

Influence de la densité spectrale sur la synthèse de sons bruités

Pierre HANNA et Myriam DESAINTE-CATHERINE
SCRIME¹

LaBRI², Université Bordeaux I
351, cours de la Libération
F-33405 Talence cedex – France
{hanna,myriam}@labri.fr

Résumé

Les modèles de synthèse spectrale de sons bruités posent comme hypothèse que ces sons sont des bruits blancs filtrés. Nous proposons dans ce papier une méthode de synthèse spectrale basée sur des paramètres statistiques permettant de définir les distributions des fréquences et des phases de façon à contrôler indépendamment de l'enveloppe spectrale un facteur, mis en avant par des travaux menés par des psychoacousticiens, et considéré implicitement jusqu'alors comme constant : la densité spectrale. Ces paramètres de synthèse et leurs influences sont détaillés ainsi que l'algorithme général.

Mots clefs

Synthèse, bruits, modèle spectral, statistique, enveloppe dynamique

1. Introduction

Les méthodes de synthèse basées sur des modèles spectraux sont principalement destinées à l'analyse, la transformation et la synthèse de sons harmoniques ou quasi-harmoniques³. Pourtant, des sons bruités non harmoniques sont très souvent utilisés par des compositeurs de musique électroacoustique. La plupart du temps, de tels sons sont synthétisés à l'aide de bruits blancs générés par des tirages aléatoires d'échantillons. Ces bruits sont ensuite filtrés. Nous proposons dans cet article une nouvelle méthode de synthèse, basée sur un modèle spectral et statistique, pour pouvoir générer des sons bruités en autorisant le contrôle de l'enveloppe spectrale, mais aussi de la densité spectrale et de ses variations, propriété considérée dans les modèles existants comme constante, mais mise en avant lors d'études psychoacoustiques⁴. L'utilisation d'une approche spectrale permet de conserver une certaine homogénéité avec d'autres modèles précis et efficaces, et l'utilisation des outils statistiques permet de prendre en compte les variations complexes du spectre.

Ainsi, après avoir introduit les modèles existants et la définition des bruits sur laquelle notre travail est basé, nous définissons les paramètres statistiques pour pouvoir contrôler les distributions de fréquences et de phases. Pour chaque paramètre, nous essayons de montrer leur influence du point de vue mathématique, musical et sur le plan de la perception. Ensuite, nous présentons un algorithme général de synthèse à partir de ces paramètres.

2. Modèles spectraux

2.1. Bruits thermiques

Les bruits thermiques sont définis⁵ par :

$$X(t) = \sum_{n=1}^N C_n \sin(\omega_n t + \Phi_n)$$

¹ Studio de Création et de Recherche en Informatique et Musique Electroacoustique, soutenu par le Conseil régional d'Aquitaine, l'Université Bordeaux I, le ministère de la Culture, la Direction régionale des Actions culturelles d'Aquitaine, la Mairie de Bordeaux, l'ENSERB et le Conseil général de la Gironde.

² Laboratoire Bordelais de Recherche en Informatique, Unité Mixte de Recherche 5800 du CNRS.

³ MARCHAND S., "Sound models for computer music: analysis, transformation, synthesis of musical sound", Ph.D Thesis, LaBRI, Université Bordeaux I, 2000.

⁴ HARTMANN W.M., McADAMS S., GERZSO A. and BOULEZ P., "Discrimination of spectral density", *Journal of Acoustical Society of America*, vol.79, no.6, pp.1915–1925, 1986.

⁵ HARTMANN W.M., *Signals, Sound, and Sensation*, Modern Acoustics and Signal Processing AIP Press, 1997.

où N est le nombre de fréquences, n est un entier, ω_n sont les pulsations régulièrement espacées, C_n sont des variables aléatoires distribuées selon une loi de Rayleigh. Et enfin, Φ_n sont des variables aléatoires uniformément distribuées.

2.2. SMS

Le modèle SMS⁶ sépare le son analysé en une partie déterministe et une partie stochastique. L'hypothèse principale pour cette partie stochastique est qu'elle peut être complètement décrite par son enveloppe spectrale. Elle est donc resynthétisée par une transformée inverse de Fourier. Cette transformation n'est mathématiquement rien d'autre qu'une somme d'un nombre fixe de sinusoïdes, dont les fréquences sont proportionnelles au rapport entre le taux d'échantillonnage et la taille de la fenêtre de synthèse (R/N), et dont les amplitudes sont des valeurs de l'enveloppe spectrale. Les phases sont des variables aléatoires suivant une loi uniforme.

$$X(t) = \sum_{n=1}^{N/2} c_n \sin\left(2\pi n \frac{R}{N} t + \Phi_n\right)$$

Avec ce modèle, il est seulement permis de contrôler les variations d'enveloppe dynamique en modifiant l'enveloppe spectrale d'une fenêtre sur l'autre : ces variations sont donc dépendantes de l'enveloppe spectrale.

3. Approche spectrale et statistique

3.1. Synthèse additive

Dans ce papier, nous considérons les sons (échantillonnés selon le taux R) dans le modèle spectral comme des processus aléatoires, comme dans un précédent article⁷. Chaque composant fréquentiel F_n est une variable aléatoire avec une amplitude fixée c_n et une phase aléatoire Φ_n :

$$X_k = \sum_{n=1}^N c_n \sin\left(2\pi F_n \frac{k}{R} + \Phi_n\right)$$

où les fréquences F_i sont distribuées dans une bande de largeur W (Hz).

Dans notre approche, c_n est fixé et défini par un seul paramètre : l'enveloppe spectrale (voir partie 5.6).

3.2. Distribution des fréquences

A partir de la représentation choisie ci-dessus, nous définissons la loi de distribution des fréquences composant le signal. Pour cela, nous introduisons des paramètres liés à cette loi (représentés sur l'exemple 1), de façon à ce que, d'une part, celle-ci soit la plus générale possible et, d'autre part, les paramètres présentés puissent être reliés à des paramètres intuitifs et musicaux.

Ainsi, pour définir aléatoirement ces fréquences sur une certaine bande, nous choisissons de tirer d'abord les casiers dans lesquels vont se situer les fréquences. Puis, à l'intérieur de chaque casier, une fréquence est tirée selon une loi uniforme de largeur contrôlable.

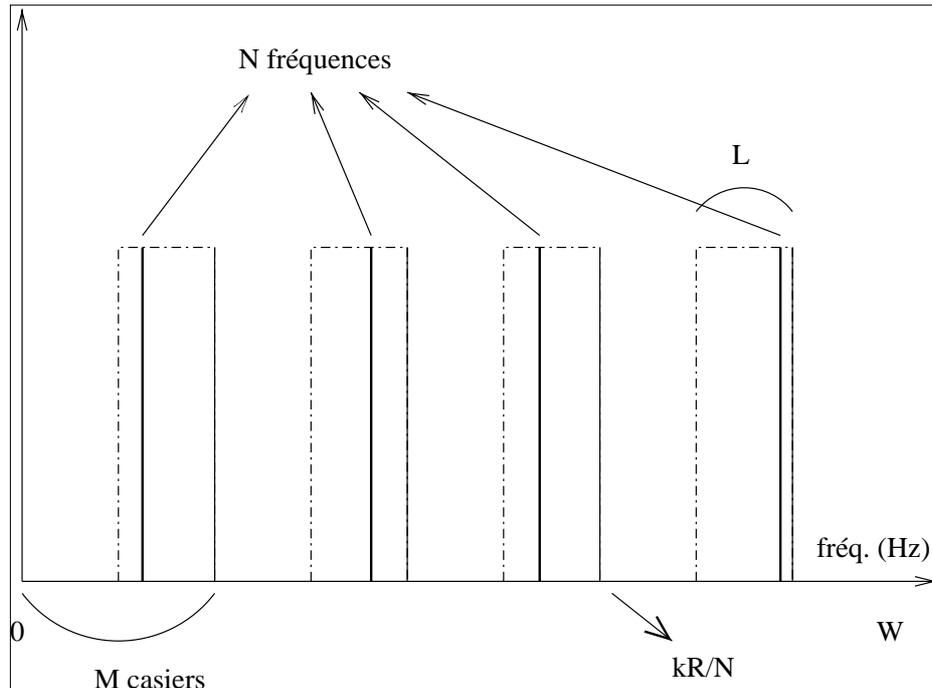
Les paramètres suivants, schématisés sur la figure 1, sont donc définis :

- N est le nombre de fréquences aléatoirement choisies dans une bande.
- W est la largeur de cette bande de fréquences (Hz).
- M est le nombre de casiers (*bins*) ($N \leq M$).
- L est la largeur de la distribution uniforme ($0 \leq L \leq 1$), dont le bord est centré sur le casier.

L représente une fraction de la taille du casier $\frac{W}{N}$.

⁶ SERRA X. and SMITH J., "Spectral modeling synthesis : a sound analysis/synthesis system based on a deterministic plus stochastic decomposition," *Computer Music Journal*, vol.14, no.4, pp.12-24, 1990.

⁷ HANNA P. and DESAINTE-CATHERINE M., "Statistical approach for sound modeling," *Proceedings of the Digital Audio Effects Workshop (DAFX'00, Verona, Italy)*, pp.91-96, 2000.



Exemple 1 : Paramètres de la distribution des fréquences.

Une infinité de distributions est donc permise. Cependant, en donnant certaines valeurs particulières à ces paramètres, il est possible de se ramener à des méthodes plus *classiques*.

3.2.1. Transformée de Fourier Inverse (SMS)

En choisissant $M=N$ et $L=0$, le choix des fréquences ne se fait plus de manière aléatoire, puisque chaque fréquence se retrouve centrée sur chaque casier. Ce choix permet donc de se ramener, dans le cas où $W=R/2$ et où la durée du signal synthétisé est égale à $2N$, à une transformée inverse de Fourier ($F_k=kR/N$).

3.2.2. Tirage uniforme

En choisissant M infini (très grand par rapport à N), les fréquences sont définies aléatoirement selon une loi uniforme sur toute la bande W de fréquences. Cette méthode induit quelques propriétés que nous décrivons dans la partie 5.

4. Algorithme de synthèse

La mise en œuvre de la méthode de synthèse d'après les paramètres introduits ci-dessus est présentée dans cette partie, après avoir insisté sur l'algorithme de tirage des casiers.

4.1. Tirage des casiers

Nous avons vu précédemment que les cas $N=M$ et $M \gg N$ se ramènent à des méthodes *classiques*. Par contre, lorsque $M > N$, il faut dans un premier temps effectuer un tirage de N casiers sur les M possibles, dans lesquels seront choisies les fréquences du spectre. Deux méthodes peuvent être envisagées, un tirage par rejet ou un tirage équivalent à celui d'une permutation aléatoire. Pour être le plus rigoureux possible, nous avons choisi la deuxième méthode. Ainsi, nous supposons la suite de casiers $(0, \dots, M-1)$, nous tirons aléatoirement un casier i et nous échangeons les casiers $M-1$ et i . Nous recommençons ensuite avec la suite $(0, \dots, M-2)$, ..., jusqu'à ce que N casiers aient été tirés.

4.2. Algorithme principal

Data : paramètres, enveloppe spectrale

Begin

Pour toutes les fenêtres *faire*

Pour $n \in 0, \dots, N-1$ faire

Tirage d'un casier parmi les $N-n$ restants ;

Tirage d'une fréquence dans le casier choisie selon une loi uniforme dépendante de L ;

Tirage d'une phase selon une loi uniforme ;

End

Synthèse additive ;

End

End

5. Influence des paramètres

Les paramètres définis dans la partie 3.2 ont une influence sur le plan perceptif et permettent donc de contrôler certaines propriétés du son synthétisé du point de vue musical. Nous étudions ici l'influence des paramètres de la distribution des fréquences, en précisant notamment quelques éléments mathématiques et psychoacoustiques.

5.1. Influence de la largeur de bande

W représente dans notre approche la largeur de la bande de fréquences. Il est évident qu'un changement de sa valeur va avoir une influence directe sur le spectre du son synthétisé, et plus particulièrement sur la densité spectrale (que nous définissons précisément ci dessous). Cette bande peut donc non seulement être élargie ou rétrécie, mais également être *transposée* le long de l'axe des fréquences.

5.2. Influence du nombre de composants fréquentiels

N représente donc le nombre de sinusoides composant le signal synthétisé.

5.2.1. Perception de la densité spectrale

Une sinusoïde de fréquence et d'amplitude fixes est perçue comme un son parfaitement stable, alors que l'amplitude de l'onde varie beaucoup au cours du temps. Cette stabilité peut être décrite par les fluctuations d'enveloppe temporelle. Ce même paramètre peut être également très utile pour décrire les bandes de bruit⁸. Une application possible de l'étude des *fluctuations d'enveloppe temporelle* est l'explication de la faculté des auditeurs à faire la distinction entre des sons de différentes densités spectrales, cette *densité spectrale* étant définie par le rapport entre le nombre de composants fréquentiels et la largeur de la bande de fréquences.

Plusieurs expériences psychoacoustiques montrent que cette capacité est principalement basée sur deux propriétés, une temporelle et une spectrale. Le caractère spectral correspond à la hauteur associée aux bandes

$$\rho = \frac{N}{W}$$

de bruits (*trous* dans la bande de fréquences), alors que le caractère temporel est lié aux fluctuations d'enveloppe. C'est pourquoi modifier le nombre de composants fréquentiels N a une influence assez importante sur la perception. Il est intéressant de noter que les modèles de bruits existants ne prennent pas en compte ces facteurs psychoacoustiques.

5.2.2. Contrôle de la densité spectrale

Il est possible de relier mathématiquement la moyenne des fluctuations d'enveloppe (plus exactement la variance de la puissance de l'enveloppe) aux paramètres de la distribution des fréquences, en faisant l'hypothèse que les phases sont uniformément distribuées⁸. Ce calcul n'est pas présenté en détail dans ce papier.

$$V(E^2) = f(W, N, M, L)$$

Des courbes expérimentales⁹ sont présentées qui montrent bien qu'un maximum de densité moyenne peut être atteint pour une largeur de bande W fixée en augmentant N , M et L . Au-delà de ces valeurs, les expériences psychoacoustiques montrent également que les densités spectrales ne sont plus différenciées.

⁸ HARTMANN W.M., *Signals, Sound, and Sensation*, Modern Acoustics and Signal Processing AIP Press, 1997.

⁹ HANNA P. and DESAINTE-CATHERINE M., "Influence of frequency distribution on intensity fluctuations of noise," *Proceedings of the Digital Audio Effects Workshop (DAFX'01, Limerick, Ireland)*, pp.125-129, 2001.

Il est d'ailleurs intéressant de noter que dans la méthode de synthèse SMS (partie stochastique), qui se fait en utilisant une transformée inverse de Fourier, la densité spectrale est toujours maximale sur toutes les bandes de fréquences. En effet, pour une fenêtre de durée $2k$ échantillons (taux d'échantillonnage R), la transformée inverse de Fourier est en fait la somme de k sinusoïdes fixées. Les différences minimales entre chaque fréquence sont donc de $R/2k$ et il est facile de montrer que cette différence ne peut pas être perçue à cause de la taille de la fenêtre de synthèse. Cette méthode permet donc de se placer implicitement dans le cas d'une densité spectrale maximale.

5.2.3. Paramètre musical

La densité spectrale est non seulement importante du point de vue de la perception, mais aussi musicalement. Shaeffer¹⁰ utilise ainsi le concept de *grain* pour décrire différents sons complexes. Il est défini entre le domaine du rythme et celui de la hauteur, comme de nombreuses petites irrégularités à la surface du son.

5.3. Influence du nombre de casiers

M est le nombre de casiers de la bande de fréquences dans lesquels les N fréquences vont être choisies plus ou moins aléatoirement.

5.3.1. Contrôle des variations de densité spectrale

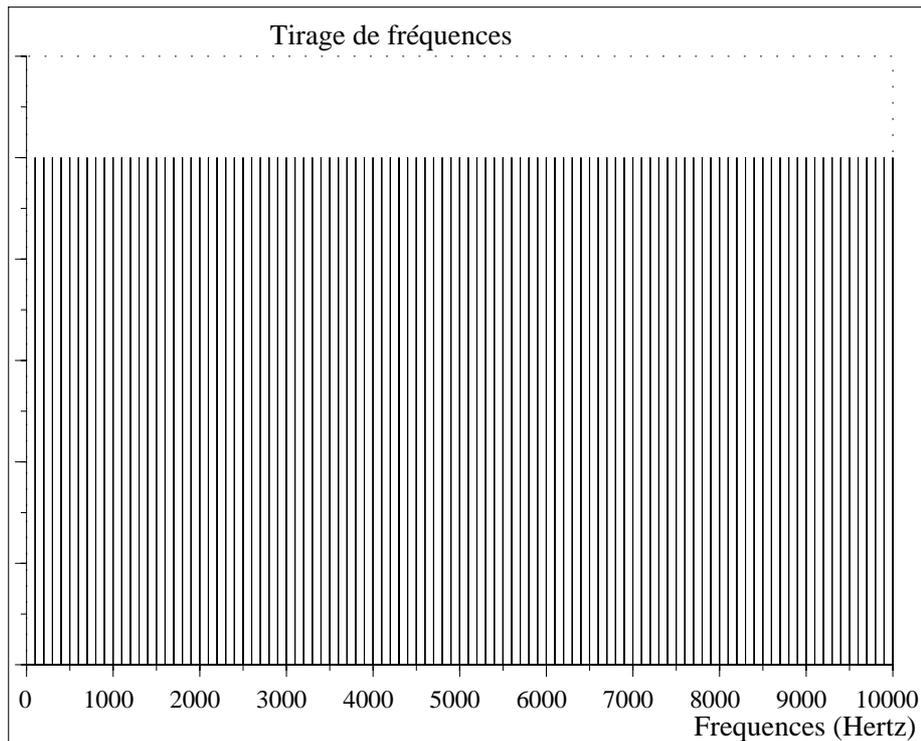
Le paramètre M n'est à priori pas lié à N et W , et donc à la notion de densité spectrale (voir partie 5.2.1). Seulement nos expériences montrent que, dans le cas d'une *bande de fréquences assez large* (plusieurs kHz), et pour une densité spectrale constante (selon notre définition), modifier la valeur de M , et plus particulièrement la choisir beaucoup plus importante que N (approximativement à partir de $M=4N$), permet de synthétiser un bruit équivalent sur le plan de la perception à un bruit dont la *densité spectrale* perçue est moindre. Ce phénomène peut être parfaitement expliqué. En choisissant M plus important que N ($L=1$), le nombre de casiers est très grand. Ceux-ci se retrouvent donc de largeur réduite. Comme le tirage sur les casiers est uniforme, il y a de fortes possibilités pour que plusieurs casiers consécutifs ne soient pas tirés, provoquant ainsi des *trous* dans le spectre qui seront audibles (voir les exemples 2 et 3). Ces *trous* étant différemment placés dans le spectre d'une fenêtre temporelle sur l'autre, une impression de mouvement dans le spectre sera perçue. Il est donc nécessaire de préciser que notre définition de la densité spectrale représente la densité spectrale *moyenne*, et qu'il est important d'étudier les variations de cette densité moyenne dans chaque bande du spectre. Cette remarque est d'ailleurs confirmée par le fait que les courbes de variation des fluctuations d'enveloppe indiquent qu'une augmentation de M accroît ces fluctuations moyennes, pour toutes les valeurs de W (particulièrement donc les grandes) : les *trous* dans le spectre sont, dans ces cas là, plus importants sur le plan de la perception.

5.3.2. Influence sur la perception

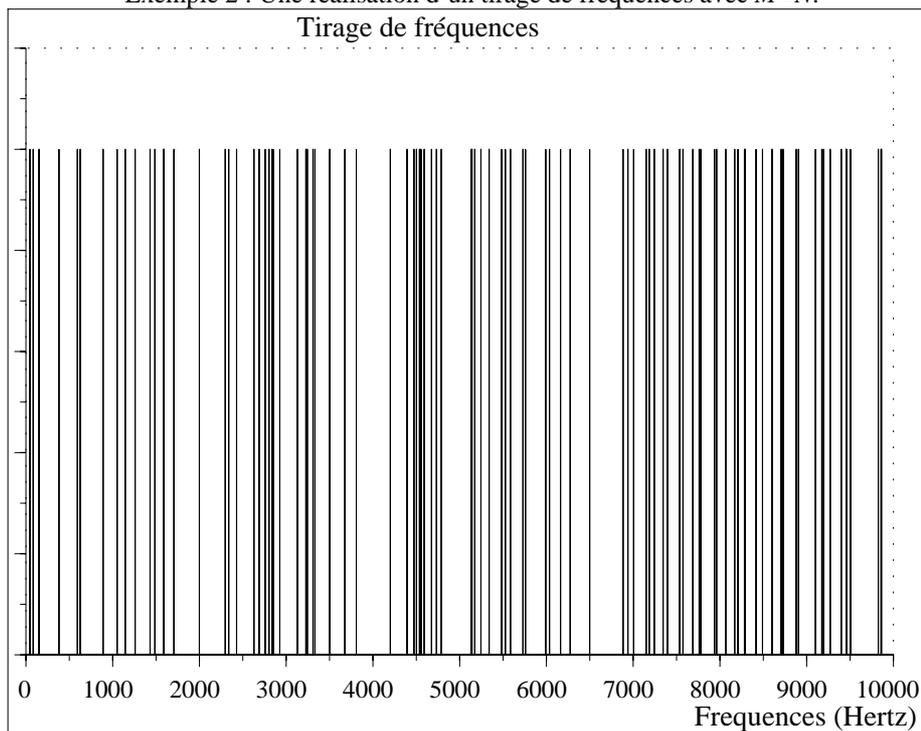
Dans le cas de bandes de fréquences larges, le paramètre M influe donc sur la densité spectrale moyenne du son. Dans le cas de bandes de fréquences plus étroites, les variations de densité dans chaque bande du spectre sont moins importantes sur le plan de la perception, car elles n'entraînent pas de vide suffisamment grand dans ce spectre. Par contre, pour une même valeur de densité spectrale moyenne non maximale, augmenter M permet d'augmenter les fluctuations d'enveloppe dynamique¹¹.

¹⁰ SCHAEFFER P., *Traité des objets musicaux*, Seuil, 1966.

¹¹ HANNA P. and DESAINTE-CATHERINE M., "Influence of frequency distribution on intensity fluctuations of noise," *Proceedings of the Digital Audio Effects Workshop (DAFX'01, Limerick, Ireland)*, pp.125-129, 2001.



Exemple 2 : Une réalisation d'un tirage de fréquences avec $M=N$.



Exemple 3 : Une réalisation d'un tirage de fréquences avec M *infini* : quelques fréquences sont isolées.

5.4. Influence de la largeur de la fonction de densité de probabilité des fréquences

Le paramètre L représente la largeur relative de la densité de probabilité des fréquences de chaque composante du signal synthétisé. Il caractérise en partie le caractère aléatoire de la distribution des fréquences.

5.4.1. Influence sur la perception

L'influence de ce paramètre L n'est audible que dans le cas de densités spectrales non maximales. Si le nombre de composants fréquentiels est suffisamment important par rapport à la largeur de la bande de

fréquence et à la taille de la fenêtre de synthèse (voir partie 5.5), et que le nombre de casiers est égal ou très peu supérieur à ce nombre de composants, le choix des fréquences à l'intérieur de chaque casier n'a, à priori, aucune influence. Par contre, dans le cas d'une densité spectrale faible, le choix de L influence la variance de l'écart entre des fréquences de deux casiers consécutifs. Ainsi, une valeur de $L=0$ supprime le caractère aléatoire et permet de placer chaque fréquence à l'extrémité de chaque casier. Les N fréquences seront alors tirées parmi les M valeurs possibles : $F_k=kW/M$. Cette propriété implique que l'écart entre deux fréquences consécutives sur le spectre sera constant (dans le cas $N=M$), voire harmonique (dans le cas $N<M$). Ainsi, l'enveloppe dynamique sera *périodique*, et cette régularité pourra être perçue sous la forme d'une *hauteur*. À l'inverse, une valeur de $L=1$ atténue ce caractère périodique puisque les écarts entre deux fréquences consécutives seront peu ou pas harmoniques.

5.4.2. Paramètre musical

Du point de vue musical, le paramètre L permet donc de contrôler l'*harmonicité* des mouvements de l'enveloppe temporelle, mais aussi la régularité des variations de la densité spectrale. Ainsi diminuer L peut amener à synthétiser des bruits réguliers, qui peuvent faire penser à des bruits issus de *machines*. Par contre, augmenter L amènera à générer des sons rappelant des *liquides*, à cause des mouvements spectraux induits.

5.5. Taille de la fenêtre de synthèse

Dans notre approche, le signal synthétisé est une suite de signaux dont la durée est précisée par un paramètre indépendant : la taille de la fenêtre de synthèse. Cette longueur représente donc une durée de temps. Elle a une influence importante sur la *périodicité* du signal et sur celle de l'enveloppe dynamique. En effet, comme cela a déjà été dit dans la partie 5.2.2, cette taille est l'explication de l'absence de perception de périodicité dans le cas de la synthèse de la partie stochastique du modèle SMS. Sur le même principe, un choix de taille de fenêtre relativement courte peut permettre d'éviter de percevoir des périodicités dans certains cas.

5.6. Enveloppe spectrale

Dans notre approche comme dans celle du modèle SMS, nous introduisons comme paramètre l'*enveloppe spectrale* ou *couleur*¹². Nous ne reviendrons pas en détail, dans ce papier, sur ce paramètre souvent décrit avec précision. Il est toutefois important de noter ici l'indépendance entre cette couleur et l'enveloppe dynamique et plus précisément ses fluctuations. En effet, la couleur est une fonction continue, et les valeurs des amplitudes de chaque composant fréquentiel est déduite de cette fonction. Mais le calcul permet de montrer que seules les *différences* entre les fréquences rentrent en compte dans la définition des fluctuations d'enveloppe, et donc dans la densité spectrale.

Il est donc possible, pour une même enveloppe spectrale, de définir plusieurs séries de fréquences différentes (modifications des paramètres décrits ci-dessus), c'est-à-dire différentes densités spectrales. Une fois les fréquences choisies, leur amplitude respective sera définie par le paramètre couleur, c'est-à-dire l'enveloppe spectrale.

¹² MARCHAND S. and DESAINTE-CATHERINE M., "Structured Additive Synthesis : towards a model of sound timbre and electroacoustics music forms", *Proceedings of International Computer Music (ICMC'99, Beijing, China)*, pp.260-263, 1999.

6. Discussion

6.1. Efficacité

L'algorithme présenté dans ce papier a été implémenté en langage C. Pour l'instant, les efforts de recherche n'ont pas porté sur la rapidité de la synthèse, mais plutôt sur la pertinence des paramètres définis. Cependant il est possible de penser que cette méthode peut être implémentée en temps réel. En effet, elle repose en grande partie sur une méthode classique de synthèse additive. Or le nombre de composants fréquentiels à synthétiser est certes important, mais raisonnable compte tenu de l'évolution de la puissance des machines actuelles : par exemple un bruit blanc filtré peut être synthétisé à l'aide de 256 oscillateurs, en considérant une fenêtre de synthèse de 1024 et un taux de recouvrement de 0,5. Le travail de recherche restant à accomplir concerne surtout la prise en compte efficace des changements des paramètres proposés et les différents tirages des fréquences et des phases de chaque partiel de la synthèse additive d'après ces mêmes paramètres.

6.2. Conclusion

Nous proposons donc une méthode de synthèse de sons bruités, basée sur une approche spectrale et statistique, permettant de contrôler indépendamment l'enveloppe spectrale et la densité spectrale. Outre l'implémentation en temps réel de cette approche, le travail restant à accomplir est relatif surtout à la mise au point de méthodes d'analyse de sons réels, permettant de définir le plus précisément possible les paramètres présentés ici. Cependant, en l'état actuel, les techniques suggérées dans ce papier peuvent toujours permettre de mettre au point un outil d'exploration du *timbre* des sons bruités complexes, mettant en avant des propriétés impossibles à prendre en compte dans les modèles de bruits blancs filtrés, tout en restant homogène avec les modèles développés pour les sons harmoniques.