Partie Bruit

A partir du code modele.c,

- ReSynthétiser la partie sinus d'un son (par exemple piano_original.wav)
- ReSynthétiser la partie bruit du même son
- Mélanger les deux et comparer au son original
- Tester sur plusieurs sons
- Tester en approximant l'enveloppe spectrale (segments)

Modèles Sinus+Bruit

SMS Serra/Smith 1989



Modèles non hybrides

Fitz 1998

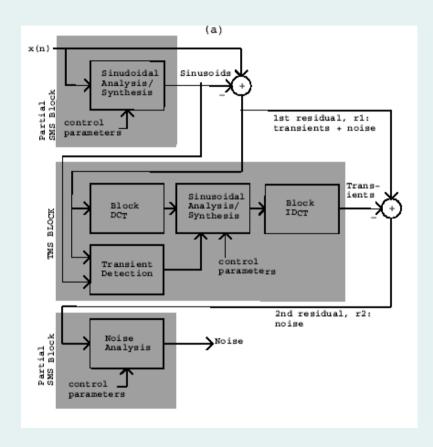
Son : somme de sinusoïdes bruitée

$$y_n = A(\sqrt{1-K} + +\sqrt{2K}[\delta_n * h_n]) \cdot \exp(j\omega_c n)$$

- Partiel centré en ω_c
- *K* proportion de bruit
- δ_n bruit blanc
- h_n réponse impulsionnelle (enveloppe du bruit)

Modèles Sinus+Attaques+Bruit

STN Verma/Smith

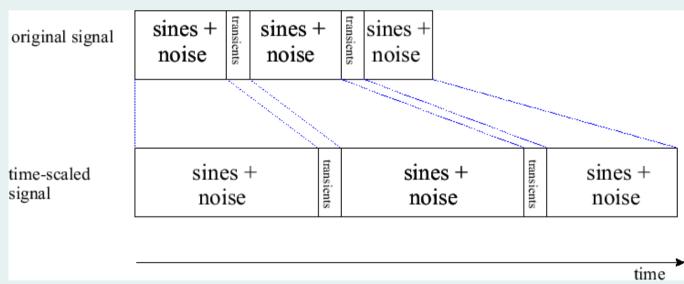


Modèles Sinus+Attaques+Bruit

Autre approche (Nsabimana/Zölzer 2008)

- Détection précise des transitoires
- Supression
- Extrapolation du signal pour combler les trous

Applications:



Modèles Sinus+Attaques+Bruit

Exemples

- Analyse
- Etirement temporel