

①9 RÉPUBLIQUE FRANÇAISE
—
**INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE**
—
COURBEVOIE
—

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

3 035 729

②1 N° d'enregistrement national : **15 53792**

⑤1 Int Cl⁸ : **G 06 T 7/00 (2015.01)**

⑫

BREVET D'INVENTION

B1

⑤4 **PROCEDE D'ANALYSE D'UNE SEQUENCE VIDEO ET EQUIPEMENT POUR LA MISE EN
OEUVRE DU PROCEDE.**

②2 **Date de dépôt** : 28.04.15.

③0 **Priorité** :

④3 **Date de mise à la disposition du public
de la demande** : 04.11.16 Bulletin 16/44.

④5 **Date de la mise à disposition du public du
brevet d'invention** : 27.11.20 Bulletin 20/48.

⑤6 **Liste des documents cités dans le rapport de
recherche** :

Se reporter à la fin du présent fascicule

⑥0 **Références à d'autres documents nationaux
apparentés** :

Demande(s) d'extension :

⑦1 **Demandeur(s)** : *ATEME Société anonyme* — FR.

⑦2 **Inventeur(s)** : LARBIER PIERRE.

⑦3 **Titulaire(s)** : *ATEME Société anonyme*.

⑦4 **Mandataire(s)** : PLASSERAUD IP.

FR 3 035 729 - B1



PROCEDE D'ANALYSE D'UNE SEQUENCE VIDEO ET EQUIPEMENT POUR LA MISE EN ŒUVRE DU PROCEDE

La présente invention se rapporte à un procédé d'analyse de séquences vidéo et un dispositif pour la mise en œuvre de ce procédé. Elle s'applique notamment à l'analyse d'une séquence vidéo en vue de traitements (codage vidéo, compression, débruitage, etc.) à effectuer sur la séquence.

Les données vidéo font en général l'objet d'un codage source visant à les compresser afin de limiter les ressources nécessaires à leur transmission et/ou à leur stockage. Il existe de nombreux standards de codage, tels que H.264/AVC, H.265/HEVC et MPEG-2, qui peuvent être utilisés à cette fin.

On considère une séquence vidéo comprenant un ensemble d'images. Nombre de traitements automatiques sur des séquences vidéo requièrent une analyse préalable. C'est le cas par exemple de la compression en deux passes à débit variable où une première passe correspondant à une phase d'analyse préalable permet principalement de déterminer la complexité de la séquence avant de l'encoder à proprement parler dans une seconde passe correspondant à une phase de traitement. Les opérations de la phase d'analyse sont bien souvent aussi complexes que celles de la phase de traitement qui suit. Par conséquent, le temps global de traitement avec analyse préalable se trouve être nettement supérieur à celui d'un traitement sans analyse préalable.

Dans le cas d'un traitement par encodage vidéo, la durée de la phase d'analyse préalable joue ainsi un rôle prépondérant pour la détermination de la durée totale nécessaire à l'encodage, autrement dit de la vitesse de traitement de la séquence traitée. Il est donc souhaitable de diminuer autant que possible la durée de la phase d'analyse préalable pour atteindre des vitesses de traitement élevées qui soient compatibles avec des impératifs de traitements de catalogue vidéo contenant par exemple plusieurs milliers de films.

Les techniques existantes d'analyse de séquence vidéo préalablement à un traitement consistent généralement à balayer successivement les images qui composent la séquence vidéo à traiter. A chaque image, des statistiques sont extraites, puis à l'issue de la phase d'analyse, un résumé contenant des informations synthétiques sur la séquence est calculé. La phase de traitement qui s'ensuit prend en entrée la séquence vidéo ainsi que le résumé calculé précédemment.

Dans le cas d'une compression vidéo en deux passes, le traitement

d'analyse effectué en première passe peut aussi consister en un encodage vidéo qui permet d'extraire pour chaque image ou chaque sous-séquence d'images une relation entre la qualité vidéo (tenant compte de la distorsion de compression) et le débit obtenu après compression. Lors de la seconde passe, l'ensemble de ces relations est utilisé par l'encodeur de manière à réguler le débit de façon optimale.

Dans le cas d'un débruiteur vidéo en deux passes, la phase d'analyse préalable permet généralement d'extraire les caractéristiques du bruit afin de guider le débruitage qui est réalisé lors de la seconde phase. Une méthode d'estimation des caractéristiques du bruit consiste à le retirer puis à mesurer des statistiques du signal dont on a extrait le bruit pour déterminer une différence entre le signal bruité et le signal après extraction du bruit caractérisant le bruit extrait. Cette opération présente généralement la complexité d'un débruitage complet, et comme dans le cas de la compression vidéo, double quasiment le temps de traitement par rapport à un traitement fait sans analyse préalable.

Le document US 2010/0027622 propose de réduire la résolution spatiale et/ou temporelle d'un flux vidéo préalablement à une première passe d'encodage dans le but de réduire le temps de calcul dans le cas particulier d'un encodage vidéo à deux passes.

La réduction de la résolution spatiale suppose de diminuer la taille des images analysées, de manière linéaire ou par sélection d'une portion de chaque image. Les statistiques extraites lors de la première passe sont ensuite extrapolées pour obtenir une estimation de ce qui aurait été obtenu si les images entières avaient été analysées.

Lorsque la résolution temporelle est modifiée, le flux d'image est réduit de manière régulière ou irrégulière et seules les images conservées sont analysées pendant la première passe. Comme précédemment, les statistiques des images qui ne sont pas analysées sont extrapolées.

Les deux méthodes peuvent être combinées, permettant ainsi de réduire encore plus le temps d'analyse. Elles sont toutes deux basées sur l'idée qu'il est possible d'extrapoler les données manquantes à partir des données analysées.

On observe cependant qu'une opération d'échantillonnage combinée à une extrapolation peut conduire à des statistiques d'une qualité faible, ce qui conduit à ne pas utiliser ce procédé dans l'immense majorité des applications pratiques.

Il existe ainsi un besoin pour un procédé d'analyse d'une séquence vidéo amélioré ne présentant pas les inconvénients exposés ci-dessus.

Il existe en outre un besoin pour un procédé d'analyse de séquence vidéo

amélioré dans la diminution du temps de calcul d'une phase d'analyse de la séquence dans le cadre d'un traitement multi-passe.

Un objet de la présente invention est de proposer un procédé d'analyse amélioré d'une séquence vidéo dans le cadre d'un traitement multi-passe.

5 Selon un premier aspect, il est proposé un procédé d'analyse d'un ensemble d'images d'une séquence vidéo en vue d'un traitement à effectuer sur la séquence, le procédé comprenant la détermination dans la séquence vidéo d'une pluralité de sous-séquences consécutives disjointes d'une ou plusieurs images successives, et l'analyse des images de chaque sous-séquence déterminée dans
10 la séquence vidéo, dans lequel les sous-séquences sont déterminées en fonction du type de traitement à effectuer et en fonction du contenu de la séquence vidéo.

On constate que les méthodes d'analyse conventionnelles méconnaissent plusieurs aspects importants, ce qui nuit fortement à leur mise en œuvre :

Dans le cas général, les images ne sont pas homogènes spatialement.
15 Leur centre par exemple, qui est le point d'intérêt privilégié, n'a pas la même complexité que leurs bords. Les résultats d'analyse, pouvant comprendre des données statistiques extraites après un sous-échantillonnage spatial, n'ont souvent qu'un rapport lointain avec les résultats d'une analyse effectuée sur les images initiales. Un exemple typique est le bruit, qui change de caractéristiques lorsque la
20 taille des images est réduite.

Le sous-échantillonnage temporel pose également le problème d'homogénéité du contenu qui rend difficile l'extrapolation des statistiques. En effet, tous les traitements se basant sur la cohérence temporelle des séquences vidéo perdent en efficacité voire deviennent inapplicables lorsque les séquences sont
25 sous-échantillonnées temporellement. C'est le cas par exemple de l'estimation de mouvement d'un compresseur vidéo qui perd en précision à mesure que la distance temporelle entre les images augmente.

Les procédés conventionnels peuvent ainsi mener à une régulation de débit peu précise, et qui est difficilement utilisable sans apprentissage préalable, ce qui
30 explique l'utilisation limitée qui est faite de ces procédés dans des applications pratiques.

Le procédé proposé présente l'avantage de privilégier un sous-échantillonnage temporel, afin d'éviter les écueils inhérents au sous-échantillonnage spatial évoqués ci-dessus.

35 Le procédé proposé tient en outre avantageusement compte du contenu de la séquence vidéo pour son analyse, les inventeurs ayant identifié le problème

d'homogénéité du contenu mentionné ci-dessus. Le procédé proposé facilite dès lors l'extrapolation des statistiques lorsqu'un sous-échantillonnage temporel est mis en œuvre lors d'une phase d'analyse d'une séquence vidéo. Le procédé proposé peut ainsi par exemple tenir compte du type de contenu (film, sport, spectacle musical, etc.) analysé dans le cadre de la première passe d'un traitement multi-passe.

De plus, au contraire du procédé conventionnel du document US 2010/0027622, qui fait partie intégrante d'un encodeur multi-passe et n'est donc pas d'un usage général, le procédé proposé tient avantageusement compte du type de traitement à effectuer sur la séquence vidéo (compression, débruitage, etc.) lors d'une phase de traitement utilisant les résultats d'analyse générés lors d'une phase d'analyse utilisant le procédé proposé.

Le procédé proposé présente de ce fait l'avantage de pouvoir être adapté selon le traitement à effectuer pour une séquence vidéo, en tenant compte pour la phase d'analyse de la séquence du type de traitement (compression, filtrage, etc.) réalisé ultérieurement sur la séquence.

Le procédé proposé convient particulièrement bien, bien que de façon non exclusive, pour l'encodage ou la compression d'une séquence vidéo selon un schéma du type H.264, H. 265, H. 262, MPEG-2, AVC, ou HEVC. Mais il convient aussi pour l'encodage d'images selon tout schéma d'encodage vidéo en deux passes (une passe d'analyse et une passe d'encodage), ou pour tout traitement d'une séquence vidéo en deux passes.

Dans une mise en œuvre particulière du procédé proposé, les tailles respectives des sous-séquences et les écarts respectifs entre deux sous-séquences voisines sont déterminés en fonction du type de traitement à effectuer et en fonction du contenu de la séquence vidéo.

Dans une mise en œuvre particulière du procédé proposé, les sous-séquences pourront avoir une taille identique, à l'exception de la dernière sous-séquence de la pluralité de sous-séquences consécutives.

De préférence, la taille de la dernière sous-séquence sera choisie supérieure ou égale à la taille des autres sous-séquences, que les autres sous-séquences aient une taille unique ou non.

Dans une mise en œuvre particulière du procédé proposé, à l'exception de l'écart entre la dernière sous-séquence de la pluralité de sous-séquences consécutives et la sous-séquence voisine de la dernière sous-séquence, les écarts respectifs entre deux sous-séquences voisines pourront être choisis identiques.

Dans une mise en œuvre particulière, le procédé proposé comprend en outre la génération, par extrapolation des résultats d'analyse des sous-séquences de la séquence vidéo, des résultats d'analyse de la séquence vidéo.

5 Dans une mise en œuvre particulière du procédé proposé, au moins une des sous-séquences pourra ne contenir qu'une seule image.

Dans une mise en œuvre particulière du procédé proposé, les sous-séquences sont en outre déterminées en fonction de la vitesse d'analyse ou de la précision de l'analyse.

10 Selon un deuxième aspect, il est proposé un dispositif d'analyse d'un ensemble d'images d'une séquence vidéo en vue d'un traitement à effectuer sur la séquence, comprenant une interface d'entrée configurée pour recevoir la séquence vidéo, et une unité d'analyse de séquence, comprenant un processeur couplé de manière opérationnelle à une mémoire, configurée pour déterminer dans la séquence vidéo une pluralité de sous-séquences consécutives disjointes d'une ou
15 plusieurs images successives, en fonction du type de traitement à effectuer et en fonction du contenu de la séquence vidéo et analyser les images de chaque sous-séquence déterminée dans la séquence vidéo.

Selon un autre aspect, il est proposé un programme d'ordinateur, chargeable dans une mémoire associée à un processeur, et comprenant des
20 portions de code pour la mise en œuvre des étapes du procédé proposé lors de l'exécution dudit programme par le processeur, ainsi qu'un ensemble de données représentant, par exemple par voie de compression ou d'encodage, ledit programme d'ordinateur.

Un autre aspect concerne un support de stockage non-transitoire d'un
25 programme exécutable par ordinateur, comprenant un ensemble de données représentant un ou plusieurs programmes, lesdits un ou plusieurs programmes comprenant des instructions pour, lors de l'exécution desdits un ou plusieurs programmes par un ordinateur comprenant une unité de traitement couplée de manière opérationnelle à des moyens mémoire et à un module d'interface
30 entrées/sorties, conduire l'ordinateur à analyser les images d'une séquence vidéo selon le procédé proposé.

D'autres particularités et avantages de la présente invention apparaîtront dans la description ci-après d'exemples de réalisation non limitatifs, en référence aux dessins annexés, dans lesquels :

35 - la figure 1 est un schéma illustrant l'architecture d'un dispositif d'analyse de séquence vidéo selon un mode de réalisation du procédé proposé ;

- la figure 2 est un schéma illustrant l'architecture d'un dispositif de traitement de séquence vidéo selon un mode de réalisation du procédé proposé ;
- la figure 3 est un diagramme illustrant le procédé proposé selon un mode de réalisation ;
- les figures 4a et 4b sont des diagrammes illustrant le sous-échantillonnage d'une séquence vidéo selon un mode de réalisation du procédé proposé ;
- les figures 5a, 5b, 5c et 5d sont des diagrammes illustrant le sous-échantillonnage d'une séquence vidéo selon un mode de réalisation du procédé proposé ;
- la figure 6 représente un mode de réalisation d'un système informatique pour la mise en œuvre du procédé proposé.

Dans la description détaillée ci-après de modes de réalisation de l'invention, de nombreux détails spécifiques sont présentés pour apporter une compréhension plus complète. Néanmoins, l'homme du métier peut se rendre compte que des modes de réalisation peuvent être mis en pratique sans ces détails spécifiques. Dans d'autres cas, des caractéristiques bien connues ne sont pas décrites en détail pour éviter de compliquer inutilement la description.

On entend ici par « sous-échantillonnage » toute opération réalisant l'extraction ou la sélection de sous-séquences au sein d'une séquence vidéo, sans limitation relative à la méthode particulière ou à un paramètre particulier de sous-échantillonnage (période, récurrence, etc.), sauf indication expresse. Le sous-échantillonnage se distingue ainsi de la décimation, qui suppose une extraction régulière de sous-séquences (par exemple, extraction d'une image toutes les n images de la séquence vidéo).

En référence à la figure 1, le dispositif d'analyse de séquence vidéo 100 reçoit en entrée 102 une séquence vidéo d'entrée 101 à analyser dans le cadre d'un traitement multi-passe. Le dispositif d'analyse 100 comprend un contrôleur 103, couplé de manière opérationnelle à l'interface d'entrée 102, qui pilote une unité de sous-échantillonnage 104 et une unité d'analyse 105. Les données reçues sur l'interface d'entrée 102 sont transmises en entrée de l'unité de sous-échantillonnage.

L'unité de sous-échantillonnage 104 effectue un sous-échantillonnage de la séquence vidéo selon le procédé proposé en déterminant dans la séquence vidéo une pluralité de sous-séquences consécutives disjointes d'une ou plusieurs images

successives. L'unité 104 génère après sous-échantillonnage des données représentant la pluralité de sous-séquences consécutives et disjointes d'images de la séquence déterminée, qui sont traitées par le contrôleur 103 qui fournit en entrée de l'unité d'analyse 105 les images de la pluralité de sous-séquences disjointes d'images de la séquence vidéo sélectionnées par l'unité de sous-échantillonnage 104. L'unité d'analyse 105 génère après analyse des images de la pluralité de sous-séquences d'images de la séquence vidéo reçue en entrée des données statistiques 107 relatives à la séquence vidéo d'entrée 101, qui sont fournies par le contrôleur 103 sur une interface de sortie 106 du dispositif d'analyse 100.

Dans un ou plusieurs modes de réalisation du procédé proposé, la génération des données statistiques 107 peut inclure une extrapolation de données statistiques extraites en utilisant des résultats de l'analyse des images de la pluralité de sous-séquences d'images de la séquence vidéo reçues en entrée afin d'obtenir des résultats d'analyse pour l'ensemble des images de la séquence vidéo, et non pas seulement pour celles ayant été effectivement analysées. Par « extrapolation », on entend ici toute opération permettant de générer des données statistiques pour les images n'ayant pas été analysées (c'est-à-dire les images n'ayant pas été sélectionnées lors du sous-échantillonnage de la séquence) en utilisant notamment des données statistiques extraites pour les images analysées (c'est-à-dire les images de la pluralité de sous-séquences d'images de la séquence vidéo). Le dispositif d'analyse 100 peut ainsi fournir en sortie un ensemble de statistiques qui ne font pas apparaître le découpage en sous-séquences de la séquence vidéo initiale. Les traitements ultérieurs peuvent alors être effectués en utilisant les résultats de la phase d'analyse générés par le dispositif d'analyse 100 sans qu'ils aient à connaître le découpage de la séquence vidéo réalisé par le dispositif d'analyse 100.

Le contrôleur 103 est configuré pour piloter l'unité de sous-échantillonnage 104 et l'unité d'analyse 105, et en particulier les entrées/sorties de ces unités. L'architecture du dispositif d'analyse 100 illustrée sur la figure 1 n'est toutefois pas limitative. Par exemple, l'interface d'entrée 102 du dispositif d'analyse 100 pourrait être couplée de manière opérationnelle à une interface d'entrée de l'unité de sous-échantillonnage 104. De même, l'unité de sous-échantillonnage 104 pourrait comprendre une sortie couplée de manière opérationnelle à une entrée de l'unité d'analyse 105, et l'unité d'analyse 105 pourrait comprendre une sortie couplée de manière opérationnelle à l'interface de sortie 106.

Le dispositif d'analyse 100 peut être un ordinateur, un réseau d'ordinateurs,

un composant électronique, ou un autre appareil comportant un processeur couplé de manière opérationnelle à une mémoire, ainsi que, selon le mode de réalisation choisi, une unité de stockage de données, et d'autres éléments matériels associés comme une interface de réseau et un lecteur de support pour lire un support non-transitoire de stockage amovible et écrire sur un tel support (non représentés sur la figure). Le support de stockage amovible peut être, par exemple, un disque compact (CD), un disque vidéo/polyvalent numérique (DVD), un disque flash, une clé USB, etc.

En fonction du mode de réalisation, la mémoire, l'unité de stockage de données ou le support de stockage amovible contient des instructions qui, lorsqu'elles sont exécutées par le contrôleur 103, amènent ce contrôleur 103, à effectuer ou contrôler les parties interface d'entrée 102, sous-échantillonnage 104, analyse 105 et/ou interface de sortie 106 des exemples de mise en œuvre du procédé proposé décrits dans les présentes. Le contrôleur 103 peut être un composant implémentant un processeur ou une unité de calcul pour l'analyse d'images selon le procédé proposé et le contrôle des unités 102, 104, 105 et 106 du dispositif d'analyse 100.

Le dispositif d'analyse 100 peut ainsi être mis sous forme d'un logiciel qui, lorsqu'il est chargé dans une mémoire et exécuté par un processeur, met en œuvre l'analyse d'une séquence vidéo selon le procédé proposé.

En outre, le dispositif d'analyse 100 peut être mis en œuvre sous forme logicielle, comme décrit ci-dessus, ou sous forme matérielle, comme un circuit intégré spécifique application (ASIC), ou sous forme d'une combinaison d'éléments matériels et logiciels, comme par exemple un programme logiciel destiné à être chargé et exécuté sur un composant de type FPGA (Field Programmable Gate Array).

La figure 2 est un diagramme illustrant un dispositif de traitement de séquence vidéo. En référence à la figure 2, le dispositif de traitement de séquence vidéo 200 reçoit en entrée 202 une séquence vidéo d'entrée 201 à traiter dans le cadre d'un traitement multi-passe. Le dispositif d'analyse 100 comprend un contrôleur 203, couplé de manière opérationnelle à l'interface d'entrée 202, qui pilote une unité de traitement 204 et une unité d'analyse 205. Les données reçues sur l'interface d'entrée 202 sont transmises en entrée de l'unité d'analyse 205. Les données de sortie 207 du dispositif de traitement 200 sont générées sur une interface de sortie 206.

L'ensemble contrôleur 203, unité d'analyse 205 et interfaces d'entrée et de

sortie 202/206 forme une unité d'analyse pouvant correspondre à l'unité d'analyse 100 décrite en référence à la figure 1, et configurée pour mettre en œuvre le procédé proposé.

5 L'unité d'analyse 205 génère après sous-échantillonnage de la séquence vidéo selon le procédé proposé des données représentant une pluralité de sous-séquences d'images de la séquence déterminée par le sous-échantillonnage, qui sont traitées par le contrôleur 203 ou par un contrôleur de l'unité d'analyse, pour analyse des images des sous-séquences. Dans un ou plusieurs mode de réalisation du procédé proposé, les résultats d'analyse des images des sous-séquences issues du sous-échantillonnage peuvent être extrapolés afin de générer des résultats d'analyse de l'ensemble des images de la séquence vidéo.

10 Le contrôleur 203 est configuré pour piloter l'unité d'analyse 205 et l'unité de traitement 204, et en particulier les entrées/sorties de ces unités. Dans un ou plusieurs modes de réalisation, les résultats d'analyse produits par l'unité d'analyse 205 sont fournis, sous la supervision du contrôleur 205, en entrée de l'unité de traitement 204 pour traitement de la séquence vidéo 201 d'entrée dans le cadre d'un traitement multi-passe, l'analyse effectuée par l'unité d'analyse 205 correspondant à une première passe et le traitement effectué par l'unité de traitement correspondant à une deuxième passe.

20 L'architecture du dispositif de traitement 200 illustrée sur la figure 2 n'est toutefois pas limitative. Par exemple, l'interface d'entrée 202 du dispositif de traitement 200 pourrait être couplée de manière opérationnelle à une interface d'entrée de l'unité d'analyse 205 et à une interface d'entrée de l'unité de traitement 204. De même, l'unité d'analyse 205 pourrait comprendre une sortie couplée de manière opérationnelle à une entrée de l'unité de traitement 204, et l'unité de traitement 204 pourrait comprendre une sortie couplée de manière opérationnelle à l'interface de sortie 206 pour y produire des données correspondants à la séquence vidéo 201 traitée.

30 De plus, le dispositif de traitement 200 peut être un encodeur vidéo multi-passe, un dispositif de débruitage vidéo, ou tout autre dispositif de traitement multi-passe de séquence vidéo dans lequel au moins une passe comprend une analyse d'une séquence vidéo d'entrée préalablement à son traitement.

35 Le dispositif de traitement 200 peut être un ordinateur, un réseau d'ordinateurs, un composant électronique, ou un autre appareil comportant un processeur couplé de manière opérationnelle à une mémoire, ainsi que, selon le mode de réalisation choisi, une unité de stockage de données, et d'autres éléments

matériels associés comme une interface de réseau et un lecteur de support pour lire un support de stockage amovible et écrire sur un tel support (non représentés sur la figure). En fonction du mode de réalisation, la mémoire, l'unité de stockage de données ou le support de stockage amovible contient des instructions qui, lorsqu'elles sont exécutées par le contrôleur 203, amènent ce contrôleur 203, à effectuer ou contrôler les parties interface d'entrée 202, analyse 205, traitement 204 et/ou interface de sortie 206 des exemples de mise en œuvre du procédé proposé décrits dans les présentes. Le contrôleur 203 peut être un composant implémentant un processeur ou une unité de calcul pour le traitement d'images comprenant une analyse selon le procédé proposé et le contrôle des unités 202, 204, 205 et 206 du dispositif de traitement 200.

En outre, le dispositif de traitement 200 peut être mis en œuvre sous forme logicielle, comme décrit ci-dessus, ou sous forme matérielle, comme un circuit intégré spécifique application (ASIC), ou sous forme d'une combinaison d'éléments matériels et logiciels, comme par exemple un programme logiciel destiné à être chargé et exécuté sur un composant de type FPGA (Field Programmable Gate Array).

La figure 3 montre un diagramme illustrant le procédé proposé selon un mode de réalisation.

Un ou plusieurs paramètres relatifs au type de traitement à réaliser sur la séquence vidéo d'entrée sont entrés (301) afin d'être pris en compte lors de la phase de sous-échantillonnage. De même, un ou plusieurs paramètres relatifs au contenu de la séquence vidéo à analyser sont entrés (302) afin d'être pris en compte lors de la phase de sous-échantillonnage. La séquence vidéo à analyser est produite en entrée (303) afin d'être analysée. Cette séquence vidéo est ensuite sous-échantillonnée (304) pour y déterminer une pluralité de sous-séquences consécutives disjointes d'images d'une ou plusieurs images successives en fonction du type de traitement à effectuer et du contenu de la séquence vidéo, sur la base du/des paramètres relatifs au type de traitement et du/des paramètres relatifs au contenu de la séquence vidéo. Les images de chaque sous-séquence ainsi déterminée sont ensuite analysées (305) selon un procédé d'analyse prédéterminé correspondant au traitement à effectuer sur la séquence, pour fournir des résultats d'analyse des sous-séquences. Dans un ou plusieurs modes de réalisation, l'analyse des images des sous-séquences peut être suivie de l'extrapolation (306) des résultats de cette analyse pour générer des résultats d'analyse de l'ensemble des images de la séquence vidéo d'entrée.

Les figures 4a et 4b illustrent le sous-échantillonnage d'une séquence vidéo selon un mode de réalisation du procédé proposé.

La figure 4a représente schématiquement une séquence vidéo (400) comprenant un ensemble de N images réparties sur une durée D_s entre la première image (401) de la séquence et la dernière image (402) de la séquence.

La figure 4b représente la séquence vidéo (400) après sous-échantillonnage. Le sous-échantillonnage de la séquence (400) a opéré une détermination de sous-séquences consécutives et disjointes (403a, 403b, 403c, 403d) de la séquence, parfois référencées dans la présente demande sous les termes « chunk », « sous-ensemble » ou « paquet ».

Les sous-séquences (403a, 403b, 403c, 403d) déterminées sont disjointes, en ce que deux sous-séquences voisines sont respectivement séparées par des « trous » (404a, 404b, 404c), chaque trou contenant au moins une image de la séquence vidéo (400). Ces trous (404a, 404b, 404c) correspondent aux groupes d'images de la séquence initiale (400) qui, selon le procédé proposé, ne seront pas analysées.

Les sous-séquences (403a, 403b, 403c, 403d) déterminées sont consécutives, en ce qu'elles résultent de l'échantillonnage d'une séquence vidéo correspondant à une durée. L'échantillonnage de la séquence vidéo pourra être effectué selon le déroulement temporel de la séquence, pour déterminer une suite de sous-séquences, parmi lesquelles une première séquence correspond au début de la séquence vidéo, et une dernière séquence correspond à la fin de la séquence vidéo.

Les sous-séquences déterminées par le sous-échantillonnage ne sont pas nécessairement de tailles égales, en ce qu'elles ne contiennent par exemple pas toutes un même nombre d'images. Sur l'exemple illustré par la figure 4b, les sous-séquences 403a, 403b et 403c sont de taille égale, cette taille étant inférieure ou égale à celle de la dernière sous-séquence (403d) de la séquence (400).

Ainsi, dans un ou plusieurs modes de réalisation, le procédé réalise un découpage temporel en « chunks » de la séquence vidéo à analyser. Les chunks sont des sous-séquences d'images consécutives qui ne contiennent pas nécessairement le même nombre d'images. Les images qui ne sont pas dans des chunks ne subiront pas le traitement d'analyse.

Les tailles respectives de la séquence vidéo et des sous-séquences peuvent être exprimées en nombre d'images, ou bien sous la forme d'une durée, les deux mesures étant liées par le nombre d'images par seconde de la séquence

vidéo considérée. Il en est de même de l'écart entre deux sous-séquences voisines, qui peut être, en fonction de l'implémentation, exprimée en nombre d'images ou comme une durée.

Le sous-échantillonnage pourra être effectué sur la base de différents paramètres, en fonction de l'implémentation.

Dans un ou plusieurs modes de réalisation, on pourra déterminer les sous-séquences issues du sous-échantillonnage en fonction d'une taille de sous-séquence, par exemple exprimée par un nombre d'images, identique pour toutes les sous-séquences éventuellement à l'exception d'une sous-séquence (de préférence la dernière), et d'une fréquence de sous-échantillonnage ou d'une période de sous-échantillonnage définissant l'écart entre deux sous-séquences voisines.

En variante, on pourra déterminer les sous-séquences en fonction d'une taille de sous-séquence et d'un taux de sous-échantillonnage. Par exemple, on pourra déterminer les sous-séquences en fixant une taille unique (éventuellement à l'exception de la dernière sous-séquence) égale à une image, et un taux de sous-échantillonnage à 1/6. Les sous-séquences seront donc déterminées en sélectionnant une image toutes les 6 images de la séquence vidéo à traiter aux fins de l'analyse. Pour une séquence vidéo d'une durée d'une heure, la durée d'analyse sera ainsi réduite à 10 minutes. Des résultats d'analyse pour les parties non analysées (c'est-à-dire, ne correspondant pas à une sous-séquence) pourront être déduit des résultats d'analyse des sous-séquences, par exemple par un procédé d'extrapolation.

Dans un ou plusieurs modes de réalisation, on pourra déterminer les sous-séquences issues du sous-échantillonnage en fonction d'une taille de sous-séquence, par exemple exprimée par un nombre d'images, identique pour toutes les sous-séquences éventuellement à l'exception d'une sous-séquence, d'un nombre de sous-séquences, et de la taille de la séquence.

En fonction de l'implémentation, le procédé proposé pourra déterminer des paramètres de sous-échantillonnage en tenant compte du type de traitement à effectuer et du contenu de la séquence vidéo à traiter.

Dans un ou plusieurs modes de réalisation, la taille de la dernière sous-séquence, correspondant à la fin de la séquence vidéo à traiter, pourra être choisie supérieure ou égale à celles des autres sous-séquences. Ainsi, en fonction de la taille de la séquence, le procédé proposé pourra fusionner les deux dernières sous-séquences dans le cas où la taille de la dernière sous-séquence serait autrement

inférieure à un seuil prédéterminé, à la taille des autres sous-séquences si elle est unique, ou à la taille d'au moins une autre sous-séquence.

Dans le cas d'un sous-échantillonnage régulier de la séquence, la position d'un chunk dans une période de sous-échantillonnage pourra varier. La figure 5a illustre le cas d'une période de sous-échantillonnage (501) contenant un chunk (500) positionné en début de période.

La figure 5b illustre le cas d'une période de sous-échantillonnage (503) contenant un chunk (502) positionné en fin de période.

Comme expliqué ci-dessus, la taille du dernier chunk de la séquence pourra être choisie supérieure ou égale à celle des autres chunks. Comme illustré sur la figure 5c, dans le cas de chunks positionnés en début de période d'échantillonnage, le dernier chunk (504) pourra être prolongé pour couvrir toute la période (505), de manière à assurer une analyse plus précise de la fin de la séquence vidéo.

Il en sera de même du cas de chunks positionnés en fin de période, ou à toute autre position de la période, comme illustré sur la figure 5d. Le dernier chunk (506) pourra être prolongé pour couvrir toute la période (507), de manière à assurer une analyse plus précise de la fin de la séquence vidéo.

La répartition des chunks ainsi que leur taille exprimée en nombre d'images permet d'en déduire un gain de vitesse du procédé d'analyse. En effet, si la charge de calcul est proportionnelle au nombre d'images analysées (ce qui est généralement le cas dans la plupart des applications), la proportion d'images constituant les chunks peut directement fournir une estimation de ce gain.

Dans un ou plusieurs modes de réalisation, la taille des chunks ainsi que leur répartition est dépendante de l'application (du traitement effectué sur la séquence). Si, en fonction de l'application, la méthode d'analyse utilise des informations temporelles comme le mouvement d'objets dans la scène, les chunks seront de préférence constitués d'un nombre significatifs d'images consécutives. Un chunk pourra par exemple être considéré de grande taille lorsqu'il comprend plusieurs milliers d'images consécutives. Par exemple, dans le cadre d'une analyse effectuée pour une application de compression vidéo, certains chunks pourront comprendre environ 3000 images. Au contraire, si l'analyse traite les images indépendamment, comme par exemple une mesure de luminosité, les chunks pourront être de taille réduite, et par exemple être réduits à une seule image.

Dans un ou plusieurs modes de réalisation, les résultats d'analyse, par exemple les données statistiques générées par le procédé d'analyse, sont

extrapolées pour les sous-séquences d'images qui ne sont pas analysées. Il est préférable dans ce cas que les images ignorées soient de même nature que les images analysées du point de vue des statistiques renvoyées. Grâce à cette extrapolation, le procédé d'analyse peut fournir à l'étage de traitement qui suit un ensemble de statistiques complet, c'est-à-dire comprenant des statistiques correspondant aux images qui n'ont pas été analysées. Plusieurs méthodes d'extrapolation sont possibles et il est judicieux de choisir celles qui sont adaptées au type de statistiques à mesurer. Les statistiques peuvent par exemple être extrapolées linéairement. On estimera ainsi que leurs valeurs sont placées sur une droite qui relie la dernière image d'un chunk à la première image du chunk suivant.

Comme indiqué ci-dessus, la taille du dernier chunk sera de préférence choisie supérieure ou égale à celle des autres chunks, et par exemple suffisamment grande pour corriger les erreurs d'analyse de la séquence vidéo d'entrée. La taille de cette fenêtre d'analyse de fin peut en effet influencer la qualité d'analyse de l'ensemble du contenu, et il est préférable de choisir une taille suffisante pour pouvoir compenser les erreurs d'analyse au début de la séquence imputables au sous-échantillonnage et à la taille choisie pour les fenêtres d'analyse de début de la séquence.

L'opération de sous-échantillonnage d'une séquence vidéo tenant compte du type de traitement à effectuer et du contenu de la séquence vidéo est illustrée au travers des deux exemples suivants : le filtrage d'une séquence vidéo dans le contexte de la compression vidéo, et la compression vidéo en deux passes.

Filtrage d'un bruit pour débruiter une séquence vidéo dans le contexte de la compression vidéo

Un des aspects importants de l'appréciation de la qualité vidéo d'une séquence vidéo compressée est son homogénéité visuelle. La qualité vidéo sera d'autant plus élevée que les défauts inhérents à la compression sont uniformes au cours de la séquence vidéo compressée. Lorsqu'il est prévu de retirer du bruit avant d'effectuer la tâche de compression, celui-ci doit donc de préférence être retiré de façon homogène tout au long de la séquence vidéo à débruiter.

Une méthode de débruitage consiste à analyser préalablement toute la séquence vidéo de façon à identifier des caractéristiques du bruit (par exemple des caractéristiques statistiques du bruit), puis dans un second temps, d'effectuer un filtrage à l'aide des caractéristiques recueillies, pour réduire le bruit. Ce procédé de débruitage comprend donc une première phase d'analyse du bruit, suivie d'un traitement de débruitage, par exemple par filtrage du bruit, qui utilise les résultats

de la phase d'analyse préalable. Suivant le type de filtrage réalisé, différentes caractéristiques peuvent être acquises durant la phase d'analyse préalable, comme par exemple l'énergie du bruit et son amplitude spectrale lorsqu'il est considéré comme étant un bruit blanc additif gaussien.

5 Dans un ou plusieurs modes de réalisation du procédé proposé, on tient compte du type de contenu de la séquence vidéo à débiter pour déterminer dans la séquence vidéo une pluralité de sous-séquences consécutives disjointes d'une ou plusieurs images successives, lors de la phase d'analyse préalable au débruitage.

10 Par exemple, dans le cas où le contenu de la séquence vidéo à traiter est un film de cinéma, le bruit à retirer sera typiquement du grain de pellicule (argentique ou artificiellement ajouté dans le cas du cinéma numérique). Les caractéristiques de ce bruit peuvent être considérées comme étant homogènes tout au long du film. En revanche, il n'est pas ajouté de manière linéaire. Il peut aussi
15 être plus important dans les scènes sombres que dans les scènes claires par exemple. Dans ce cas des paramètres non-linéaires peuvent être calculés. Ils peuvent permettre une identification complète du bruit, en utilisant par exemple la technique décrite dans la demande de brevet américain US 2004/005365. Par exemple, des paramètres non-linéaires tels que ceux transportés dans les
20 standards de compression vidéo H.264 et HEVC pourront être calculés. Ceux-ci peuvent en effet être utilisés pour modéliser des bruits multiplicatifs avec des coefficients dépendant de l'intensité lumineuse.

Dans ce cas particulier non limitatif d'une séquence de vidéo dont le contenu est un film de cinéma, on peut effectuer pour l'analyse préalable de la
25 séquence des mesures de caractéristiques statistiques du bruit, de préférence à intervalle régulier, en utilisant des sous-séquences, ou « chunks », de quelques images. Un sous-échantillonnage de la séquence vidéo est effectué en déterminant dans la séquence vidéo une pluralité de sous-séquences consécutives disjointes, les sous-séquences comprenant une ou plusieurs images successives. Ces sous-
30 séquences sont déterminées en fonction du type de traitement à effectuer (dans cet exemple un filtrage du bruit) et en fonction du contenu de la séquence vidéo (dans cet exemple un film de cinéma). En particulier, on peut choisir selon le procédé proposé d'utiliser des sous-séquences de taille réduite à quelques images, du fait que pour l'exemple d'un film de cinéma on peut considérer que les caractéristiques
35 du bruit sont homogènes sur l'ensemble de la séquence vidéo. Dans un exemple particulier d'implémentation, l'utilisation de « chunks » de 4 images consécutives

permet d'acquérir des caractéristiques précises du bruit de films de cinéma, tout en s'affranchissant des variations locales très brutales de luminosité tels que les flashes.

De même, on peut choisir pour ce cas particulier un écart constant entre deux sous-séquences adjacentes, ainsi que la durée de cet écart. La durée moyenne des scènes dans les films de cinéma est en effet de l'ordre de deux secondes. Par conséquent, effectuer par exemple une mesure toutes les secondes permet d'acquérir de l'information sur toutes les scènes du film et d'obtenir ainsi un profil précis des caractéristiques du bruit sur l'ensemble de la séquence vidéo.

Comme discuté ci-dessus, la taille (exprimée en nombre d'images ou en unité de temps) de la dernière sous-séquence, ainsi que l'écart entre cette dernière sous-séquence et la sous-séquence précédente, pourront être choisis différents des autres sous-séquences déterminées par le sous-échantillonnage.

On pourra par exemple choisir des chunks de taille $N_c,1$, éventuellement à l'exception du dernier chunk de la séquence qui pourra être choisi de taille supérieure ou égale $N_c,2$ afin d'assurer une qualité globale d'analyse comme discuté ci-dessus. L'échantillonnage pourra être effectué toutes les secondes ce qui permettra de calculer le nombre total de chunks en fonction de la durée totale D_s de la séquence vidéo.

Dans un ou plusieurs modes de réalisation, une extrapolation des résultats d'analyse de chunks est effectuée. Différents procédés d'extrapolation des statistiques acquises peuvent, le cas échéant, être employés. Par exemple, on peut considérer que ces statistiques sur les images non-analysées sont identiques à celles des images des extrémités des chunks adjacents. Cette méthode est particulièrement valide lorsque les chunks sont composés de peu d'images, comme c'est le cas dans l'exemple de réalisation décrit précédemment où ils sont limités à 4 images.

Dans un exemple de mise en oeuvre, pour une séquence vidéo à 24 Hz (24 images par seconde) correspondant à un film, le procédé d'analyse pourra être configuré pour effectuer des mesures sur 4 images consécutives (les chunks seront donc de taille égale à 4 images, éventuellement à l'exception du dernier chunk) toutes les 24 images qui représentent une seconde de film. L'analyse est donc réalisée 6 fois plus rapidement que si l'ensemble des images de la séquence étaient analysées. Dans ce mode de réalisation, le découpage de la séquence est adapté au contenu comme au type de traitements réalisés ultérieurement. En effet, la période des *chunks* comme leurs durées proposées ci-dessus sont bien adaptés

à des séquences vidéo de type film.

Dans le cas d'autres types de contenus, comme des événements sportifs télévisés, on choisira une période de découpage plus longue car le montage de ce type de séquence vidéo est nettement moins rapide. De ce fait, l'invention permet d'optimiser la vitesse d'analyse en fonction du type de contenu, ce qui constitue un avantage important dans le cas de traitements par lots.

Compression vidéo en deux passes

La compression vidéo en deux passes a généralement pour objectif de garantir une qualité vidéo constante sur toute la séquence vidéo compressée, tout en respectant une taille totale spécifiée. La première passe est une phase d'analyse, ayant par exemple pour objet la détermination de la complexité de la séquence. La seconde passe est la phase de compression proprement dite. Cette dernière utilise les résultats de la phase d'analyse, dans l'exemple précédent les complexités obtenues, pour optimiser le débit localement et maintenir à la fois une qualité constante et une taille totale.

Dans un ou plusieurs modes de réalisation, la phase d'analyse peut comprendre la détermination à chaque image d'une relation entre un paramètre de compression, un quantificateur, et un débit correspondant. Cette relation est obtenue généralement en compressant la séquence vidéo, de sorte que la phase d'analyse s'avère être quasiment aussi coûteuse que la phase de traitement (compression) elle-même.

Plusieurs relations entre un paramètre de compression, un quantificateur, et un débit correspondant, ont été proposées. On peut par exemple considérer des fonctions du type : $Q_p = a \cdot \log(\text{Débit}) + b$, où Q_p est un quantificateur moyen des images de la séquence vidéo à compresser, et a et b sont deux paramètres à identifier dépendant des images de la séquence vidéo à encoder. En variante, on peut aussi considérer une relation entre un paramètre de compression, un quantificateur, et un débit correspondant, construite par extrapolation logarithmique à partir de points $Q_p, \text{Débit}$ mesurés.

Dans un ou plusieurs modes de réalisation, le résumé statistique réalisé à l'issue de la phase d'analyse est constitué pour chaque image d'une relation *Quantificateur, Débit*. Selon le procédé proposé, on détermine dans la séquence vidéo à compresser une pluralité de sous-séquences consécutives disjointes d'une ou plusieurs images successives. Les images de chaque sous-séquence sont analysées pour produire, pour chaque image de chaque sous-séquence analysée,

une relation *Quantificateur, Débit*.

Dans un ou plusieurs modes de réalisation, des statistiques des parties non analysées peuvent être générées par extrapolation en utilisant les résultats d'analyse des sous-séquences. Dans l'exemple précité, plusieurs méthodes de
 5 génération de résultats d'analyse pour les parties de la séquence qui n'ont pas été analysées sont possibles. Par exemple, des résultats d'analyse peuvent être générés en interpolant linéairement les relations *Quantificateur, Débit* obtenues par l'analyse des images des sous-séquences. Cela revient à faire varier linéairement les paramètres a et b de la relation $Q_p = a \cdot \log(\text{Débit}) + b$ décrite
 10 précédemment.

Le procédé proposé permet de réduire considérablement la complexité de cette phase d'analyse du fait qu'il est procédé, dans le cadre de la phase d'analyse, à un sous-échantillonnage qui tient compte du type de traitement à effectuer (dans cet exemple une compression) et du contenu de la séquence vidéo. Dans un ou
 15 plusieurs modes de réalisation, des paramètres du sous-échantillonnages sont choisis en tenant compte du type de traitement à effectuer et du contenu de la séquence vidéo à traiter.

Par exemple, dans le cas où la séquence vidéo à compresser est un film de cinéma (par exemple dans le cadre d'une application de vidéo à la demande), la relation entre le(s) paramètre(s) de compression et le débit obtenu varie peu. En
 20 d'autres termes, les films présentent des caractéristiques sensiblement homogènes, au moins localement. De ce fait, il n'est pas nécessaire d'effectuer l'analyse sur tout le film car des sous-séquences de quelques secondes correctement réparties suffisent à fournir l'information nécessaire à la seconde
 25 passe.

La séquence vidéo contenant un film peut par exemple être découpée en sous-séquences de durée égale à dix secondes, l'analyse n'étant effectuée que toutes les minutes de contenu, ce qui détermine l'écart entre deux sous-séquences voisines. Des résultats d'analyse (par exemple la complexité) pour les zones qui ne
 30 sont pas analysées peuvent être obtenus par extrapolation des résultats d'analyse des sous-séquences, ce qui introduit une imprécision qui pourra être compensée lors de la phase de traitement. Dans un ou plusieurs modes de réalisation, pour faciliter cette compensation et obtenir la taille requise, l'extrême fin de la séquence est complètement analysée, en choisissant une taille pour la dernière sous-
 35 séquence plus importante que celle des autres sous-séquences. Par exemple, la dernière sous-séquence peut être choisie correspondant aux dernières minutes de

la séquence, dans le cas où le contenu de celle-ci est un film. Cette dernière séquence peut en particulier être composée des deux dernières minutes de la séquence, soit environ 3000 images.

Dans le cas d'un film de deux heures, le choix des paramètres de sous-échantillonnage proposé ci-dessus pour tenir compte du contenu de la séquence dans le cadre d'un traitement de compression (tailles des sous-séquences et écarts entre deux sous-séquences voisines) conduit à n'analyser qu'environ 22 minutes du film, soit 18% seulement des images. Le gain en termes de temps de calcul est donc d'environ 5.5.

Des études sur des centaines de films de cinéma ont montré que cette méthode de découpage permettait d'obtenir une qualité indiscernable d'une analyse complète.

Dans un ou plusieurs modes de réalisation, il est en outre tenu compte d'un objectif de vitesse de l'analyse, de précision de l'analyse, ou d'un critère représentant un compromis entre la vitesse et la précision de l'analyse. Il est en effet possible de modifier la répartition des zones analysées pour accélérer encore la phase d'analyse, ou améliorer la précision de l'analyse. Par exemple, on détermine un compromis entre la vitesse de l'analyse et sa précision, et on modifie la répartition des sous-séquences déterminées en fonction du type de traitement à effectuer et en fonction du contenu de la séquence vidéo à traiter, selon que l'on souhaite privilégier la vitesse de l'analyse, ou la qualité de celle-ci.

Lorsque la vitesse d'analyse est privilégiée, la durée respective des sous-séquences déterminée en fonction du type de traitement à effectuer et en fonction du contenu de la séquence vidéo à traiter peut être réduite, par exemple d'un facteur prédéterminé correspondant au gain de vitesse d'analyse recherché. Dans un exemple particulier de réalisation, ceci peut être obtenu en réduisant la durée des sous-séquences analysées de moitié, en réduisant par exemple la durée des sous-séquences de 10s à 5s.

A l'inverse, dans le cas où la précision de l'analyse est privilégiée, ce qui peut être rendu nécessaire lorsque les contenus sont très inhomogènes (comme c'est le cas par exemple de séquences d'images de télévision) la durée respective des sous-séquences déterminée en fonction du type de traitement à effectuer et en fonction du contenu de la séquence vidéo à traiter peut être augmentée, par exemple d'un facteur prédéterminé correspondant au gain de précision d'analyse recherché. Dans un exemple particulier de réalisation, ceci peut être obtenu en augmentant la durée des sous-séquences analysées ou en augmentant la durée de

la sous-séquence finale.

Des modes de réalisation du procédé d'analyse d'une séquence vidéo peuvent être, au moins pour partie, mis en œuvre sur pratiquement n'importe quel type d'ordinateur, indépendamment de la plate-forme utilisée. Par exemple, comme
5 cela est représenté sur la figure 6, un système informatique (600), qui peut correspondre aux unités d'analyse de séquence vidéo et de traitement de séquence vidéo représentées sur les figures 1 et 2 ou être couplé de manière opérationnelle à ces éléments, comprend une unité de traitement de données (601), qui comprend un ou plusieurs processeurs (602), comme une unité centrale
10 (CPU) ou un autre processeur matériel, une mémoire associée (603) (par exemple, une mémoire vive (RAM), une mémoire cache, une mémoire flash, etc.), un dispositif de stockage (604) (par exemple un disque dur, un disque optique comme un CD ou un DVD, une clé de mémoire flash, etc.), et de nombreux autres éléments et fonctionnalités typiques des ordinateurs actuels (non représentés).

15 L'unité de traitement de données (601) comprend également un module d'interface entrées/sorties (605) qui pilote les différentes interfaces entre l'unité (601) et des moyens d'entrée et/ou de sorties du système (600). Le système (600) peut en effet également comprendre des moyens d'entrée, comme un clavier (606), une souris (607), ou un microphone (non représenté). En outre, l'ordinateur (600)
20 peut comprendre des moyens de sortie, comme un moniteur (608) (par exemple, un moniteur d'affichage à cristaux liquides (LCD), un moniteur d'affichage à LEDs ou un moniteur à tube cathodique (CRT)). Le système informatique (600) peut être relié à un réseau (609) (par exemple, un réseau local (LAN), un réseau étendu (WAN) comme Internet, ou n'importe quel autre type similaire de réseau) par
25 l'intermédiaire d'une connexion d'interface de réseau (non représentée). L'homme du métier peut se rendre compte qu'il existe de nombreux types différents de systèmes informatiques (par exemple, un ordinateur de bureau, un ordinateur portable, ou n'importe quel autre système informatique capable d'exécuter des instructions lisibles par ordinateur), et les moyens d'entrée et de sortie
30 susmentionnés peuvent prendre d'autres formes, actuellement connues ou développées ultérieurement.

En général, le système informatique (600) comprend au moins les moyens minimaux de traitement, d'entrée et/ou de sortie nécessaires pour mettre en pratique un ou plusieurs modes de réalisation du procédé d'analyse proposé. Par
35 exemple, le processeur (602) est apte à être configuré pour exécuter un programme d'ordinateur comprenant des portions de code pour la mise en œuvre

d'un analyseur, configuré pour effectuer l'analyse d'une séquence vidéo fournie en entrée selon les différents mode de réalisation du procédé d'analyse proposé. Le dispositif de stockage (604) sera choisi de préférence apte à stocker les données correspondant aux résultats de l'analyse et du traitement de la séquence vidéo.

5 L'homme du métier peut se rendre compte qu'un ou plusieurs éléments du système informatique (600) susmentionné peuvent se trouver à un emplacement distant et être reliés à d'autres éléments sur un réseau. De plus, un ou plusieurs modes de réalisation peuvent être mis en œuvre sur un système distribué comportant une pluralité de nœuds, où chaque portion de la mise en œuvre peut
10 être située sur un nœud différent à l'intérieur du système distribué. Dans un ou plusieurs modes de réalisation, le nœud correspond à un système informatique. En variante, le nœud peut correspondre à un processeur avec une mémoire physique associée. Le nœud peut également correspondre à un processeur avec une mémoire partagée et/ou des ressources partagées. En outre, des instructions
15 logicielles pour effectuer un ou plusieurs modes de réalisation peuvent être stockées sur un support non transitoire lisible par ordinateur comme un disque compact (CD), une disquette, une bande, ou n'importe quel autre dispositif de stockage lisible par ordinateur.

En fonction du mode de réalisation choisi, certains actes, actions, évènements ou fonctions de chacune des méthodes décrites dans le présent document peuvent être effectués ou se produire selon un ordre différent de celui dans lequel ils ont été décrits, ou peuvent être ajoutés, fusionnés ou bien ne pas être effectués ou ne pas se produire, selon le cas. En outre, dans certains modes de réalisation, certains actes, actions ou évènements sont effectués ou se
25 produisent concurremment et non pas successivement.

Bien que décrits à travers un certain nombre d'exemples de réalisation détaillés, le procédé d'encodage proposé et l'équipement pour la mise en œuvre du procédé comprennent différentes variantes, modifications et perfectionnements qui apparaîtront de façon évidente à l'homme de l'art, étant entendu que ces
30 différentes variantes, modifications et perfectionnements font partie de la portée de l'invention, telle que définie par les revendications qui suivent. De plus, différents aspects et caractéristiques décrits ci-dessus peuvent être mis en œuvre ensemble, ou séparément, ou bien substitués les uns aux autres, et l'ensemble des différentes combinaisons et sous combinaisons des aspects et caractéristiques font partie de
35 la portée de l'invention. En outre, il se peut que certains systèmes et équipements décrits ci-dessus n'incorporent pas la totalité des modules et fonctions décrits pour

les modes de réalisation préférés.

REVENDEICATIONS

1. Procédé d'analyse d'un ensemble d'images d'une séquence vidéo en vue d'un traitement à effectuer sur la séquence, le procédé comprenant :

5 déterminer dans la séquence vidéo une pluralité de sous-séquences consécutives disjointes d'une ou plusieurs images successives, lesdites sous-séquences consécutives disjointes étant séparées par des trous contenant au moins une image ;

10 analyser les images de chaque sous-séquence déterminée dans la séquence vidéo ;

générer, par extrapolation des résultats d'analyse des sous-séquences de la séquence vidéo, des résultats d'analyse de l'ensemble des images de la séquence vidéo ;

15 dans lequel les sous-séquences sont déterminées en fonction du type de traitement à effectuer et en fonction du contenu de la séquence vidéo,

dans lequel les sous-séquences sont en outre déterminées en fonction de la vitesse d'analyse, et

20 dans lequel la durée respective des sous-séquences est déterminée en fonction du type de traitement à effectuer et en fonction du contenu de la séquence vidéo, et est réduite d'un facteur prédéterminé correspondant à un gain de vitesse d'analyse recherché.

2. Procédé selon la revendication 1, dans lequel les tailles respectives des sous-séquences et les écarts respectifs entre deux sous-séquences voisines sont
25 déterminés en fonction du type de traitement à effectuer et en fonction du contenu de la séquence vidéo.

3. Procédé selon la revendication 2, dans lequel, à l'exception de la dernière sous-séquence de la pluralité de sous-séquences consécutives, les sous-
30 séquences ont une taille identique.

4. Procédé selon la revendication 3, dans lequel la taille de la dernière sous-séquence est choisie supérieure ou égale à la taille des autres sous-séquences.

5 5. Procédé selon l'une quelconque des revendications 2 à 4, dans lequel, à l'exception de l'écart entre la dernière sous-séquence de la pluralité de sous-séquences consécutives et la sous-séquence voisine de la dernière sous-séquence, les écarts respectifs entre deux sous-séquences voisines sont identiques.

10

6. Procédé selon l'une quelconque des revendications précédentes, dans lequel au moins une des sous-séquences contient une seule image.

15 7. Procédé selon l'une quelconque des revendications précédentes, dans lequel les sous-séquences sont en outre déterminées en fonction de la précision de l'analyse.

8. Dispositif d'analyse d'un ensemble d'images d'une séquence vidéo en vue d'un traitement à effectuer sur la séquence, comprenant :

- 20
- une interface d'entrée configurée pour recevoir la séquence vidéo ;
 - une unité d'analyse de séquence, comprenant un processeur couplé de manière opérationnelle à une mémoire, configurée pour :
 - 25 ○ déterminer dans la séquence vidéo une pluralité de sous-séquences consécutives disjointes d'une ou plusieurs images successives, en fonction du type de traitement à effectuer et en fonction du contenu de la séquence vidéo, lesdites sous-séquences consécutives disjointes étant séparées par des trous contenant au moins une image;
 - analyser les images de chaque sous-séquence déterminée dans la séquence vidéo ;

○ générer, par extrapolation des résultats d'analyse des sous-séquences de la séquence vidéo, des résultats d'analyse de l'ensemble des images de la séquence vidéo,

5 dans lequel les sous-séquences sont en outre déterminées en fonction de la vitesse d'analyse, et

dans lequel la durée respective des sous-séquences est déterminée en fonction du type de traitement à effectuer et en fonction du contenu de la séquence vidéo, et est réduite d'un facteur prédéterminé correspondant à un gain de vitesse d'analyse recherché.

10

9. Programme d'ordinateur, chargeable dans une mémoire associée à un processeur, et comprenant des portions de code pour la mise en œuvre des étapes d'un procédé selon l'une quelconque des revendications 1 à 7 lors de l'exécution dudit programme par le processeur.

15

10. Ensemble de données représentant, par exemple par voie de compression ou d'encodage, un programme d'ordinateur selon la revendication 9.

20 11. Support de stockage non-transitoire d'un programme exécutable par ordinateur, comprenant un ensemble de données représentant un ou plusieurs programmes, lesdits un ou plusieurs programmes comprenant des instructions pour, lors de l'exécution desdits un ou plusieurs programmes par un ordinateur comprenant une unité de traitement couplée de manière opérationnelle à des moyens mémoire et à un module d'interface entrées/sorties, conduire l'ordinateur à
25 analyser les images d'une séquence vidéo selon le procédé de l'une quelconque des revendications 1 à 7.

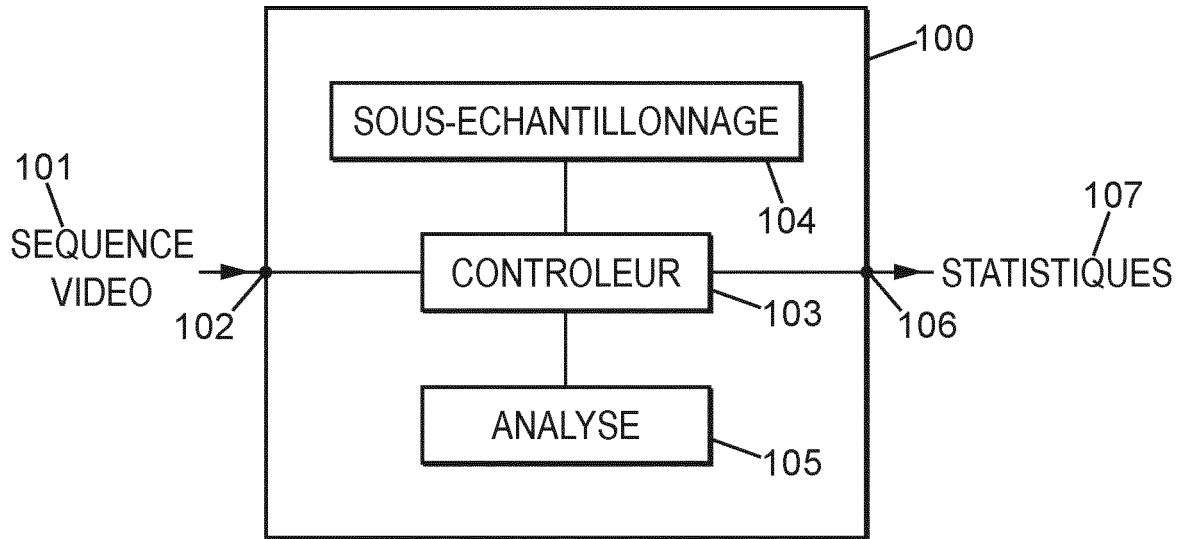


FIG. 1

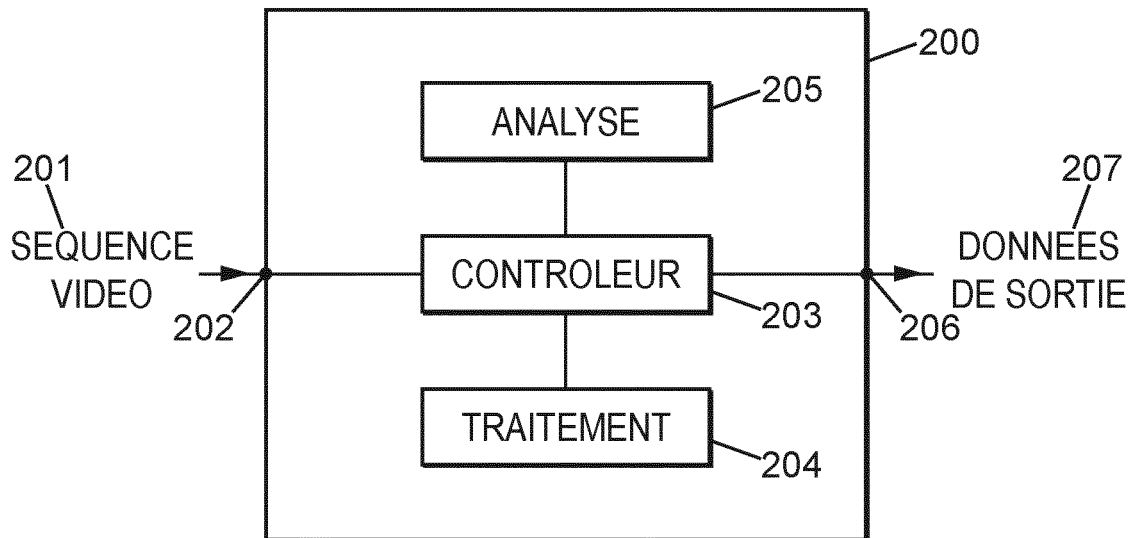


FIG. 2

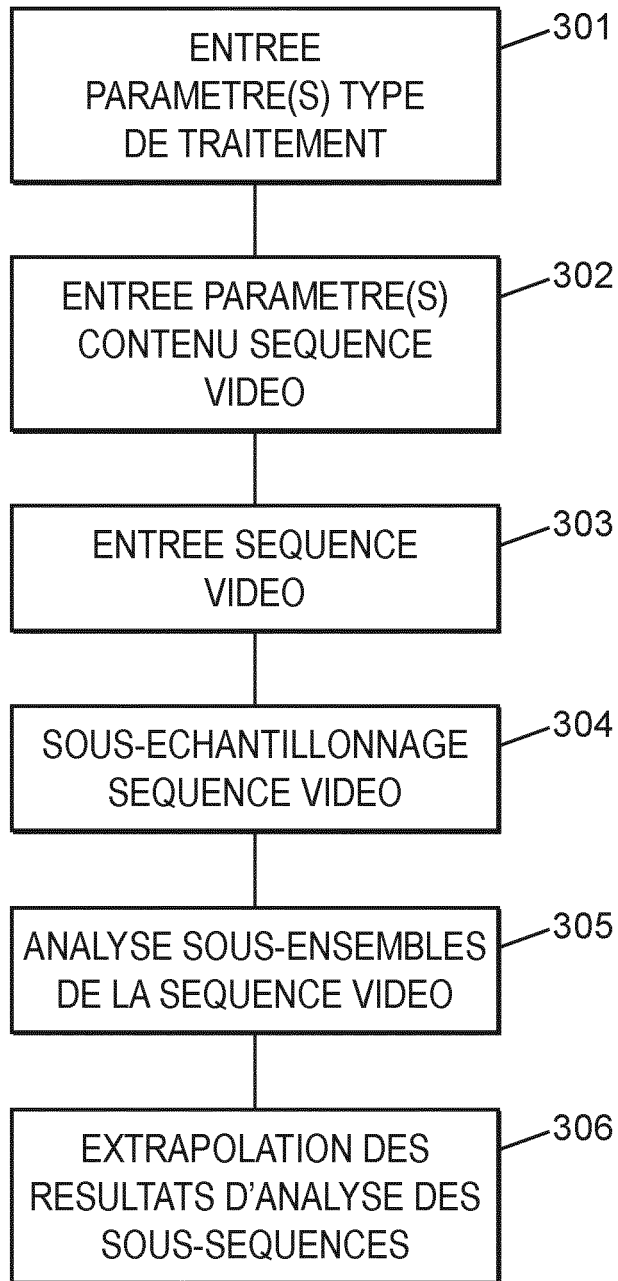


FIG. 3

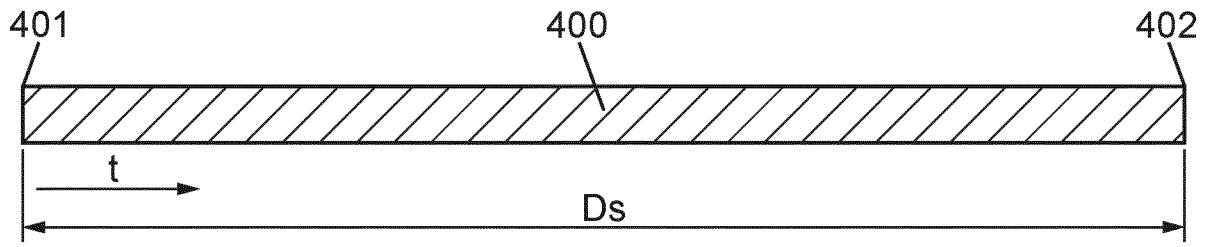


FIG. 4a

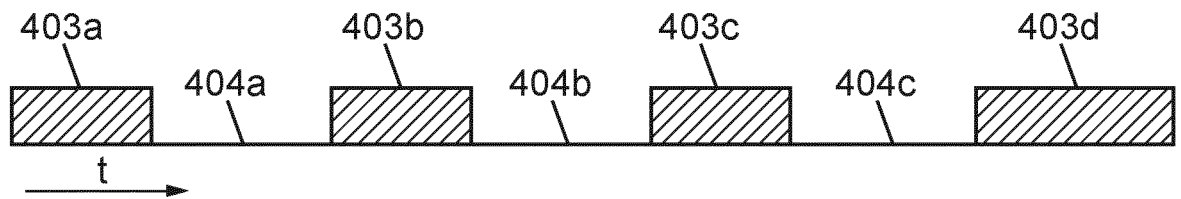


FIG. 4b

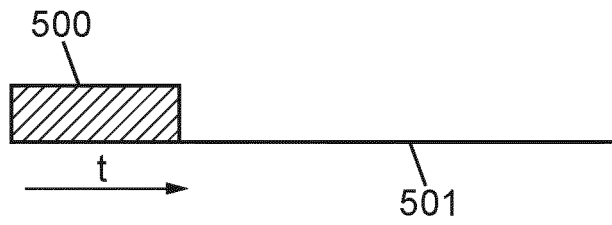


FIG. 5a

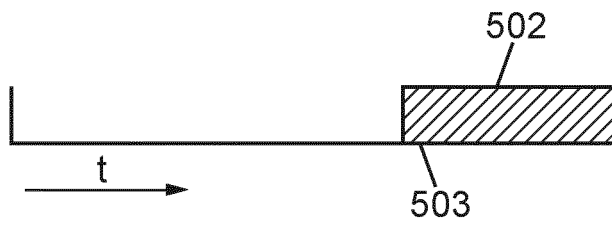


FIG. 5b

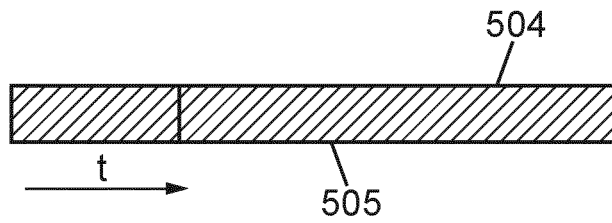


FIG. 5c

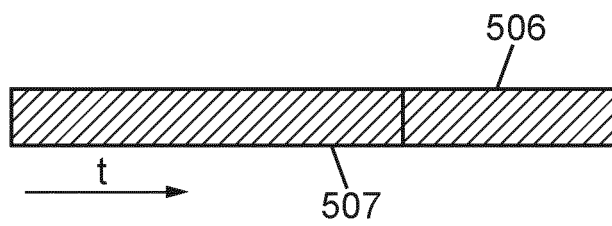


FIG. 5d

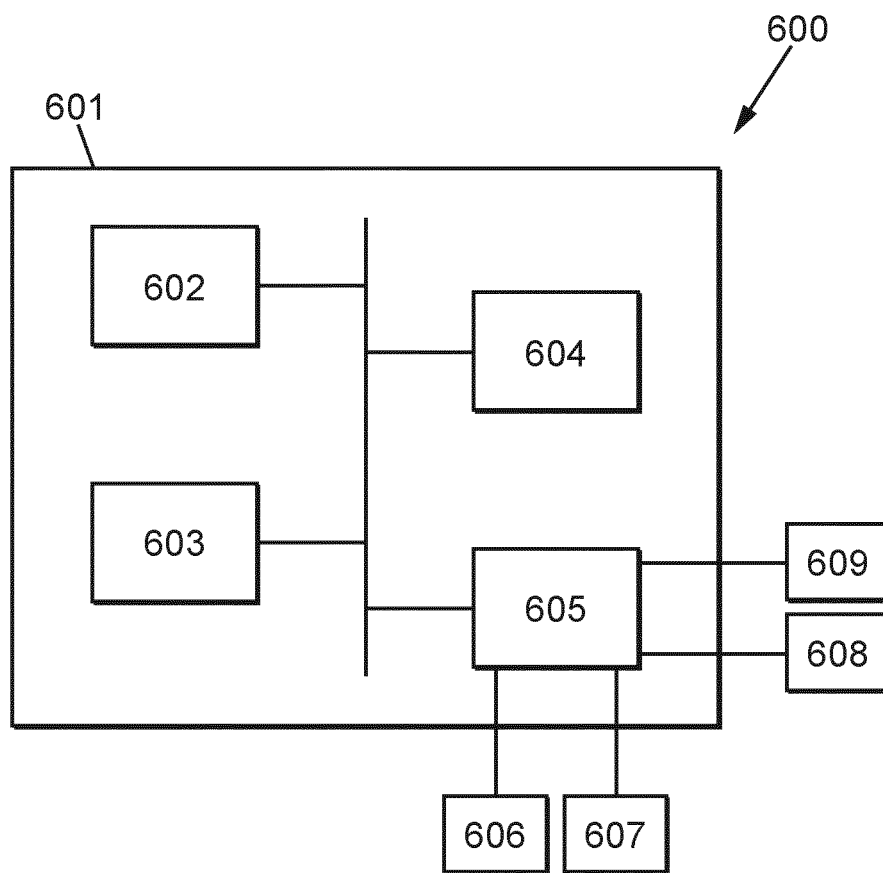


FIG. 6

RAPPORT DE RECHERCHE

articles L.612-14, L.612-53 à 69 du code de la propriété intellectuelle

OBJET DU RAPPORT DE RECHERCHE

L'I.N.P.I. annexe à chaque brevet un "RAPPORT DE RECHERCHE" citant les éléments de l'état de la technique qui peuvent être pris en considération pour apprécier la brevetabilité de l'invention, au sens des articles L. 611-11 (nouveau) et L. 611-14 (activité inventive) du code de la propriété intellectuelle. Ce rapport porte sur les revendications du brevet qui définissent l'objet de l'invention et délimitent l'étendue de la protection.

Après délivrance, l'I.N.P.I. peut, à la requête de toute personne intéressée, formuler un "AVIS DOCUMENTAIRE" sur la base des documents cités dans ce rapport de recherche et de tout autre document que le requérant souhaite voir prendre en considération.

CONDITIONS D'ETABLISSEMENT DU PRESENT RAPPORT DE RECHERCHE

Le demandeur a présenté des observations en réponse au rapport de recherche préliminaire.

Le demandeur a maintenu les revendications.

Le demandeur a modifié les revendications.

Le demandeur a modifié la description pour en éliminer les éléments qui n'étaient plus en concordance avec les nouvelles revendications.

Les tiers ont présenté des observations après publication du rapport de recherche préliminaire.

Un rapport de recherche préliminaire complémentaire a été établi.

DOCUMENTS CITES DANS LE PRESENT RAPPORT DE RECHERCHE

La répartition des documents entre les rubriques 1, 2 et 3 tient compte, le cas échéant, des revendications déposées en dernier lieu et/ou des observations présentées.

Les documents énumérés à la rubrique 1 ci-après sont susceptibles d'être pris en considération pour apprécier la brevetabilité de l'invention.

Les documents énumérés à la rubrique 2 ci-après illustrent l'arrière-plan technologique général.

Les documents énumérés à la rubrique 3 ci-après ont été cités en cours de procédure, mais leur pertinence dépend de la validité des priorités revendiquées.

Aucun document n'a été cité en cours de procédure.

**1. ELEMENTS DE L'ETAT DE LA TECHNIQUE SUSCEPTIBLES D'ETRE PRIS EN
CONSIDERATION POUR APPRECIER LA BREVETABILITE DE L'INVENTION**

CHONG-WAH NGO ET AL: "Video Partitioning by Temporal Slice Coherency", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE SERVICE CENTER, PISCATAWAY, NJ, US, vol. 11, no. 8, 1 août 2001 (2001-08-01), XP011014230, ISSN: 1051-8215

US 2003/004966 A1 (BOLLE RUDOLF M [US] ET AL) 2 janvier 2003 (2003-01-02)

US 2003/033347 A1 (BOLLE RUDOLF M [US] ET AL) 13 février 2003 (2003-02-13)

BERTINI M ET AL: "Content-based indexing and retrieval of TV news", PATTERN RECOGNITION LETTERS, ELSEVIER, AMSTERDAM, NL, vol. 22, no. 5, 1 avril 2001 (2001-04-01), pages 503-516, XP004233002, ISSN: 0167-8655, DOI: 10.1016/S0167-8655(00)00113-6

**2. ELEMENTS DE L'ETAT DE LA TECHNIQUE ILLUSTRANT L'ARRIERE-PLAN
TECHNOLOGIQUE GENERAL**

NEANT

**3. ELEMENTS DE L'ETAT DE LA TECHNIQUE DONT LA PERTINENCE DEPEND
DE LA VALIDITE DES PRIORITES**

NEANT