

19 RÉPUBLIQUE FRANÇAISE  
INSTITUT NATIONAL  
DE LA PROPRIÉTÉ INDUSTRIELLE  
COURBEVOIE

11 N° de publication :  
(à n'utiliser que pour les  
commandes de reproduction)

3 028 086

21 N° d'enregistrement national : 14 60622

51 Int Cl<sup>8</sup> : G 10 H 1/00 (2016.01)

12

## DEMANDE DE BREVET D'INVENTION

A1

22 Date de dépôt : 04.11.14.

30 Priorité :

43 Date de mise à la disposition du public de la demande : 06.05.16 Bulletin 16/18.

56 Liste des documents cités dans le rapport de recherche préliminaire : *Se reporter à la fin du présent fascicule*

60 Références à d'autres documents nationaux apparentés :

Demande(s) d'extension :

71 Demandeur(s) : UNIVERSITE DE BORDEAUX Etablissement public — FR, CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) Etablissement public — FR et INSTITUT POLYTECHNIQUE DE BORDEAUX — FR.

72 Inventeur(s) : HANNA PIERRE, FERRARO PASCAL, ROBINE MATTHIAS et ALLALI JULIEN.

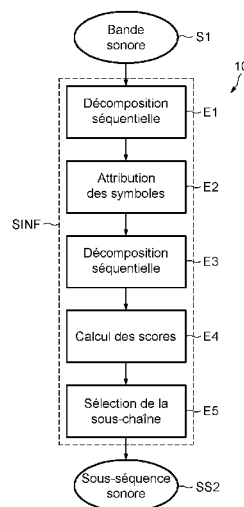
73 Titulaire(s) : UNIVERSITE DE BORDEAUX Etablissement public, CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) Etablissement public, INSTITUT POLYTECHNIQUE DE BORDEAUX.

74 Mandataire(s) : CASALONGA & ASSOCIES.

54 PROCÉDE DE RECHERCHE AUTOMATISE D'AU MOINS UNE SOUS-SEQUENCE SONORE REPRESENTATIVE AU SEIN D'UNE BANDE SONORE.

57 Procédé de recherche automatisée d'au moins une sous-séquence sonore (SS2) représentative au sein d'une bande sonore (S1), comprenant :

- une décomposition séquentielle (E1),
- une attribution (E2) d'un symbole choisi dans un alphabet,
- une décomposition séquentielle (E3),
- un calcul (E4) d'un score correspondant à un cumul de taux d'identité de séquence de la sous-chaîne par rapport aux autres sous-chaînes
- une détermination (E5) de la sous-chaîne la plus représentative.



FR 3 028 086 - A1



**Procédé de recherche automatisée d'au moins une sous-séquence  
sonore représentative au sein d'une bande sonore**

L'invention concerne de façon générale les méthodes d'analyse  
5 et de traitement des bandes sonores.

Des bandes sonores peuvent correspondre à des productions  
musicales, par exemple des morceaux de musique, et elles peuvent être  
commercialisées sous la forme de CD audio. Il est également possible  
de commercialiser les bandes sonores par des moyens numériques de  
10 vente en ligne.

Généralement, aux fins de promouvoir la vente d'œuvres  
musicales, des extraits des titres sont mis à la disposition d'acheteurs  
potentiels. Ces extraits, de durées relativement courtes, par exemple  
allant de 30 à 45 secondes, sont destinés à offrir un aperçu des œuvres  
15 proposées à la vente.

Le choix de ces extraits peut être réalisé de manière plus ou  
moins arbitraire, par exemple en optant pour un échantillon provenant  
des premiers instants des titres. Les extraits ainsi choisis, c'est-à-dire  
de manière arbitraire, sont rarement représentatifs des œuvres dont ils  
20 sont issus, et ils peuvent donner un aperçu erroné aux clients. Certains  
acheteurs potentiels peuvent ainsi être dissuadés de procéder à l'achat.  
D'autres acheteurs potentiels pourraient être déçus d'avoir acquis des  
œuvres qui ne répondraient nullement à leurs attentes. Ces acheteurs  
frustrés risquent alors de se détourner définitivement d'un tel système  
25 de vente de musique.

Alternativement, il a été proposé de sélectionner des extraits  
qui correspondent à des refrains et/ou des couplets des titres. Ces  
passages, du fait de leurs répétitions au sein d'un même morceau, sont  
censés être musicalement les plus attractifs voire les plus  
30 représentatifs du morceau considéré. Des moyens automatisés ont ainsi  
été développés en vue d'identifier ces sous-séquences sonores  
répétées, qui peuvent alors être utilisées comme résumé sonore.

À cet égard, le document FR 2 856 817 décrit le traitement  
automatisé d'une bande sonore dans lequel un traitement de

transformée spectrale permet d'identifier une sous-séquence répétée, et de localiser le début et la fin de cette sous-séquence répétée. Les sous-séquences répétées coïncident généralement soit avec les refrains soit avec les couplets des morceaux ou des titres analysés, lorsque ces derniers comprennent effectivement un refrain et des couplets.

La solution décrite dans le document FR 2 856 817 a pour inconvénient de fournir des sous-séquences ayant des durées non-standardisées, qui peuvent grandement varier d'un titre à l'autre. La durée des sous-séquences est totalement indépendante de la volonté de l'opérateur, elle est en effet intrinsèque au morceau/titre de référence. Un autre inconvénient à cette solution est qu'elle ne peut être généralisée à tout type de bande sonore, car toutes les bandes sonores ne sont pas constituées d'un refrain et de couplets.

La présente invention a donc pour but de remédier aux inconvénients présentés ci-avant, et en particulier de permettre la détermination automatique d'une sous-séquence représentative d'une bande sonore au sein de cette bande sonore de manière simple, avec une durée qui peut être choisie.

L'invention a donc pour objet un procédé de recherche automatisée d'au moins une sous-séquence sonore au sein d'au moins une bande sonore, la sous-séquence sonore recherchée étant représentative de ladite bande sonore, comprenant une phase élémentaire de recherche automatisée comportant:

- une décomposition séquentielle de la bande sonore en une succession ordonnée de séquences élémentaires éventuellement partiellement chevauchantes, de sorte que :

$$N = \frac{(D-d)}{d} \cdot \frac{1}{(1-\alpha)} + 1$$

avec :

D, la durée de la bande sonore,  
d, la durée de chacune des séquences élémentaires,

$\alpha$ , le taux de chevauchement de chaque séquence élémentaire avec la séquence élémentaire qui la précède,  $\alpha$  étant supérieur ou égal à 0 et inférieur à 1, et

5 N, le nombre de séquences élémentaires formant ladite bande sonore,

- une attribution à chaque séquence élémentaire d'un symbole choisi dans un alphabet en fonction d'au moins un paramètre inhérent de la séquence élémentaire de façon à obtenir une chaîne de symboles représentative de la bande sonore,

10 - une décomposition séquentielle de ladite chaîne de symboles en une suite régulière de sous-chaînes consécutives ayant une durée  $d_{SC}$  correspondant à n séquences élémentaires,  $d_{SC}$  étant supérieure à d,

15 ladite décomposition étant mise en œuvre de façon à ce que le début de la première sous-chaîne de ladite suite coïncide avec une séquence élémentaire particulière de ladite chaîne de symboles,

- pour chaque sous-chaîne de ladite suite régulière de sous-chaînes, un calcul d'un score correspondant à un cumul de taux d'identité de séquence de la sous-chaîne par rapport aux autres sous-chaînes,

20 - une sélection de la sous-chaîne ayant le score le plus élevé, la sous-chaîne ayant le score le plus élevé étant la sous-séquence sonore représentative recherchée.

25 Par décomposition séquentielle, on entend notamment une décomposition ou un découpage ordonné(e), c'est-à-dire d'éléments successifs.

30 La décomposition séquentielle en une suite de sous-chaînes est mise en œuvre de sorte que les séquences élémentaires de début de deux sous-chaînes consécutives sont séparées par n-1 séquences élémentaires. n correspond au saut de séquences élémentaires entre deux sous-chaînes consécutives.

Ainsi, contrairement au procédé de traitement décrit dans document FR 2 856 817, qui est basé sur une identification de passages répétés, le procédé selon l'invention détermine une sous-séquence

ayant une durée  $d_{sc}$  choisie préalablement, et qui a une meilleure homogénéité musicale avec l'ensemble de la bande sonore.

5 Les sous-séquences obtenues au moyen du procédé décrit ci-avant peuvent être utilisées pour faire la promotion d'œuvres musicales.

10 On peut noter que le procédé peut être appliqué, soit directement soit moyennant quelques adaptations évidentes, à tout type de bande sonore, par exemple des fichiers informatiques audio, ou encore des représentations symboliques de suites de notes ou des fichiers de tablature de guitare, par exemple des fichiers MIDI.

A titre indicatif, lors du traitement d'une série de notes ou d'accords formant une bande sonore, on peut obtenir une sous-séquence représentative sous la forme de descripteurs tonaux (HPCP).

15 On peut aussi noter que le procédé décrit ci-avant est adapté quelle que soit la longueur de la sous-séquence recherchée, c'est-à-dire la durée de l'extrait (ou l'échantillon). Typiquement, le procédé selon l'invention est adapté pour une longueur de 45 secondes, mais il peut également être utilisé pour n'importe quelle autre longueur de sous-séquence, qu'elle soit plus courte ou plus longue que 45  
20 secondes. Ainsi, le procédé selon l'invention permet par exemple de choisir une deuxième longueur (différente de 45 secondes), qui peut être déterminée en fonction de l'application visée. Aujourd'hui, il est classique de proposer uniquement des échantillons de la même longueur, par exemple uniquement des échantillons de 45 secondes.

25 Par ailleurs, contrairement au procédé du document mentionné ci-dessus, il est possible ici d'obtenir une sous-séquence d'un morceau de musique qui comporte à la fois le refrain ainsi que quelques secondes précédant le refrain et quelques secondes suivant ce refrain (si le refrain a une longueur inférieure à la deuxième longueur). On  
30 peut obtenir ainsi une sous-séquence qui peut être plus agréable à écouter qu'une sous-séquence qui démarre directement avec le refrain.

On peut noter que pour obtenir la sous-séquence sonore la plus représentative, on peut déterminer à quel instant de la bande sonore correspond le début de la sous-chaîne ayant le score le plus élevé, pour

ensuite identifier dans la bande sonore la sous-séquence représentative.

Ledit au moins un paramètre inhérent d'échantillon est choisi dans le groupe formé par la tonalité, le rythme, le timbre, l'accord, les  
5 paroles et le contexte tonal.

Le calcul du score correspondant à un cumul de taux d'identité de séquence peut être effectué au moyen d'un algorithme d'alignement, par exemple l'algorithme d'alignement local Smith-Waterman.

10 Alternativement, le calcul du score correspondant à un cumul de taux d'identité de séquence peut être effectué au moyen d'un algorithme de chaînage tel que celui décrit dans l'ouvrage de Dan Gusfield intitulé « Algorithms on Strings, Trees and Sequences ». Cambridge University Press, 1997,59,60,78,81,82,83.

15 Il est possible d'utiliser des algorithmes connus de l'homme du métier qui permettent d'obtenir ces scores. À titre indicatif, on peut mettre en œuvre l'algorithme BLAST (décrit dans l'article de Altschul SF, Gish W, Miller W, Myers EW et Lipman DJ intitulé « Basic local alignment search tool » J Mol Biol. 1990) ou l'algorithme FASTA  
20 (décrit dans l'article « Rapid and sensitive protein similarity searches » de D J Lipman et W R Pearson, Science 04/1985), utilisés dans le domaine de la biologie.

On peut notamment utiliser pour la détermination de la sous-chaîne la plus répétée un algorithme de détermination de  
25 correspondance de chaîne. On obtient ainsi de manière automatique un score de répétition permettant de déduire la sous-chaîne la plus répétée. La séquence élémentaire particulière de ladite chaîne de symboles peut ne pas être la première séquence élémentaire de ladite chaîne de symboles. Ainsi, on ne prend pas en compte le début de la  
30 bande sonore.

La durée  $d$  peut être comprise entre 50 et 1000 millisecondes.

Selon un mode de mise en œuvre,  $\alpha$  peut être égal à 0 et les séquences élémentaires ne sont alors pas chevauchantes. En variante,  $\alpha$  est compris entre 0,1 et 0,9 et les séquences sont chevauchantes.  $\alpha$  est

typiquement de l'ordre de  $0,5.n$  peut être compris entre 1 et 100. De préférence,  $n$  est compris entre 20 et 50.

5 Le procédé selon l'invention, de mise en œuvre et d'utilisation particulièrement simples et rapides, permet avantageusement de générer un extrait sonore particulier, d'une durée ajustable (c'est-à-dire qui peut être librement fixée par un opérateur), éventuellement standardisée, musicalement représentatif d'un morceau et/ou d'un titre musical spécifique.

10 Selon une autre application particulière, le procédé de recherche automatisé défini ci-avant peut être avantageusement utilisé en vue de générer un « résumé » d'un ensemble défini de bandes sonores (notamment, des titres d'un même album, des titre d'une compilation d'albums, des titres d'une « playlist », l'œuvre complète ou partielle d'un artiste/groupe...). Une génération d'un tel résumé  
15 comprend alors une compilation d'extraits, chacun obtenu au moyen de la phase élémentaire de recherche automatisée du procédé de recherche automatisée défini ci-avant.

Pour cela, on peut rechercher automatiquement grâce à  $N'$  phases élémentaires de recherche automatisée,  $N'$  sous-séquences sonores répétitives dans respectivement  $N'$  séquences ou bandes sonores. Les  $N'$  sous-séquences sonores répétitives peuvent former un résumé de l'ensemble des  $N'$  séquences sonores. Cette génération de résumé peut comprendre en outre une concaténation desdites  $N'$  sous-séquences sonores répétitives sous la forme d'une seule sous-séquence.  
25 Ainsi, dans le cas du traitement d'un album de musique comprenant  $N'$  séquences ou bandes sonores, on peut obtenir une sous-séquence qui est un résumé de l'album entier. On peut ainsi obtenir une sous-séquence pouvant comporter plusieurs refrains, représentative de l'album entier.

30 L'invention a également pour objet un système informatique comprenant des moyens configurés pour mettre en œuvre le procédé tel que défini ci-avant, par exemple un ordinateur, comportant des moyens tels qu'une unité centrale et des moyens de mémoire, configurés pour mettre en œuvre le procédé défini ci-avant.

L'invention a également pour objet un produit programme d'ordinateur chargeable directement dans une mémoire d'un système informatique, comprenant des portions de code de logiciel pour l'exécution du procédé tel que défini ci-avant lorsque ledit programme est exécuté sur ledit système informatique.

Enfin, l'invention a pour objet un support lisible par un système informatique, ayant des instructions exécutables par ordinateur adaptées pour provoquer l'exécution par le système informatique du procédé tel que défini ci-avant.

D'autres avantages et caractéristiques de l'invention apparaîtront à l'examen de la description détaillée de modes de mise en œuvre et de réalisation, nullement limitatifs, et des dessins annexés sur lesquels les figures 1 et 2 représentent de manière schématique les étapes de différents modes de mise en œuvre d'un procédé de recherche automatisée selon l'invention.

La présente invention peut être mise en œuvre pour traiter des bandes sonores référencées S1 sur la figure 1.

Une telle bande sonore peut être un morceau de musique. Il convient de noter qu'une telle bande sonore peut être obtenue après un échantillonnage d'un signal audio, par exemple à 44,1 kHz comme tel est le cas pour les CD audio.

La bande sonore S1 a plusieurs portions bien définies dans le domaine de la musique, notamment un refrain et des couplets.

Sur la figure 1, on a représenté de manière schématique différentes étapes d'une phase élémentaire 10 d'un mode de mise en œuvre d'un procédé selon l'invention.

Par ailleurs, sur la figure 1, la référence SINP désigne globalement un système informatique, par exemple un ordinateur, comportant des moyens tels qu'une unité centrale et des moyens de mémoire, configurés pour mettre en œuvre un mode de mise en œuvre du procédé selon l'invention.

Un tel procédé selon l'invention permet de rechercher automatiquement dans une bande sonore S1 une sous-séquence représentative ayant une longueur choisie.



La séquence sonore S1 peut être un signal échantillonné stocké sur un support informatique. Pour rendre possible la recherche de la sous-séquence, une décomposition séquentielle est mise en œuvre dans une première étape E1.

5 Cette décomposition séquentielle E1 de la bande sonore comporte une décomposition séquentielle en une succession ordonnée de séquences élémentaires éventuellement partiellement chevauchantes, de sorte que la formule suivante soit vérifiée :

$$N = \frac{(D-d)}{d} \cdot \frac{1}{(1-\alpha)} + 1$$

10 avec :

D, la durée de la bande sonore (S1),

d, la durée de chacune des séquences élémentaires,

$\alpha$ , le taux de chevauchement de chaque séquence élémentaire avec la séquence élémentaire qui la précède,  $\alpha$  étant supérieur

15 ou égal à 0 et inférieur à 1, et

N, le nombre de séquences élémentaires formant ladite bande sonore.

Dans cette étape, on découpe la bande sonore S1 en une pluralité de séquences élémentaires ayant toutes une longueur choisie  
20 d, par exemple de l'ordre de 300 millisecondes. On obtient ainsi un nombre limité de séquences élémentaires à traiter.

Préalablement à la mise en œuvre du procédé, un alphabet a été défini. N'importe quel alphabet peut être utilisé pour la mise en œuvre du procédé. Cet alphabet comporte des symboles destinés à être  
25 attribués à des séquences élémentaires en fonction d'au moins un paramètre inhérent de ces séquences élémentaires. A titre indicatif, le paramètre peut être choisi dans le groupe formé par la tonalité, le rythme, le timbre, l'accord, les paroles et le contexte tonal.

On peut noter que pour la tonalité, il est possible de considérer  
30 des descripteurs de type « Pitch Class Profile » (PCP ou chroma), bien connus de l'homme du métier et qui pourra à toutes fins utiles se référer notamment à l'article de T. Fujishima, intitulé « Realtime

chord recognition of musical sound: a system using common lisp music », Proc. of ICMC, pp. 464–467 (1999).

Si le paramètre inhérent est le rythme, il est possible de considérer des descripteurs de type Meter Class Profiles (MCP) tels que ceux décrits dans l'article de M. Robine, M. Lagrange, P. Hanna, intitulé « Meter Class Profiles For Music Similarity And Retrieval », Proc. of the International Society for Music Information Retrieval Conference (ISMIR), pp. 639-644, Kobe, Japan, October 2009.

Si le paramètre inhérent est le timbre, on peut considérer des descripteurs de type Mel Frequency Cepstral Coefficients (MFCC), bien connus de l'homme de l'art.

Si le paramètre inhérent est le contexte tonal, on peut attribuer un symbole en fonction du contexte tonal, chaque symbole pouvant être associé notamment à une amplitude de spectre et à une octave.

Ainsi, lors d'une deuxième étape E2, chaque séquence élémentaire obtenue en sortie de l'étape E1 est traitée pour déterminer la valeur du ou des paramètres inhérents associés à cette séquence élémentaire pour ensuite attribuer un symbole à cette séquence élémentaire. On obtient ainsi une chaîne de symboles, qui correspond à une représentation simplifiée de la bande sonore S1 et qui est donc traitable de façon automatique en un temps raisonnable.

L'homme du métier sait déterminer les valeurs de ces paramètres pour une séquence élémentaire. A titre d'exemple, on peut mettre en œuvre une étape de mesure du paramètre de la séquence élémentaire, et lire ensuite dans une cartographie ayant en entrée des valeurs de paramètre inhérent et délivrant en sortie le symbole à associer.

Une étape E3 est ensuite mise en œuvre, dans laquelle on met en œuvre une décomposition séquentielle de la chaîne de symboles obtenue à l'étape E2.

La décomposition séquentielle E3 de ladite chaîne de symboles comporte une décomposition séquentielle en une suite régulière de sous-chaînes consécutives ayant une durée  $d_{SC}$  correspondant à  $n$  séquences élémentaires,  $d_{SC}$  étant supérieure à  $d$ . La décomposition

est mise en œuvre de façon à ce que le début de la première sous-chaîne de ladite suite coïncide avec une séquence élémentaire particulière de ladite chaîne de symboles.

5 A titre indicatif,  $d_{SC}$  peut être 45 secondes. Le choix de la deuxième longueur dépend de l'utilisateur qui met en œuvre le procédé. La séquence élémentaire particulière peut être la première séquence élémentaire ou une autre séquence élémentaire.

10 Au cours d'une étape E4, les sous-chaînes obtenues par l'étape E3 sont traitées pour qu'un calcul (E4) soit mis en œuvre. Dans l'étape E4, on calcule un score correspondant à un cumul de taux d'identité de séquence de la sous-chaîne par rapport aux autres sous-chaînes.

15 La détermination de ces scores peut être mise en œuvre au moyen d'algorithmes de détermination de correspondance de chaîne ou d'algorithmes d'alignement local. Comme on le conçoit, l'utilisation d'un alphabet restreint permet de limiter la complexité de la mise en œuvre du procédé. À titre indicatif, on peut mettre en œuvre les algorithmes BLAST ou FASTA, utilisés dans le domaine de la biologie, on encore l'algorithme Smith-Waterman. Il convient de noter qu'il est possible, en utilisant de tels algorithmes, de déterminer la sous-séquence répétitive avec une complexité notée  $O(|x|^2 \log(|x|))$ , où  $O$  correspond à la notation de Landau et  $x$  correspond à la longueur de la chaîne de symboles obtenue en sortie de l'étape E2.

25 Enfin, dans l'étape E5 on sélectionne la sous-chaîne ayant le score le plus élevé, la sous-chaîne ayant le score le plus élevé étant la sous-séquence sonore représentative recherchée SS2.

Si S1 est un morceau de musique, alors SS2 peut contenir le refrain.

30 On obtient ainsi une sous-séquence représentative dans une bande sonore du fait de sa répétition, et un extrait ayant une longueur choisie.

Selon une autre application particulière, la phase élémentaire 10 de recherche automatisé décrite ci-avant peut être avantageusement utilisée en vue de générer un « résumé » d'un ensemble défini de N' bandes sonores (notamment, des titres d'un même album, des titres

d'une compilation d'albums, des titres d'une « playlist », l'œuvre complète ou partielle d'un artiste/groupe...). Une génération d'un tel résumé comprend alors une compilation d'extraits, chacun obtenu au moyen de ladite phase élémentaire de recherche automatisée 10.

5            Pour cela, comme illustré sur la figure 2, on peut rechercher automatiquement grâce à N' phases élémentaires de recherche automatisée 10, N' sous-séquences sonores répétitives  $SS_{2_1}$ - $SS_{2_N}$  dans respectivement N' séquences ou bandes sonores  $S_{1_1}$ - $S_{1_N}$ .

10           Les N' sous-séquences sonores répétitives  $SS_{2_1}$ - $SS_{2_N}$  peuvent former un résumé de l'ensemble des N' séquences sonores  $S_{1_1}$ - $S_{1_N}$ .

Cette génération de résumé peut comprendre en outre une concaténation 20 desdites N' sous-séquences sonores répétitives  $SS_{2_1}$ - $SS_{2_N}$  sous la forme d'une seule sous-séquence SSR.

15           Ainsi, dans le cas du traitement d'un album de musique comprenant N' séquences ou bandes sonores  $S_{1_1}$ - $S_{1_N}$ , on peut obtenir une sous-séquence SSR qui est un résumé de l'album entier.

On peut ainsi obtenir une sous-séquence SSR pouvant comporter plusieurs refrains, représentative de l'album entier.

## REVENDEICATIONS

1. Procédé de recherche automatisée d'au moins une sous-séquence sonore (SS2) au sein d'au moins une bande sonore (S1), la sous-séquence sonore recherchée étant représentative de ladite au moins une bande sonore, comprenant une phase élémentaire de recherche automatisée comportant :

- 5
- une décomposition séquentielle (E1) de la bande sonore en une succession ordonnée de séquences élémentaires éventuellement partiellement chevauchantes, de sorte que :
- 10

$$N = \frac{(D-d)}{d} \cdot \frac{1}{(1-\alpha)} + 1$$

avec :

- D, la durée de la bande sonore (S1),
- 15 d, la durée de chacune des séquences élémentaires,
- $\alpha$ , le taux de chevauchement de chaque séquence élémentaire avec la séquence élémentaire qui la précède,  $\alpha$  étant supérieur ou égal à 0 et inférieur à 1, et

N, le nombre de séquences élémentaires formant ladite bande sonore,

20

- une attribution (E2) à chaque séquence élémentaire d'un symbole choisi dans un alphabet en fonction d'au moins un paramètre inhérent de la séquence élémentaire, de façon à obtenir une chaîne de symboles,
- 25

- une décomposition séquentielle (E3) de ladite chaîne de symboles en une suite régulière de sous-chaînes consécutives ayant une durée  $d_{SC}$  correspondant à n séquences élémentaires,  $d_{SC}$  étant supérieure à d,

30 ladite décomposition étant mise en œuvre de façon à ce que le début de la première sous-chaîne de ladite suite coïncide avec

une séquence élémentaire particulière de ladite chaîne de symboles,

- pour chaque sous-chaîne de ladite suite régulière de sous-chaînes, un calcul (E4) d'un score correspondant à un cumul de taux d'identité de séquence de la sous-chaîne par rapport aux autres sous-chaînes,
- une sélection (E5) de la sous-chaîne ayant le score le plus élevé, la sous-chaîne ayant le score le plus élevé étant la sous-séquence sonore représentative recherchée.

2. Procédé selon la revendication 1, dans lequel ledit au moins un paramètre inhérent d'échantillon est choisi dans le groupe formé par la tonalité, le rythme, le timbre, l'accord, les paroles et le contexte tonal.

3. Procédé selon la revendication 2, dans lequel le paramètre inhérent est le contexte tonal, et dans lequel on attribue un symbole en fonction du contexte tonal.

4. Procédé selon la revendication 3, dans lequel chaque symbole est associé à une amplitude de spectre et à une octave.

5. Procédé selon l'une quelconque des revendications précédentes, dans lequel le calcul du score correspondant à un cumul de taux d'identité de séquence est effectué au moyen d'un algorithme d'alignement.

6. Procédé selon l'une quelconque des revendications 1 à 4, dans lequel le calcul du score correspondant à un cumul de taux d'identité de séquence est effectué au moyen d'un algorithme de chaînage.

7. Procédé selon l'une quelconque des revendications précédentes, dans lequel la séquence élémentaire particulière de ladite chaîne de symboles n'est pas la première séquence élémentaire de ladite chaîne de symboles.

8. Procédé selon l'une quelconque des revendications précédentes, dans lequel la durée  $d$  est comprise entre 50 et 1000 millisecondes.

9. Procédé selon l'une quelconque des revendications précédentes, dans lequel  $\alpha$  est égal à 0 et dans lequel les séquences élémentaires ne sont pas chevauchantes.

5 10. Procédé selon l'une quelconque des revendications précédentes, dans lequel  $\alpha$  est compris entre 0,1 et 0,9 et dans lequel les séquences sont chevauchantes.

11. Procédé selon l'une quelconque des revendications précédentes, dans lequel  $n$  est compris entre 1 et 100.

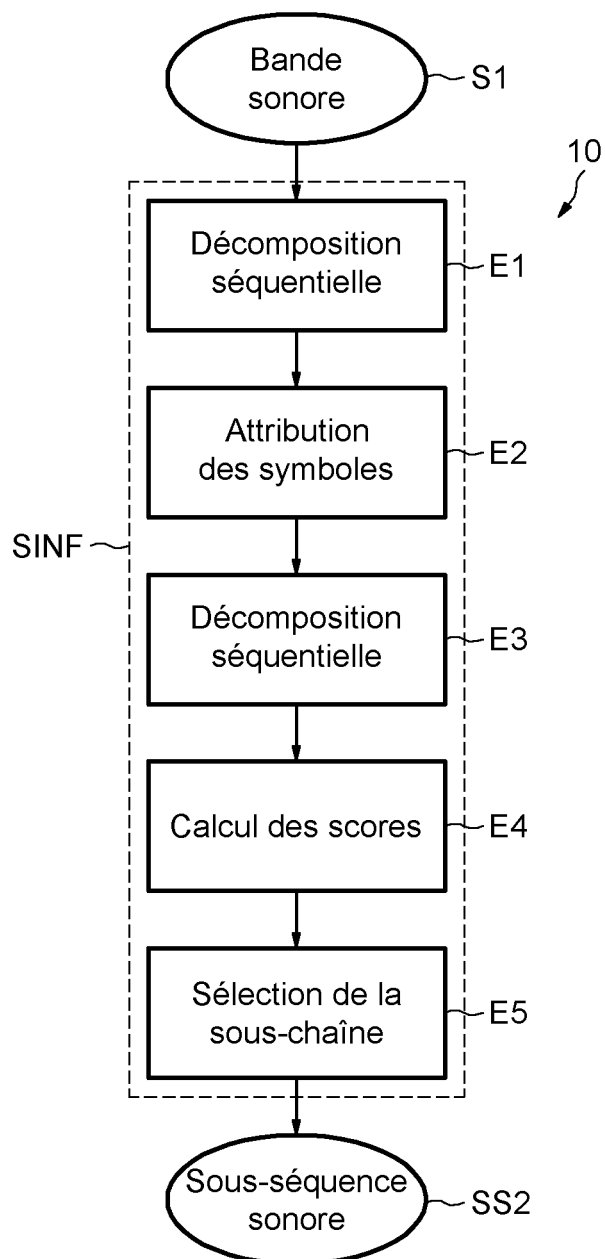
10 12. Procédé selon l'une quelconque des revendications précédentes, comprenant  $N'$  phases élémentaires de recherche automatisée délivrant respectivement  $N'$  sous-séquences sonores respectivement représentatives de  $N'$  bandes sonores de façon à générer un résumé de l'ensemble des  $N'$  bandes sonores.

15 13. Procédé selon la revendication 12, comprenant en outre une concaténation des  $N'$  sous-séquences en une seule sous-séquence.

14. Système informatique comprenant des moyens configurés pour mettre en œuvre le procédé selon l'une des revendications 1 à 13.

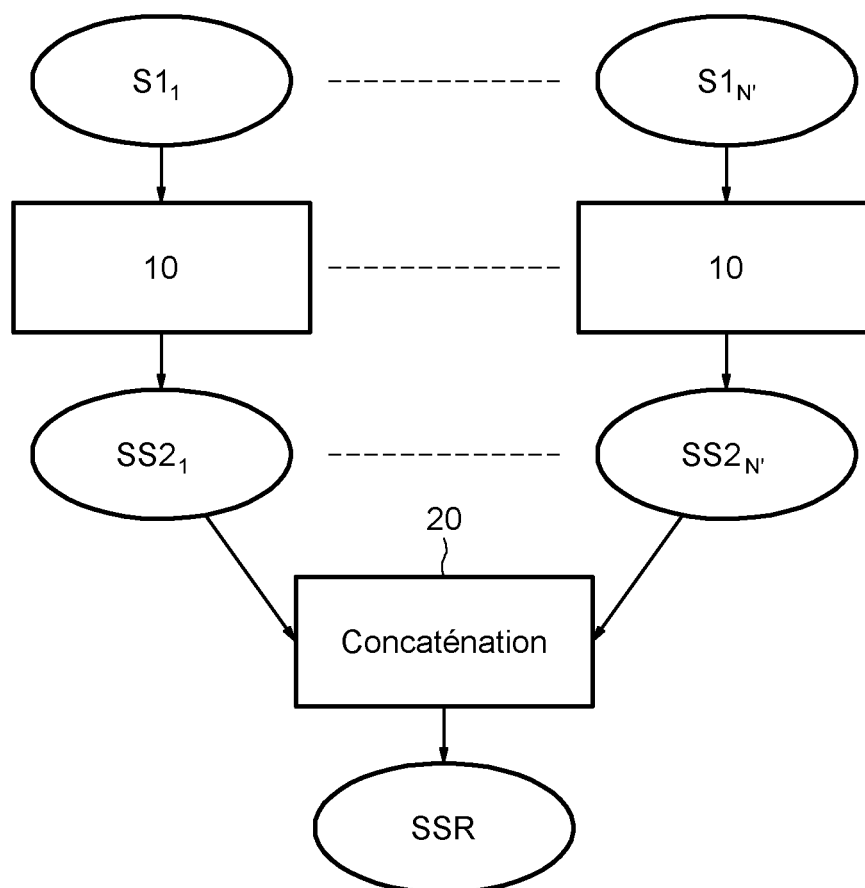
20 15. Produit programme d'ordinateur chargeable directement dans une mémoire d'un système informatique, comprenant des portions de code de logiciel pour l'exécution du procédé selon l'une des revendications 1 à 13 lorsque ledit programme est exécuté sur ledit système informatique.

25 16. Support lisible par un système informatique, ayant des instructions exécutables par ordinateur adaptées pour provoquer l'exécution par le système informatique du procédé selon l'une des revendications 1 à 13.

1/2  
**FIG. 1**



2/2  
FIG.2





**RAPPORT DE RECHERCHE  
PRÉLIMINAIRE**

N° d'enregistrement national

établi sur la base des dernières revendications déposées avant le commencement de la recherche

FA 803286  
FR 1460622

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
X Y	US 2012/167748 A1 (BAO SHENGHUA [CN] ET AL) 5 juillet 2012 (2012-07-05) * By applying classifying rules to music file, a feature segment can be identified; alinéa [0031]; figure 3 * * character string of the entire song; comparison of segments for identity of features; alinéa [0033] * * segment B is a feature segment of the song "Beijing Welcomes You"; alinéa [0038] * * classifying rules by Bayes- , NVM- ... -models; alinéa [0039] * * "Beijing Welcomes You" = AABAABBBBB, thus character string segment B corresponds to 6 music segments; alinéa [0051] *	1-11, 14-16 12,13	G10H1/00
T	ARONOWITZ H: "Segmental Modeling for Audio Segmentation", 2007 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING 15-20 APRIL 2007 HONOLULU, HI, USA, IEEE, PISCATAWAY, NJ, USA, 15 avril 2007 (2007-04-15), pages IV-393, XP031463869, ISBN: 978-1-4244-0727-9 * defining overlapping segments; alinéas [003.] - [speechsegmentation] *		DOMAINES TECHNIQUES RECHERCHÉS (IPC) G10H G06F
A,D	FR 2 856 817 A1 (FRANCE TELECOM [FR]) 31 décembre 2004 (2004-12-31) * abrégé *	1-16	
Date d'achèvement de la recherche		Examineur	
29 juin 2015		Glasser, Jean-Marc	
CATÉGORIE DES DOCUMENTS CITÉS		T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant	
X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire			

4  
EPO FORM 1503 12.99 (P04C14)



**RAPPORT DE RECHERCHE  
PRÉLIMINAIRE**

établi sur la base des dernières revendications  
déposées avant le commencement de la recherche

N° d'enregistrement  
national

FA 803286  
FR 1460622

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
Y	US 6 225 546 B1 (KRAFT REINER [US] ET AL) 1 mai 2001 (2001-05-01) * Detect main melody and generate audio summary; revendication 1; figure 6 *	12,13	DOMAINES TECHNIQUES RECHERCHÉS (IPC)
Y	CN 103 440 313 A (UNIV FUDAN) 11 décembre 2013 (2013-12-11) * revendication 1; figure 4 * * Obtaining the one-dimensional characteristic sequence based on music fingerprint; revendication 2 *	12,13	
Y	US 2011/276864 A1 (OULES GUILLAUME [FR]) 10 novembre 2011 (2011-11-10) * media motif; alinéa [0024] * * media sequence; revendication 1 *	12,13	
Date d'achèvement de la recherche		Examineur	
29 juin 2015		Glasser, Jean-Marc	
CATÉGORIE DES DOCUMENTS CITÉS		T : théorie ou principe à la base de l'invention	
X : particulièrement pertinent à lui seul		E : document de brevet bénéficiant d'une date antérieure	
Y : particulièrement pertinent en combinaison avec un		à la date de dépôt et qui n'a été publié qu'à cette date	
autre document de la même catégorie		de dépôt ou qu'à une date postérieure.	
A : arrière-plan technologique		D : cité dans la demande	
O : divulgation non-écrite		L : cité pour d'autres raisons	
P : document intercalaire		& : membre de la même famille, document correspondant	

4

EPO FORM 1503 12.99 (P04C14)

**ANNEXE AU RAPPORT DE RECHERCHE PRÉLIMINAIRE  
RELATIF A LA DEMANDE DE BREVET FRANÇAIS NO. FR 1460622 FA 803286**

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche préliminaire visé ci-dessus.

Les dits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du **29-06-2015**

Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets, ni de l'Administration française

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 2012167748 A1	05-07-2012	CN 102541965 A US 2012167748 A1	04-07-2012 05-07-2012
FR 2856817 A1	31-12-2004	EP 1636789 A2 FR 2856817 A1 JP 2007520727 A US 2006288849 A1 WO 2005004002 A2	22-03-2006 31-12-2004 26-07-2007 28-12-2006 13-01-2005
US 6225546 B1	01-05-2001	AUCUN	
CN 103440313 A	11-12-2013	CN 103440313 A WO 2015027751 A1	11-12-2013 05-03-2015
US 2011276864 A1	10-11-2011	EP 2393022 A2 FR 2959037 A1 US 2011276864 A1	07-12-2011 21-10-2011 10-11-2011