[Date]

Colmant Amaury - Lecomte Emerick

Esgi - Mémoire

Génération d’un environnement réel via un flux sonore

Table des matières

[I. Analyse de l’existant 3](#_Toc443989694)

[A. Design sonore 3](#_Toc443989695)

[Qu’est-ce que le Design sonore ? 3](#_Toc443989696)

[Technique d’analyse existante 5](#_Toc443989697)

[B. Level Design 8](#_Toc443989698)

[Les composantes scénographiques. 8](#_Toc443989699)

[Qu’est-ce qu’un environnement virtuel ? 11](#_Toc443989700)

[C. Projet existant liant son et graphisme 13](#_Toc443989701)

[Les différents types 13](#_Toc443989702)

[II. Réalisation 15](#_Toc443989703)

[A. Traitement du son 15](#_Toc443989704)

[1. La phase pré-analytique 15](#_Toc443989705)

[2. La phase Analytique 17](#_Toc443989706)

[3. Les outils d’analyses 18](#_Toc443989707)

[4. Implémentation et Optimisation 19](#_Toc443989708)

[B. Génération de l’environnement 22](#_Toc443989709)

[1. Génération d’un environnement réel 22](#_Toc443989710)

[2. Génération via un flux audio 25](#_Toc443989711)

[C. Calibrage / Adaptation au joueur 28](#_Toc443989712)

[III. Analyse Marketing 28](#_Toc443989713)

[A. Cible 28](#_Toc443989714)

[Lexique 28](#_Toc443989715)

# Analyse de l’existant

## Design sonore

### Qu’est-ce que le Design sonore ?

Selon Wikipedia :

« Le **design sonore** ou la **conception sonore** est l'art d'utiliser des éléments sonores afin d'obtenir un effet désiré. » <https://fr.wikipedia.org/wiki/Design_sonore>

Nous croisons son résultat tous les jours : lorsque nous regardons un film, écoutons une musique, allons au théâtre et, bien entendu, dans les jeux vidéo. Le travail du designer sonore consiste à arranger la musique, à la créer, à la nettoyer, mais aussi à la faire coller à un environnement donné. Par exemple, lors d’une scène triste dans un film, vous ne voudriez pas entendre une musique d’action derrière. Le designer sonore doit s’adapter au ressenti émotionnel dégagé par un lieu, un événement, un moment, ... pour y mettre au mieux en relation ces sonorités et ainsi pouvoir y dégager l'émotion qu'il souhaite transmettre.

Bien entendu chaque être humain perçoit les choses différemment, et pourtant le designer sonore doit s’appuyer sur des codes de conventions communes pour sa création.

Ces codes, quels sont-ils ?

* Le *rythme* de la musique doit être adapté au rythme de l’action. Par exemple, prenons une course poursuite, le nombre d’actions à la seconde est très important c’est pourquoi il faudra adapter la *pulsation*[[1]](#footnote-1) de la chanson en fonction de cette fréquence.
* La « teinte » de la musique doit coller avec l’environnement visuel. L'effet sonore dans une église ne sera pas le même que celui dans un théâtre, une arène ou au grand air, nous y entendrons plutôt des bruits avec échos.
* Le choix des instruments est également important, les *timbres* de ces derniers viendront imprégner des informations lier à la culture. De plus, dans un contexte historique, par exemple au moyen âge, il n'est pas naturel d'y entendre des instruments récents qui provoqueront ainsi un paradoxe temporel et donc potentiellement aller à l'encontre de l’ambiance souhaité.

Il existe d’autres codes mais ceci nous permet de définir une base commune de réflexion. Il existe une sorte de consentement implicite mutuelle entre les individus pour ces règles. Nous pouvons nous poser la question de savoir d’où viennent ces règles ?

La première réponse serait de dire que nous avons été éduqués comme cela. Effectivement, ces règles existent depuis assez longtemps pour que nous ayons toujours vécu avec sans pour autant que personne ne nous les ai énoncées. La plupart des films, pièces de théâtre, jeux et musiques les respectent et donc tout au long de notre éducation musicale, nous les avons assimilées implicitement. Nous « baignons dedans » et même si, a première vue, un individu dit ne pas faire attention aux musiques de film, il les entend quand même et son cerveau les assimile automatiquement, tout comme plein d’autres détails.

L’exemple le plus parlant nous vient des enfants en bas âges (3-5ans). Au départ, les enfants ne distingue pas les différences musicales, ils ne feront pas la distinction entre une musique rapide ou lente. Leur éducation musicale en étant qu’à son balbutiement, il ne serait pas gêné par une musique dansante pendant une scène triste dans un film. Cependant, même seuls, ils vont apprendre sans que l’on ai besoin de leur donner les codes. Cela donne un lien vers la deuxième réponse : la cohérence vis-à-vis du monde.

Dans la vie de tous les jours, nous entendons des milliers de bruits. Nos oreilles captent une « musique permanente ». Que cela soit le bruit du vent et des oiseaux chantant (auxquels nous ne prêtons pas forcement attention), ou les bruits de la ville. Nous sommes constamment entourés de sons et notre cerveau assimile et associe encore une fois ces bruits à un environnement donné. Nous allons alors associer un type de bruit, une impression générale à une vision. C’est pourquoi une musique lente, avec des bruits aigus (dans un certains *spectre*) comme des musiques celtes nous ferons penser à la nature. Alors qu’une musique rapide et plus électrique sera associé à un milieu urbain, qui correspond à notre état d’esprit et aux images assimilé par notre cerveau. On dit que la musique fait voyager, c’est en partie dû à cette association mais aussi à la culture.

Intuitivement, nous saurons reconnaitre (même sans en avoir beaucoup écouté) une musique venant des pays africains, asiatiques ou d’autres continents. Les cultures des différentes civilisations s’entendent à travers la musique. Ceci est dû à l’éducation, a l’environnement visuel mais aussi aux habitudes des civilisations et à leurs codes musicaux. La culture asiatique utilise beaucoup d’instruments que nous n’utilisons pas en Europe de l'ouest, tout simplement car notre culture ne nous a pas habitués à ça. C’est pourquoi la plupart des films asiatiques sont remasterisés pour correspondre aux codes du pays d'exportation avant de sortir chez lui.

Tous ces éléments nous ont permis de dresser une liste de lois implicites autour des codes de la musique. Bien entendu, celles-ci dépendent de la culture, c'est pourquoi les règles citées ci-dessus s’appliquent à toutes les cultures, elles sont le reflet du monde dans lequel nous vivons.

### Technique d’analyse existante

Le son étant un art faisant appel, comme tout art, à la sensibilité de l’être humain, il est de fait, subjectif. C’est pourquoi le travail d’un designer sonore est requis pour ce genre de production et qu’il est difficile de l’automatiser. Cependant, le son est avant tout un signal. Pour rappel, le son est créé par des vibrations à différentes fréquences. Nous allons distinguer 3 catégories de sons, les infra-sons (en dessous de 20Hz), les sons audibles humainement (entre 20Hz et 20 000Hz) et les ultra-sons (au-dessus de 20 000Hz). Nous intéressant à la musique, nous ne traiterons que les sons audibles. Cependant, avec un ordinateur nous pourrions traiter de la même manière les infra et ultra-sons. D’un point de vue informatique, le son sera présenté sous la forme d’un tableau à 2 dimensions mettant en parallèle la fréquence et la nuance relative. Pour exemple prenons un son constant à 1 000 Hz, nous pourrons exprimer sa puissance entre 0 (aucun son) et 1 (puissance de sortie dans les enceintes du son). Donc à 0.2 le son sera le même qu’à 0.7 mais moins audible. La combinaison de ces nuances et de ces fréquences sur le spectre audible permet de concevoir des musiques.

Cependant, tous les sons ne sont pas perçus de la même façon par l’oreille humaine. Prenons par exemple les basses (entre 20 Hz et 160 Hz voir 312 Hz en prenant large), ces dernières produisent plus de bruits audibles par l’oreille. C’est pourquoi produire un son a 20 Hz et d’une puissance 0.5 sera plus audible qu’un son a 2 000 Hz a 0.5. Ceci à un impact non négligeable sur l’analyse sonore. Intuitivement, l’homme peut dire si une musique à une sonorité grave ou aiguë et, ceci étant subjectif, tout le monde peut ne pas être en accord avec lui. Mais la machine, ne disposant pas de sensibilité mais simplement d’informations, ne peut pas déterminer la teinte d’une musique. Il faudrait pour cela définir pour chaque fréquence un taux de conversion pour niveler une musique.

La plupart des oreilles expérimentées arrivent à reconnaitre les instruments dans une musique. Bien souvent parce leur utilisation est soit différée dans le temps (quelque secondes voir millisecondes) ou bien parce qu’elles n’ont pas la même sonorité (exemple : batterie et clavier ne produisent pas des sons dans les mêmes fréquences). Pour l’ordinateur, il est impossible de les distinguer sans une intervention humaine. Cette limite est due au fait que les fréquences des instruments peuvent se superposer et donc nous ne pouvons assurer que sur une plage définie de fréquences, celles-ci feront toujours référence au même instrument. Mais dans le cas où nous arriverions à séparer les instruments, les notes pourraient être retrouvé grâce au la fréquence fondamentale. La fréquence fondamentale est définie par l’instrument, cela sera le multiple sur lesquels les notes seront calquées. Prenons pour exemple une corde de piano qui a pour fréquence fondamental 440Hz, toutes ses notes seront exprimées en multiple de 440Hz.

Maintenant que nous avons pu voire les limites de l’analyse nous pouvons déterminer que la seule façon d’analyser une musique est sur ces données pure. Une des analyses les plus communes est l’analyse de la *pulsation*.

* La *pulsation* définit en partie le rythme et il est souvent donné principalement (mais pas que) par un instrument à percutions. La *pulsation* peut être analysée en regardant l’énergie dissipée au cours du temps par une musique sur une bande.
* Les filtres « passe haut », « passe bas » et « passe bande ». Ils permettent de filtrer une partie de la musique et donc d’en extraire une partition de la fréquence. Cet outil sert notamment pour analyser la *pulsation*.
* Une autre analyse possible est aussi le développement d’intensité dans le temps. Prenons une ou plusieurs fréquences données et regardons leurs évolutions d’intensité dans le temps.
* Analyse de l’*harmonique* en fonction de la gamme utilisée pour la détection de suites de notes caractéristiques d’un genre de musique ou d’une tonalité.

Ces éléments permettent de façon simple ou combinés, de déterminer un bon nombre de paramètres pour analyser la musique et la classifier. Nous pouvons définir le genre, le rythme, et filtrer certaines fréquences. Nous développerons cette partie plus tard mais, à titre d’exemple, nous pourrions imaginer que dans un environnement désertique, une tempête fait rage et la puissance de cette tempête serait fonction du beat de la musique courante.

## Level Design

En jeu vidéo, le *level design* est l’étude et la construction de la structure d’un niveau de jeu par un Level designer. Il doit s’adapter à ou définir la mécanique de jeu que nous appellerons plus communément *gameplay*. Cette définition est spécifique au jeu vidéo, cependant, il est possible de l’appliquer dans d’autre domaine. Au théâtre, le metteur en scène avec le responsable des décors fait lui aussi du *level design*, il pose les bases d’un environnement, il en va de même pour le cinéma. Au final, nous pourrions définir le *level design* comme la création de l’environnement visuel pour le spectateur.

Dans notre cas nous allons nous intéresser aux procédés déterminant dans la création d’un environnement pour le ressenti émotionnel du joueur. Ici, il ne s’agit pas de définir un *level design* qui va impacter la manière de jouer du joueur mais un *level design* qui va **impacter la vision émotionnelle** du joueur. Nous serons donc plus proches de la mise en scènes tel quelle pourrait être créée dans le cinéma ou le théâtre que dans le jeu vidéo. C'est pourquoi nous nous baserons sur le travail d'un *scénographe* pour définir celui du *level designer*.

### Les composantes scénographiques.

Prenons la définition wikipedia de la *scénographie* :

« La **scénographie** (du grec σκηνη (skene) *scène* et γραφειν (graphein) *écrire*) désigne aujourd'hui l'art de l'organisation de l'espace scénique, grâce à la coordination des moyens techniques et artistiques. »

Ainsi nous pourrons noté dans un premier temps que le *scénographe* a la possibilité de manipuler cinq composantes technique de la manière la plus cohérente possible afin d'en dégager l'impression émotionnel recherché. Il compose avec des volumes, des objets, des couleurs, des lumières, et des textures.

1. Les Volumes

Avant toute chose, il y a la définition des volumes. C'est ce qui va permettre de cadrer les proportions de toutes choses dans un espace donné. Ceux-ci sont donc basés sur la différence entre la vision de l'utilisateur/spectateur et les limites de l'espace auxquelles ce dernier est confiné.

1. Les objets

"Objets" est un terme général qui désigne toutes choses contenue dans un volume. La taille de ces derniers correspond de manière proportionnelle, bien entendu, au volume auquel ils appartiennent. Les objets de par leur présence ou absence parlent à leur manière. Aucun objet n'est jamais muet, mais il faut parfois une analyse assez minutieuse, pour en décrypter tout le message particulier. Les objets symbolisent donc l'essence même de la scène, ce sont eux qui économisent le narratif propre aux différents genres littéraires ou bien ici, dans le cas d'un jeu vidéo, à un scénario. Un agencement d'objets et d'éléments sont indépendants les uns des autres, mais constituant d’un tout, souvent éphémère, qui permet en un regard de découvrir le lieu et la temporalité de l'action.

1. Les couleurs

Les couleurs sont d'une complexité lorsqu'il s'agit de comprendre ce qu'elles représentent. Mais une fois admises, elles représentent un suprême potentiel de transmission de message. Dans un premier temps, il faut considérer que la couleur n'existe que dans notre cerveau. Il ne s'agit en fait que de fréquences qui vont inconditionnellement agir sur l'ensemble des êtres humains de la même manière. En effet, les analogies de structure de nos rétines, de nos systèmes nerveux vont répondre de manière plus ou moins automatique (par réflexe) à ces fréquences et ainsi agir, entre autre, sur notre rythme cardiaque et d'une manière plus générale, sur notre état d'esprit. C'est donc un bon moyen de conditionnement certes, mais il existe aussi des adaptations personnelles qui expliquent des préférences individuelles. En effet, dans un second temps, les couleurs permettent de véhiculer un message en fonction de la culture, l'éducation et l'expérience du monde et la nature qui l'entoure. Le transfert de l'image photographiée par l'œil, emprunte de nombreuses voies jusqu'à être transmise au cortex. Celui-ci analyse l'image et utilise les zones frontales associatives qui, en fonction des acquis mémorisés, vont tirer les conséquences pour alimenter, en réponse, les impressions qui lui semblent subjectivement associées.

1. Les lumières

La lumière est un atout majeur pour déterminer l’environnement dans lequel nous voulons plonger le joueur. D’un coup d’œil, le spectateur peut identifier l’environnement dans lequel il se trouve. Est-il inquiétant ou bien réaliste. Dans quel moment de la journée nous trouvons nous. En intérieur ou bien en extérieur. La lumière permet de passer rapidement toutes ces informations. Pour prendre un exemple, en intérieur la lumière sera plus tamisée et plus artificielle. Elle ne viendra pas du soleil et donc surement plusieurs sources différentes, c’est pourquoi les ombres seront multipliés et les températures (Kelvin) autres que celles du soleil. Cela peut renforcer un sentiment de malaise ou bien au contraire un sentiment de réconfort selon le type de source lumineuse.

1. Les textures

Une des choses qui apportent plus de crédibilité. Les textures sont appliquées aux objets en même temps que le choix des couleurs et de l'éclairage, ils sont tous deux étroitement lié. Après définition géométrique des éléments d'une scène, des propriétés s'appliquent aux objets qui leurs confèrent la rugosité, la couleur, la réflexion etc. Ces ensembles de propriétés sont inclus dans ce qui est techniquement connu sous le nom “Matériel”. Le matériel comprend donc différents concepts associés avec le comportement de la lumière sur l'objet.

Cet ensemble de composantes va ainsi définir le lieu et l'environnement dans lequel se passe l'action, c'est ce que nous appelons le décor. Le décor bien entendu est une composante essentielle de l’identification. C’est pourquoi le décor nous donne des éléments visuels qui nous permettent d’identifier notre environnement, cependant, ces éléments doivent être cohérents entre eux pour définir un environnement. Une télévision n’aurait pas sa place dans un environnement médiéval par exemple. Et pourtant, ce décor permet un certain degré de liberté. Il serait possible de trouver une télévision dans un château, si le joueur explore à notre époque des « ruines ». Cela permet de créer une certaine ambiance dans un contexte établie. Dans un seconds temps, le scénographe vat travailler sur

1. La mise en scène

La mise en scène peut se décomposer en deux parties, l’agencement scénique des décors et la cohérence de ceux-ci. Ces deux éléments permettent de définir l’univers. Prenons par exemple un cactus dans un désert. Jusqu’ici, rien de spécial, il est normal de trouver ces éléments ensembles. Maintenant prenons un cactus qui se déplace dans un désert. Ici, le spectateur sait directement que l’environnement se place dans un contexte onirique ou surréaliste. Le décor et/ou la lumière ne lui apprennent rien là-dessus, c’est bien la mise en scène qui s’en occupe et plus particulièrement la cohérence par rapport à l’environnement habituel que nous connaissons tous. Prenons maintenant comme exemple une salle, comme une salle à manger. Si tous les éléments sont au sol, le contexte semble être réel, mais sinon plaçons ces éléments sur le plafond, le contexte se modifie. L’agencement permet de définir ce contexte. Il peut parfois rejoindre la cohérence.

1. Espace scénique

L’espace scénique concerne pour nous l’endroit où le joueur pourra ou non se déplacer. Pourra-t-il traverser les murs et voler ? La réponse à ces choix impacte la vision que le joueur aura de lui-même et donc l’appréhension de son environnement. Est-il humain ou fantôme ? Dans le deuxième cas nous nous trouvons dans un environnement surréaliste.

Ces composantes vont nous aider à définir l’environnement du joueur et donc l’impact des choix que nous faisons sur lui. L’émotion ressentie par le joueur ne sera pas la même dans un environnement réel que dans un environnement imaginaire.

### Qu’est-ce qu’un environnement virtuel ?

Un environnement virtuel est une représentation d’une mise en scène, d'un décor de façon informatique. Principalement, il s’agit du croisement d’un jeu vidéo et d’un réseau social. Il s’agit d’immerger le spectateur dans un environnement réel et lui laisser le libre de choix de ce qu’il va en faire. L’environnement virtuel le plus connu est le « jeu » : Second Life. Ici, le joueur est représenté par un avatar. Ce jeu pourrait être vu comme un simulateur de vie ou le joueur peut effectuer une partie des actions qu’il ferait dans la réalité, comme par exemple travailler, allez boire un verre avec ces « amis », etc… C’est ici que nous voyons la dimension réseau social arriver. Le but est d’interagir avec les autres joueurs du monde. Outre cela, le joueur peut être amené à visiter différentes régions comme dans un jeu MMORPG (massivement multi-joueurs online role play game).

Le but d’un environnement virtuel est de reproduire la réalité. Comme vu précédemment, les contextes doivent donc être réels au niveau des composantes scénographiques.

## Projet existant liant son et graphisme

### Les différents types

Les projets de jeu vidéo, liant le son et les graphismes, se font assez rares pour la simple et bonne raison que l’analyse d’un signal désordonné a du mal à produire un graphisme ordonné de façon automatique. La musique a un impact émotionnel qu’il est difficile de traduire pour un ordinateur. La machine ne peut pas ressentir des émotions, c’est pourquoi il est nécessaire d’avoir une intervention humaine. Cependant, certains s’y sont essayés. Bien souvent dans le domaine de l’abstrait ou alors dans un but rythmique de gameplay.

1. Les jeux abstraits

Dans ces jeux, le but est de générer un environnement virtuel en fonction de paramètres ou d’actions. Dans cette catégorie nous pourrions citer :

Panoramical : Jeu dans lequel il faut contrôler via des jauges, la puissance de certaines fréquences et qui génère plus ou moins un environnement.



Cet environnement n’a rien de réaliste, il relève plus du concept artistique que d’un environnement virtuel, de plus la seule interaction possible est celle de modifier les fréquences.

1. Les jeux basant leur gameplay sur la musique

Cette deuxième catégorie met en avant différents gameplay possibles en rapport avec la musique.

Guitar Hero : Jeu de rythme dans lequel nous possédons une guitare réelle avec des couleurs associées aux boutons, il s’agit d’appuyer sur les boutons au bon moment pour que la note se joue correctement. Ici, si vous vous trompez de note, la musique fait une fausse note et l’environnement (les spectateurs) réagissent en conséquence.

Flower : Jeu poétique qui prend place dans un environnement réel, une prairie. Vous jouez un étal de fleur volant au vent et qui doit collecter d’autres pétales auprès des fleurs. A chaque fois que vous ramassez un pétale cela joue un son, à vous de trouver le bon rythme pour produire une jolie musique.



Crypt of the Necrodancer : Ici nous nous rapprochons plus de notre concept qui est d’analyser la musique pour en tirer des évènements. Necrodancer vous laisse jouer avec vos propres musiques. Il va analyser ces dernières pour en déterminer le *beat*. Vous ne pouvez effectuer vos actions que sur un *beat* de la musique. Les ennemis et vous doivent avancer au rythme de la musique. Il s’agit d’un jeu de type dungeons crawler ou le but est de se déplacer de salle en salle tout en éliminant les ennemis.

# Réalisation

## Traitement du son

Problématique : Comment est-il possible d’extraire le son et ces composantes fondamental d’analyses ?

### La phase pré-analytique

#### Définition

Pouvons-nous résumer une chanson à une simple suite de note de musique prise au hasard dans un tout ?

Commençons par définir ce qui résume une chanson. Distinguons deux points de réflexion :

* Du point de vue humain
* Du point de vue analytique

Du point de vue humain, ce qui résume une musique est ce qui va lui rappeler cette même musique. Prenons pour exemple : « Cette soirée-là » de Yannick, et « Cette année-là » de Claude-François, la première étant une reprise de la seconde. Ici nous pouvons dire que pour les gens qui connaissent l’original de Claude François, la première chanson est un résumé de la seconde, car malgré des « instruments » différents la première reprend les mêmes notes et sonorité que la seconde et donc fait appel au souvenir de la chanson original.

D’un point de vue analytique ces deux chansons ne seront pourtant pas semblable car les fréquences utiliser ne seront pas les mêmes (malgré que les notes soit les mêmes un La en guitare ne sera pas le même qu’un La en flutes). Nous pourrons toujours définir la suite de note en prenant la fréquence fondamental et en expriment les notes en fonction de celle-ci. De façon analytique deux musiques seront semblable si elle utilise les mêmes notes répétées (d’un point de vue légal une suite de 8 mêmes notes est considérée comme un plagia).

Ce qui définit une musique est donc une suite de note consécutive. Maintenant, il nous faut savoir d’où nous pouvons extraire cette suite de notes ?

Lorsque que l’être humain se souvient d’une chanson c’est souvent par son refrain. Le refrain étant une suite de note revenant à intervalle régulier sert souvent caractérisé une chanson car de son aspect répétitif elle marque plus facilement l’utilisateur. Mais le refrain seul n’est pas suffisant, car il n’est pas toujours présent et selon la taille de la chanson peut avoir un impact moindre. C’est pourquoi la meilleure façon de caractériser une chanson est en déterminant la plus longue suite de note répétée dans une chanson.

#### Transformation du son en spectre

Les fichiers sonores courant sont des fichiers binaires. L’analyse musicale passe par l’analyse d’un spectre sonore. C’est pourquoi il va nous falloir transformer ces fichiers audio en spectre audio.

*Qu’est-ce qu’un spectre sonore ?*

Le spectre sonore est un graphique tri dimensionnel qui met en correspondance une fréquence avec une amplitude. Cette amplitude définie la puissance d’une fréquence a un moment donnée. Nous avons donc l’évolution au cours du temps de la puissance d’une fréquence.

Attention : l’amplitude d’une fréquence ne définit pas sa puissance de sortie dans les enceintes (les décibels). Effectivement les fréquences étant différentes il faudra une amplitude plus ou moins élevé selon la fréquence.

Dans un premier temps il va nous falloir décompresser le fichier audio. Vu le nombre de fichier existant, nous n’allons pas tous décrire leur système de décompression, cependant voici un exemple pour les fichiers MP3 (MPEG ½ audio layer 3 de son nom complet).

Le fichier MP3 contient un fichier d’entête qui permet de définir son contenu. Ce header a une taille fixe de : 32 bit soit 4 octet. Dans lesquels se trouvent différentes informations comme le taux d’échantillonnage (sample rate) qui définit la plage de fréquence audio (cf. spectre sonore), le type de sortie : mono, stéréo, dual Channel, etc… Mais aussi diverses informations sur la musique en elle-même : copyright, style de musique et auteur.

(Source : <http://mpgedit.org/mpgedit/mpeg_format/mpeghdr.htm>)

Pour la décompression nous allons surtout nous intéresser au taux d’échantillonnage et au débit (bit rate) qui nous permettrons de définir le nombre de bits pour un état du spectre dans le temps.

La formule étant la suivante : (en fonction du layer définie dans le header)

For Layer I files us this formula:

FrameLengthInBytes = (12 \* BitRate / SampleRate + Padding) \* 4

For Layer II & III files use this formula:

FrameLengthInBytes = 144 \* BitRate / SampleRate + Padding

Nous savons maintenant que nous pouvons lire le reste du fichier et découper par tranche de taille FrameLengthInBytes pour décoder notre musique. Nous savons que dans une « Frame » nous avons « SampleRate » fréquence soit nous avons donc FrameSize = FrameLengthInBytes / SampleRate. Ce qui nous permet de définir un chiffre réel (compris entre -1 et 1 selon le système de décompression). Le moyen le plus utiliser pour une décompression rapide est les transformations rapide de Fourier (Fast Fourier Transform : FFT). Qui permet de passer d’une représentation binaire en une sinusoïde discrète pour un coup en mémoire moindre.

### La phase Analytique

#### Analyse Globale

Comme dit précédemment, la musique peut être caractérisée par une suite de notes répétées. Cependant ce n’est pas la seule donnée qui peut être extraite dans une analyse globale du son.

* Le tempo : Le tempo défini le rythme de la musique ou autrement le nombre de battements par minutes. Il ne peut être évalué précisément que par une analyse globale de la musique. Nous devons réussir à déterminer les « pics » de son par rapport à une moyenne sur un temps donné. Prenons par exemple une grosse caisse de batterie simple frappant à intervalle régulier toutes les secondes.
* La fréquence fondamentale : en analysant la source entièrement nous pouvons déterminer la fréquence la plus utilisé dans la chanson.
* La gamme de note : il suffit d’exprimer les fréquences en fonction de la fréquence fondamentale

Pour extraire ces informations nous avons besoin d’analyser le spectre sonore entièrement. C’est donc une analyse qui ne peut pas se réaliser en temps réel (en même temps que la musique est jouée). Il nous permettrons de définir une tendance globale sur la musique. La gamme de note nous permettrait d’identifier les styles de musiques, par exemple les musiques de l’été ont très souvent les même notes utilisés.

#### Analyse Spécifique d’une frame

Lors de l’analyse spécifique d’une frame nous ne possédons les informations de la frame actuelle et des frames précédentes s’il y en a eu. Les analyses ici sont surtout des analyses d’évolutions dans le temps :

* Evolution dans le temps (diminution, augmentation de l’amplitude)
* Détection d’un battement (beats) si le son est X fois plus puissant que la moyenne des Y dernières frames, dans le même style détection d’un creux dans le son
* Evolution globale de la somme des amplitudes (le son a-t-il tendance à être plus puissant ou moins puissant dans le temps)

Combiner à une analyse globale, ces fonctions vont nous permettre d’affiner la visualisation en temps réel, et permettre de coller à un instant T à ce que l’utilisateur entend.

### Les outils d’analyses

#### Les filtres

Le son était des signaux, il est possible de définir des filtres au niveau des fréquences :

* Les filtres passe haut et passe bas : qui permet de ne laisser passer qu’une fréquence supérieur ou inférieurs à un seuil données. C’est filtre vont s’avérer très utiles les fréquences basses et hautes n’ayant pas les mêmes puissances sonore si nous voulons déterminer l’impact sur l’utilisateur nous devront souvent user de ces filtres pour comparer les deux puissances sur un pied d’égalité
* Les filtres passe bande : qui sont la combinaison d’un filtre passe haut et un filtre passe bas permet d’extraire une fréquence dans un intervalle définie.

Outre les filtres, nous pouvons aussi avoir besoin de déclencheurs (triggers) qui émettront un signal positif si le son dépasse ou est inférieure à une certaine amplitude, dans le cas contraire ils émettront un signal négatif. C’est trigger pourront aussi servir pour analyser la tenue d’une fréquence dans le temps, par exemple, nous pourrions définir un trigger qui se déclencherait si la fréquence est tenue dans le temps.

#### Les Descripteurs

La famille des descripteurs sert à décrire l’évolution ou l’état actuel du signal. Prenons par exemple les variations descripteurs, il peut être intéressant de savoir si sur une période donnée la musique a plutôt eu tendance à augmenter/diminué ou varié constamment.

Ces outils-là se place comme un inspecteur. Ils vont stocker à chaque frame la valeur de la fréquence ou des fréquences sélectionner selon les modèles mathématiques classiques.

### Implémentation et Optimisation

#### Implémentation

D’un point de vue purement fonctionnel nous allons utiliser la librairie FMOD. Cette dernière étant très utilisée en Jeu vidéo notamment dans les moteurs Unity et Unreal Engine. Ce choix se base sur 3 caractéristiques :

* Possibilité de son environnent
* Possibilité de son 3D
* Lecture de différent format (+ normalisation)

Le Son environnent se démarque du son 3D par le faite que celui-ci ne sera pas spatialisé. La spatialisation est le fait de donné une certaine portée au son pour reproduire un effet de réalité. Si une balle tombe à un endroit, le son partira donc de cette endroit pour ensuite rebondir sur les parois du monde et peut être arriver à l’oreille du joueur. Contrairement au son environnent qui lui sera toujours présent dans les oreilles du joueur. Ici, l’idée est de faire de la musique courante une sorte de musique d’ambiance mais aussi d’ajouter des sons 3D en fonction de évènements sur le terrain.

Pour ce qui est de la lecture et de la normalisation du format, nous voulons laissez le plus de choix possible au joueur, et nous ne pouvons donc pas prédire le format de ses fichiers audio. Les formats de fichier audio sont souvent propriétaire et l’extraction des données est spécifique à chacune. Ainsi avec FMOD nous bénéficions de toutes ces transformations pré faites.

Comme dit plus haut, il va falloir transformer les données binaires en spectre pour pouvoir être analysé. FMOD contient des outils pour réaliser la FFT rapidement.

#### Optimisation

Précédemment, nous avons vu qu’il nous devions pouvoir réaliser deux types d’analyses :

* Analyse en temps réel
* Pré-Analyse

La première optimisation possible est sur la plage des fréquences. Plus la plage sera grande, plus la précision sera grande mais plus le temps de traitement sera long. Reprenons la définition du 1.b, ici la pas de fréquence est défini par le taux d’échantillonnage et la taille de la fenêtre. La taille de la fenêtre va définir la précision que nous voulons en bit pour l’ensemble des fréquences. Soit :

FrequencyStep = windowSize / (sampleRate / 2)

SampleRate / 2 : définie le nyquistRate (en générale 22 000Hz)

On va donc exprimé le buffer de la fenêtre en fonction du nombre de fréquence et nous pourrons ainsi définir le pas de fréquence.

Les paramètres d’une musique étant représentés par un tableau deux dimensions de taille :

* Durée de la musique \* windowSize

Nous pouvons aussi optimiser le temps de calcule en utilisant du multi-threading. Le principe du multithreading est de faire travailler en parallèle ou en pseudo-parallèles le CPU. Ici, n’ayant à traiter que des données nous pouvons facilement paralléliser le traitement de la FFT et ainsi gagner en temps d’exécution.

Pour ce qui est de l’analyse en temps réel, nous allons extraire les données à chaque frame, les traiter avec les mêmes fonctions que vu précédemment, et les passer ensuite a FMOD pour qu’il les joues.

## Génération de l’environnement

### Génération d’un environnement réel

Problématique : Comment générer un environnement cohérent via une donnée aléatoire.

Dans cette partie nous nous intéresseront aux méthodes possibles dans la génération procédurale. Notre outil doit être capable de construire son niveau en fonction de n’importe quelles analyses musicales et donc d’une donnée pseudo aléatoire.

La génération procédurale consiste à placer différentes ressources du monde de façon aléatoire. Cette dernière obéit à des règles prédéfinies ayant une probabilité de sortie différentes.

#### La génération par itération (exemple : no man’s sky) :

No man’s sky est un jeu encore en développement qui a pour ambition d’utiliser la génération procédural pour générer 18,446,744,073,709,551,616 monde différent (source : <http://iq.intel.co.uk/no-mans-sky-procedural-generation/> ) . Leur algorithme se divise en plusieurs parties :

* Création de la planète ou ici carte des hauteurs. Ils vont ainsi définir le relief global de la planète de façon aléatoire. (Perlin Noise)
* Choix des biomes présent sur la planète. En fonction de la carte des reliefs ils vont se baser sur des faits pseudo scientifique pour déterminer les types de paysage que la planète aura : par exemple une carte ayant peu d’eau, aura surement des biomes plutôt arides.
* Placement des éléments de décors (arbres, eau, nuage, neige etc…). Ici, le choix est fait en fonction des biomes défini, et de différent paramètre comme la hauteur ou la distance à l’eau la plus proche. Mais aussi une partie d’aléatoire.
* Pour terminer, il implémente les créatures en fonction de leur besoin vital (air ou eau, carnivore, herbivore).

#### La génération par automate cellulaire

Notamment utiliser dans le jeu Terraria pour la génération de cave, les automates cellulaires sont représenté par une grille de X cases par Y dans laquelle chaque case à un état (souvent « vivant » ou « mort »). A cette grille sont associé un ensemble de règles qui définissent l’état suivant. Prenons pour exemple le jeu de la vie de John Horton Conway défini par les règles suivante :

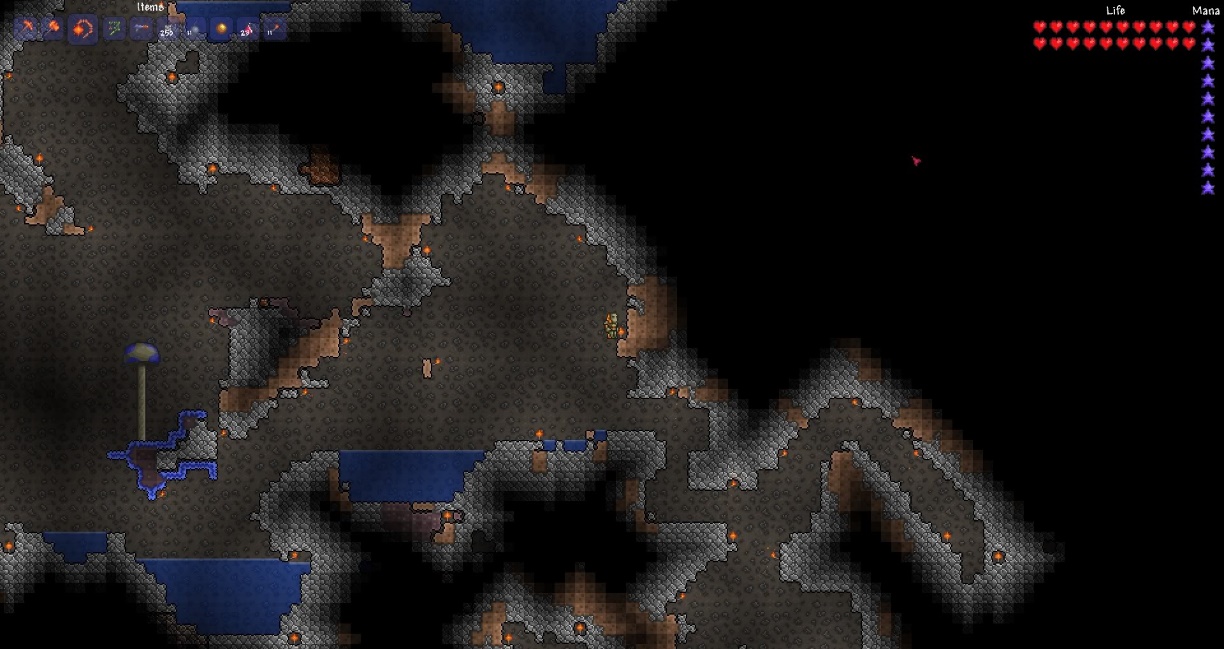
* Chaque cellule vivante avec moins de deux cellules vivantes autour d’elle meurt pour cause de sous population.
* Chaque cellule avec 2 ou 3 voisines vies à la génération suivante
* Chaque cellule avec plus de 3 voisines meurt pour cause de surpopulation
* Chaque cellule morte avec exactement 3 voisine vie à la prochaine génération.

Avec ces règles et un pattern de départ, différents types d’automates arrivent à emerger :

* Automate stable : après une génération X ils n’évoluent plus
* Automate oscillatoire : leur états oscille entre deux ou plus état en boucle.
* Automate « vaisseaux » : leur états sont récurrent après un certain nombre de générations
* Automate « mathusalems » : il est défini comme un petit objet qui explose en de nombreux objets stables.

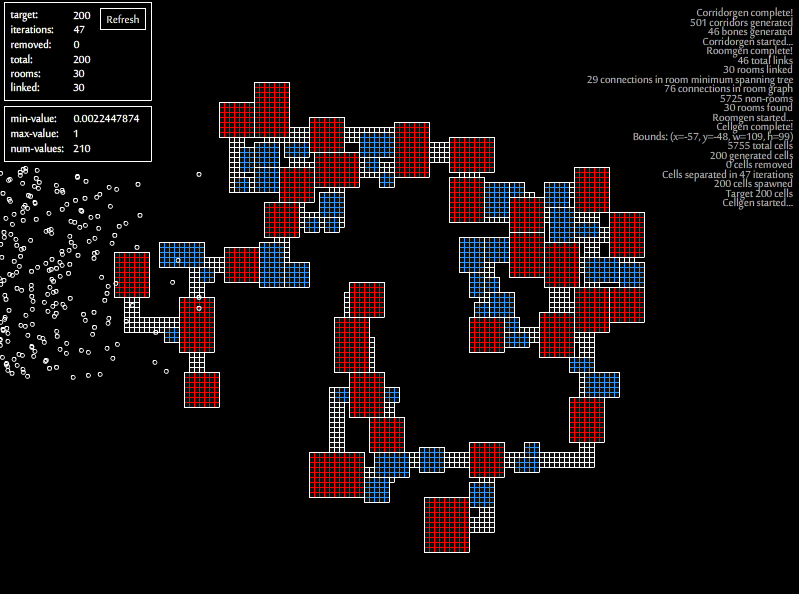
Il existe d’autres structures pouvant apparaitre mais étant plus rare nous ne nous attarderons pas dessus.

Pour en revenir à la génération procédurale une structure stable du jeu de la vie peut amener à des motifs représentant des caves, de vastes espaces plus ou moins irréguliers, comme dans l’exemple ci-dessous.



#### Assemblages de parties prédéfinies.

Cette technique a notamment été utilisée dans la série des jeux Diablo, et plus généralement dans la génération de donjons aléatoires. Il s’agit ici de disposer d’une bibliothèque de salles, avec un marqueur définissant les entrées et sorties de chaque salles, nous allons ensuite mettre bout à bout les salles pour générer un donjon. Cet algorithme permet de gérer plus facilement les choix possibles et les chemins possibles par l’utilisateur et il est moins couteux en mémoire. Cependant il peut aussi être manipulé en forçant la chance.



Ces différentes techniques permettent de générer des environnements virtuels statiques.

#### Monde dynamique

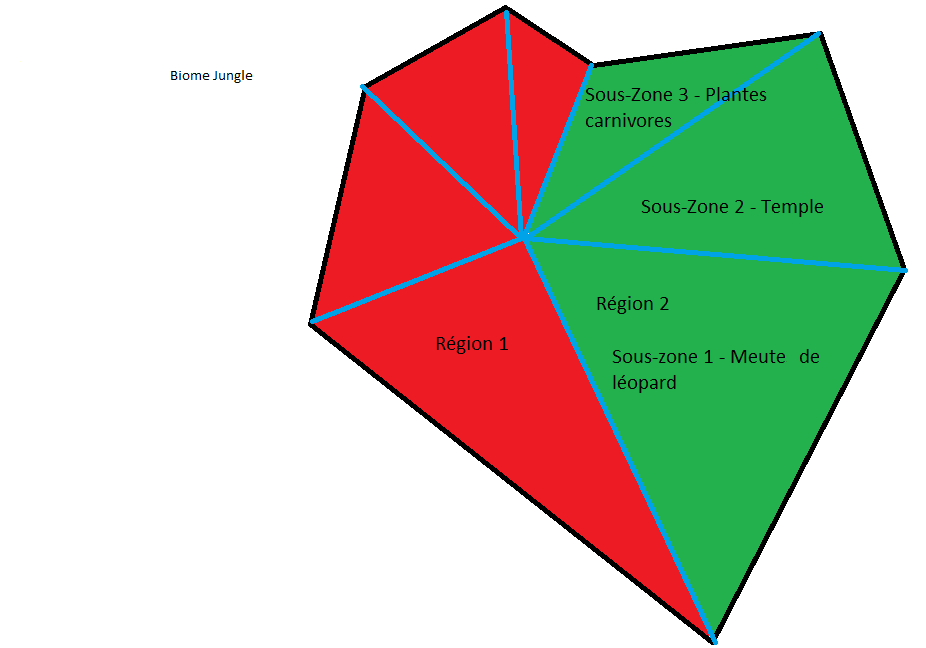
Les mondes dynamiques se construisent au fur et à mesure du besoin du joueur, sur une « seed » communes. Une « seed » est un chiffre généré aléatoirement permettant d’influencer la partie aléatoire de la génération. Ainsi avec une même « seed » nous aurons des mondes semblables mais pas identiques. C’est le cas pour le jeu Minecraft qui génère un monde faisant 8x la taille de la terre. Ce monde n’est pas généré en une seule fois mais dynamiquement en fonction des déplacements du joueur. Pour que le monde reste cohérent, le système s’appuie sur une « seed » qui permettra de générer un monde similaire.

### Génération via un flux audio

#### Définitions Lexical

Pour introduire notre concept de génération d’environnement via un flux audio nous allons devoir commencer par définir les concepts de bases :

* Les biomes : Nous allons définir par biomes les éléments de plus haut niveau, définissant une zone de même type dans leurs caractéristiques globales. Les caractéristiques globales définissent le type de zone tel que désert, jungle, maritimes, urbain etc… Elles nous permettent de définir le type de décors, les types d’évènements, l’éclairage et la skybox. Elles donnent un univers global à une partie de l’environnement crée. Par exemple dans le cas d’un biome pour la jungle nous auront :
  + Densité de la faune et flore
  + Illumination
  + La skybox (le paysage lointain)
* Les régions : Elles sont un niveau en dessous des biomes, elles en prennent donc les caractéristiques globales. Mais subdivise les biomes en plusieurs parties. Les régions servent à définir un ensemble de sous régions ayant les mêmes propriétés au niveau musicale.
  + Disposition des zones de flores et faunes
* Les sous régions / Les zones d’évènements : Elles sont un niveau en dessous des régions. Elles définissent des zones de dans lequel un évènement particulier à lieu. Ces évènements sont liés au biome auquel il appartient, par exemple pour une jungle cela pourrait être :
  + Eléments mobile : Meutes d’animaux
  + Eléments fixe : temple maya
  + Eléments évolutif : plante qui pousse



L’exemple ci-dessus montre une répartition possible d’un biome.

### Implémentation et optimisation

#### Implémentation

Dans un premier temps, nous allons implémenter un système de génération procédurale dynamique qui se créera en même temps que l’utilisateur se déplacera dans l’environnement. L’idée est d’utiliser un diagramme de Voronoï pour découper les zones dans une région donnée. Le diagramme de Voronoï est un algorithme de subdivision de plan permettant de définir des zones d’influences par rapport à des germes. Nous tirons deux avantages de cette technique :

* L’ajout Delaunay : qui nous permettra d’ajouter des germes pour étendre notre environnement en fonction des déplacements du joueur, et donc de rendre l’environnement dynamique.
* Le déplacement des germes : pour les sous régions qui permettent aux éléments mobiles de voir le jour.

L’avantage du diagramme de Voronoï est qu’il permet un déplacement dynamique des zones pour les régions mobiles, ce qui nous permet de « forcer » l’utilisateur à vivre certaines expériences en fonction de ce que nous analyserons. Etant divisé en zone nous pouvons aussi gérer facilement le Culling (technique de rendu permettant de ne pas afficher les parties qui ne seront pas visible par le joueur). Aidé de cette technique nous définirons une zone autour du joueur qui sera prédéterminé, et nous pourrons ainsi charger/décharger des zones en fonction de la musique courante. Cette technique a été imaginée pour limiter le possible décalage qu’il y aurait entre la musique courante et l’environnement dans le cas ou l’utilisateur ne bougerait pas trop. Ainsi, même un utilisateur statique verra des évènements arriver vers lui et le monde autour de lui évoluer.

#### Optimisation

## Calibrage / Adaptation au joueur

# Analyse Marketing

## Cible

# Lexique

Pulsation : Le terme de pulsation désigne, l'accent intervenant de manière cyclique au début de chaque temps. Ainsi la régularité de la pulsation garantit donc l'égalité des temps, et par conséquent, un certain tempo.

Temps : le temps est l'unité de mesure de la durée musicale. Il n'existe pas de temps étalon, c'est le tempo qui va fixer la durée exacte des temps. Pour remplir sa fonction, le temps doit pouvoir être rigoureusement délimité par la pulsation.

Tempo : Fixe, pour un passage musical donné, la durée exacte des temps. C'est ainsi qu'un tempo rapide détermine des temps courts tandis qu'un tempo lent détermine des temps longs. C'est l'un des quatre principaux éléments de la partition écrite.

Mouvement : Le mouvement est le qualificatif d'un morceau de musique quelconque qui peut définir plusieurs méthodes de structuration. On peut ainsi renvoyer par exemple à la structure compositionnel (tel que le canon), au nombre de parties (duo, trio, ...), au caractère (lamento, cantabile, ...), et bien d'autre encore. Mais dans une mesure plus commune le mouvement qualifie le tempo (vif, lent, ...).

Rythme (en musique) : Le rythme est classiquement quantifié par une horloge interne (dites pulsation), qui détermine des temps, ce qui permet de comptabiliser les différentes figures de notes et de silences. C'est l'un des quatre principaux éléments de la partition écrite.

Timbre : Le timbre désigne l'ensemble des caractéristiques sonores qui permettent d'identifier un instrument.

Nuance : Une nuance est l'intensité relative d'une note, d'une phrase, ou encore d'un passage entier d'une œuvre musicale. On distingue ainsi les nuances générales (qui détermine l'intensité fixe à une ou plusieurs pulsations) des nuances pour un ensemble de notes (qui détermine l'augmentation ou la diminution de l'intensité en fonction du temps). C'est l'un des quatre principaux éléments de la partition écrite.

Harmonique (physique acoustique) :

Harmonique (musique) :

Harmonie : L'harmonie est l'émission simultanée de plusieurs sons différents. On parle alors de l'aspect « horizontale » de la musique par opposition à la dimension « vertical » qui correspond à l'enchaînements des sons.

Mélodie : On peut voir la mélodie comme la partie instrumentale ou vocale de premier plan par rapport aux accompagnements. Une mélodie est composée de plusieurs motifs.

C'est l'un des quatre principaux éléments de la partition écrite.

Motif (musique) : Un motif est une phrase musicale ou un fragment complet se répétant de façon régulière et continue au sein d'une œuvre.

Phrase (musique) : Une phrase est une partie d'une ligne mélodique ou d'une idée musicale naturellement délimitée, significative du point de vue de la déclamation, de l'articulation et de la respiration.

Gameplay : Le gameplay (Anglicisme pour : la jouabilité) caractérise des éléments d'une « expérience vidéoludique », c'est-à-dire le ressenti du joueur lorsqu'il utilise un jeu vidéo.

Level Design :

Level Designer :

Scénographe :

Scénographie :

Source utilisé :

Les objets au théâtre - Hélène Catsiapis - Communication et langages - 1979 - Volume 43 Numéro 1 pp. 59-78

1. Un rappel ou définition de l'ensemble de la terminologie lié au solfège (et d'une manière plus globale des mots en italique) se trouve dans le Glossaire. [↑](#footnote-ref-1)