

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
7 septembre 2007 (07.09.2007)

PCT

(10) Numéro de publication internationale
WO 2007/099050 A1

(51) Classification internationale des brevets :
H04N 5/76 (2006.01) *G11B 27/00* (2006.01)
H04N 5/445 (2006.01)

(21) Numéro de la demande internationale :
PCT/EP2007/051650

(22) Date de dépôt international :
21 février 2007 (21.02.2007)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
06/50747 3 mars 2006 (03.03.2006) FR

(71) Déposant (pour tous les États désignés sauf US) :
THOMSON LICENSING [FR/FR]; 46 Quai A. Le
Gallo, F-92100 Boulogne-Billancourt (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : DE-
MARTY, Claire-Hélène [FR/FR]; 8 rue du Grand

Chevreuil, F-35520 La Mezière (FR). CHEVALLIER,
Louis [FR/FR]; 6, rue des Mimosas, F-35520 La Mezière
(FR). VIGOUROUX, Jean-Ronan [FR/FR]; 11A, rue
d'Echange, F-35000 Rennes (FR).

(74) Mandataires : LE DANTEC, Claude etc.; Thomson, 46,
Quai Alphonse Le Gallo, F-92100 Boulogne-Billancourt
(FR).

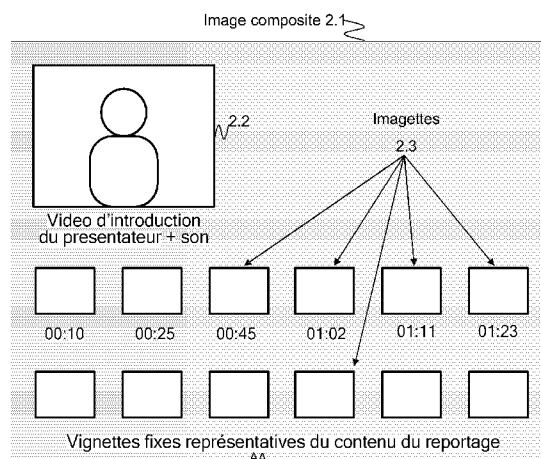
(81) États désignés (sauf indication contraire, pour tout titre de
protection nationale disponible) : AE, AG, AL, AM, AT,
AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO,
CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB,
GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP,
KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT,
LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU,
SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR,
TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) États désignés (sauf indication contraire, pour tout titre
de protection régionale disponible) : ARIPO (BW, GH,

[Suite sur la page suivante]

(54) Title: METHOD FOR DISPLAYING DATA EXTRACTED FROM A DOCUMENT CONSISTING OF REPORTS AND
RECEIVER IMPLEMENTING SAID METHOD

(54) Titre : PROCEDE D’AFFICHAGE D’INFORMATIONS EXTRAITES D’UN DOCUMENT COMPOSE DE REPORTAGES
ET RECEPTEUR METTANT EN ŒUVRE LE PROCEDE



2.1... composite image
2.2... video introducing the announcer and sound
AA... thumbnail images representing the content of the report

(57) Abstract: The invention concerns a method for displaying images extracted from an audio-visual document recorded in an apparatus. The recorded document comprises at least one report including at the beginning a first video sequence having a particular characteristic, such as the image of an announcer. Said first video sequence comprising only a limited informative content, the invention consists in selecting at least one extracted image of the report but not of the first video sequence, and in creating and displaying a composite image including both the first video sequence with the announcer and at least one image extracted from the report. An improvement consists in that said first video sequence is derived from a summary of the document, images of the report are added to the images of the summary to increase the informative content thereof. The invention also concerns a receiver for implementing said method.

[Suite sur la page suivante]

WO 2007/099050 A1



GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

Publiée :

— avec rapport de recherche internationale

(57) Abrégé : L'invention concerne un procédé d'affichage d'images extraites d'un document audiovisuel enregistré dans un appareil. Le document enregistré comporte au moins un reportage comprenant au début une première séquence vidéo présentant une caractéristique particulière, tel que l'image d'un présentateur. Cette première séquence vidéo ne comportant qu'un contenu informatif limité, l'invention consiste à sélectionner au moins une image extraite du reportage mais pas de la première séquence vidéo, et à créer et afficher une image composite comportant à la fois la première séquence vidéo avec le présentateur et une au moins image extraite du reportage. Un perfectionnement consiste en ce que cette première séquence vidéo provient d'un sommaire du document, des images du reportage sont rajoutées aux images du sommaire pour en augmenter le contenu informatif. L'invention concerne également le récepteur apte à mettre en œuvre le procédé.

PROCEDE D'AFFICHAGE D'INFORMATIONS EXTRAITES D'UN DOCUMENT COMPOSE DE REPORTAGES ET RECEPTEUR METTANT EN OEUVRE LE PROCEDE

5 **DOMAINE TECHNIQUE DE L'INVENTION**

L'invention concerne un procédé d'affichage d'informations extraites d'un document enregistré composé de reportages, et un récepteur mettant en œuvre le procédé.

10 **ETAT ANTERIEUR DE L'INVENTION**

De nos jours, de nombreux appareils permettent la réception et la reproduction d'œuvres audiovisuelles enregistrées. Récemment, des disques durs de grande capacité ont été intégrés à ces récepteurs audiovisuels ; on peut y stocker des contenus téléchargés soit d'un réseau de diffusion unidirectionnel, soit d'un réseau point à point (Internet par exemple). Parmi ces appareils, les PVR (acronyme de Personal Video Recorder) sont généralement dotés d'un Guide Electronique de Programme (EPG en abrégé) qui permet à un utilisateur d'afficher les contenus enregistrés dans le disque dur et de les reproduire à l'aide d'une commande.

15 Aujourd'hui, la taille des disques durs permet de stocker plus de 200 heures de programmes. Cette quantité d'information disponible nécessite l'ajout de nouvelles fonctionnalités dans les PVR permettant une navigation inter et intra rapide. L'EPG permet de regrouper les documents selon de nombreux critères et de les présenter par groupes, par exemple : le thème, le type, la

20 langue, le contrôle parental, etc. La classification s'effectue à l'aide de métadonnées associées au contenu audiovisuel. Selon une variante, le récepteur analyse le contenu et est capable de le classifier. Un autre exemple de moyen de stockage concerne les magnétoscopes qui enregistrent les œuvres audiovisuelles sur des bandes magnétiques. Dans

25 ce cas, l'accès au contenu audiovisuel est séquentiel. Sur ces produits, il est possible, soit d'enregistrer immédiatement la vidéo en cours de réception

30

sur le canal courant, soit de programmer l'enregistrement d'un événement diffusé ultérieurement. Quelle que soit la façon de mémoriser le contenu, l'interface utilisateur permet de contrôler la sélection et la reproduction des documents enregistrés. Dans un premier temps, l'utilisateur sélectionne à l'aide d'un guide de programmes l'émission qu'il désire visualiser. Puis, en appuyant sur la touche « lecture » il déclenche la visualisation de l'œuvre.

Dans le cas d'enregistrements numériques (par exemple : un DVD), l'œuvre est segmentée en plusieurs chapitres. De cette façon, l'utilisateur peut choisir un chapitre et commencer la visualisation au début du chapitre sélectionné. En appuyant sur la touche « Plage suivante » la visualisation saute au segment suivant. Certaines émissions reçues par un réseau de diffusion sont également segmentées, de sorte que, une fois enregistrées dans le disque dur du récepteur, l'utilisateur peut sélectionner un segment précis pour le visualiser ou naviguer d'un segment à un autre en ne visualisant que les premières secondes pour rechercher le segment désiré. Si l'émission est reçue non segmentée, il existe des procédés de segmentation permettant de le faire au niveau du récepteur. La demande de brevet référencée FR02.13977 déposée en France le 28 octobre 2002 décrit un procédé de segmentation qui exploite le contenu audiovisuel de l'œuvre pour en extraire des parties ayant une unité sémantique. Le procédé décrit dans cette demande incluse par référence à la présente demande, permet d'élaborer une segmentation de vidéos de sport selon plusieurs niveaux. Il existe également des procédés permettant d'extraire de manière automatique la structure arborescente d'un film ou d'une vidéo documentaire. Classiquement, la vidéo est découpée en plans en se basant sur des changements abrupts du contenu colorimétrique des images. Puis, les plans sont regroupés en scènes en se basant sur la ressemblance visuelle des contenus des plans et leur proximité temporelle.

Lorsque le document enregistré est un journal télévisé, l'indexation selon les procédés actuels va permettre d'accéder directement aux

reportages. Lorsqu'un utilisateur sélectionne un reportage, celui-ci est extrait de la mémoire de son appareil et envoyé à l'écran. Les reportages débutent visuellement par un gros plan sur le présentateur du journal. Si le message sonore qui l'accompagne est important car il introduit le reportage, l'image
5 ne donne aucune information pertinente et ne permet pas à l'utilisateur de distinguer un reportage d'un autre, ni d'en avoir une vue d'ensemble, de sorte que l'utilisateur a connaissance du reportage que par sa perception auditive. Une façon de raccourcir ce temps consisterait à couper la séquence introductive du reportage où le présentateur est seul à l'écran et
10 de débiter directement par les images du reportage. Mais cette technique a deux inconvénients : le message sonore introductif et informatif est perdu, et aucune représentation complète du reportage n'est finalement affichée.

Le document US 2004/0197088 publié le 7 octobre 2004 décrit un système pour découper un document en une pluralité de parties, par
15 exemple correspondant à la structure d'un match de football. La structure du document est présentée à l'écran avec des imageries (en Anglais : « thuringiens ») associées aux parties du document. Des documents, par exemple des séquences publicitaires, sont téléchargés de l'extérieur de l'appareil et visualisés à la suite du contenu vidéo et non au même moment,
20 ce qui rallonge le temps de visualisation du document complet. L'utilisateur peut sélectionner les séquences téléchargées pour par exemple commander les objets visualisés. De ce fait, ce système nécessite une communication extérieure pour télécharger les séquences audiovisuelles qui sont rajoutées au document.

25 Le document US 6 404 925 publié le 11 juin 2002 décrit une méthode pour segmenter des documents audiovisuels. Ces segments servent d'index pour accéder à des documents enregistrés. L'indexation s'effectue en analysant le signal audio et en identifiant la personne qui parle. On peut ainsi définir des groupes de segments audiovisuels présentant la même
30 personne qui parle.

RESUME DE L'INVENTION

La présente invention propose une interface originale de représentation d'un reportage qui conserve la structure et les éléments du reportage et qui permet d'informer plus rapidement l'utilisateur de l'objet du
5 reportage.

L'invention a pour objet un procédé de création d'un nouveau document audiovisuel à partir d'un document audiovisuel enregistré dans un appareil, le document enregistré comportant une séquence vidéo présentant une caractéristique particulière, tel que l'image d'un présentateur ;
10 caractérisé en ce qu'il comporte les étapes suivantes :

- création d'une séquence vidéo composite composée des images extraites de la séquence vidéo enregistrée combinées à au moins une image extraite du document à un moment n'appartenant pas à la séquence vidéo enregistré.

De cette façon, le message sonore délivré par le présentateur est toujours présent mais son image associée ne monopolise plus l'écran, car celui-ci est enrichi de parties visuelles du reportage. L'enrichissement du contenu visuel de la séquence vidéo enregistré par l'insertion d'images
15 extraites du document à un moment n'appartenant pas à cette séquence vidéo permet de ne pas rallonger la durée du document.

Selon un premier perfectionnement, les reportages commencent par un sommaire comportant plusieurs premières séquences vidéo affichant le présentateur. Le procédé consiste alors à découper le document enregistré en plusieurs reportages, à découper le sommaire en parties annonçant un
25 reportage, chaque partie de sommaire étant ainsi associé à un reportage et enfin, à afficher la séquence vidéo composite en affichant en même temps une partie du sommaire et une image précédemment sélectionnée dans le reportage associé à cette partie de sommaire. De cette façon, l'utilisateur peut visualiser le sommaire d'un journal en ayant à l'écran aussi des images
30 clef des reportages annoncés dans ce sommaire. Selon un

perfectionnement, l'utilisateur peut naviguer dans le sommaire et notamment déclencher la visualisation plein écran du reportage associé

Selon un autre perfectionnement, le document comportant plusieurs reportages, l'utilisateur peut interrompre l'affichage de la séquence vidéo composite ce qui déclenche une nouvelle étape d'affichage de la séquence vidéo composite correspondant au reportage suivant. De cette façon, l'utilisateur peut rapidement atteindre un reportage qui l'intéresse.

Selon un autre perfectionnement, l'utilisateur peut sélectionner une image sélectionnée comme étant une image clef du reportage, ce qui déclenche la visualisation du plan séquence contenant cette image clef. De cette façon, l'utilisateur peut rapidement atteindre une partie du reportage dont l'image clef a retenu son attention.

Selon un autre perfectionnement, l'affichage d'une séquence vidéo composite dure une durée déterminée au début de chaque première séquence vidéo. Selon un autre perfectionnement, l'étape de sélection d'une image au moins du reportage consiste à sélectionner une image médiane dans un plan séquence. Selon un autre perfectionnement, l'étape d'affichage d'une séquence vidéo composite consiste également à afficher des informations textuelles diffusées dans des informations de services. Selon un autre perfectionnement, l'étape d'affichage d'une séquence vidéo composite consiste également à afficher en association avec chaque image sélectionnée une indication horaire montrant le moment d'apparition de cette image au sein du reportage.

L'invention a également pour objet un récepteur audiovisuel comprenant une unité centrale, un moyen de reproduction d'au moins une partie d'un document audiovisuel enregistré dans un moyen de stockage, le document comportant au moins une première séquence vidéo présentant une caractéristique particulière, tel que l'image d'un présentateur ; caractérisé en ce qu'il comporte un moyen de création d'une séquence vidéo composite composée des images extraites de la séquence vidéo enregistrée

combinées à au moins une image extraite du document à un moment n'appartenant pas à la séquence vidéo enregistrée.

DESCRIPTION DES FIGURES

5 D'autres caractéristiques et avantages de l'invention apparaîtront maintenant avec plus de détails dans le cadre de la description qui suit d'exemples de réalisation donnés à titre illustratif en se référant aux figures annexées qui représentent :

10 - la figure 1 est un diagramme bloc d'un récepteur audiovisuel pour la mise en œuvre de l'invention,

- la figure 2 est une apparence d'écran montrant le début d'un reportage selon un exemple préféré de réalisation,

- la figure 3 est un organigramme des principales étapes effectuées par le navigateur selon un exemple préféré de réalisation,

15 - la figure 4 est un organigramme des principales étapes effectuées par le navigateur selon une variante de réalisation,

- la figure 5 est une illustration du découpage d'un document enregistré.

DESCRIPTION D'UN MODE DE REALISATION DE L'INVENTION

20 On décrira tout d'abord le fonctionnement d'un récepteur audiovisuel 1 muni d'un dispositif d'affichage 2. Le récepteur comprend une unité centrale 3 reliée à une mémoire de programme 12, et une interface 5 pour la communication avec un bus numérique local à haut débit 6 permettant de
25 transmettre des données audio/vidéo en temps réel. Ce réseau est par exemple un réseau IEEE 1394. Le récepteur peut également recevoir des données audio/vidéo d'un réseau de diffusion à travers une antenne de réception associée à un démodulateur 4. Le récepteur comprend en outre un récepteur de signaux infrarouges 7 pour recevoir les signaux d'une
30 télécommande 8, une mémoire 9 pour le stockage des données téléchargées, et une logique de décodage audio/vidéo 10 pour la génération

des signaux audiovisuels envoyés à l'écran de télévision 2. Selon un exemple de réalisation, la mémoire 9 est un disque dur de grande capacité permettant d'enregistrer des documents audiovisuels. Selon un autre exemple de réalisation, le document audiovisuel est contenu dans un support amovible tel qu'un « compact disk » ou un disque DVD , le support étant inséré dans un périphérique lecteur (non représenté sur la figure 1) incorporé dans le récepteur. La télécommande est dotée des touches de direction : ↑, ↓, → et ← et des touches de fonction ou de validation telles que : « Sommaire », « OK » et « STOP ». Selon un mode préféré de réalisation, les flèches de direction permettent de déplacer un curseur à l'écran, les commandes accessibles à l'utilisateur sont présentées sous forme d'icônes et sont activées lorsque l'utilisateur appuie sur « OK » avec le curseur positionné sur l'icône correspondant à la commande souhaitée.

Le récepteur comprend également un circuit 11 d'affichage de données sur l'écran, appelé souvent circuit OSD, de l'Anglais "On Screen Display" (signifiant littéralement "affichage sur l'écran"). Le circuit OSD 11 est un générateur de texte et de graphisme qui permet d'afficher à l'écran des menus, des pictogrammes (par exemple, un numéro correspondant à la chaîne visualisée) et qui permet d'afficher les menus de navigation conformément à la présente invention. Le circuit OSD est contrôlé par l'Unité Centrale 3 et un navigateur 12. Le navigateur 12 est avantageusement réalisé sous la forme d'un module de programme enregistré dans une mémoire morte. Il peut aussi être réalisé sous la forme d'un circuit spécialisé de type ASIC par exemple.

Le bus numérique 6 et/ou le réseau de diffusion transmettent au récepteur des données comprenant des documents audiovisuels. Eventuellement, ces documents sont associés à des données descriptives. Ces données proviennent soit d'un réseau de diffusion, soit du réseau numérique 6. Les données descriptives comprennent notamment des informations de segmentation définissant les différents segments du document. Avantageusement, les données descriptives permettent une

décomposition fine du document et comportent du texte permettant de donner au moins un titre à chaque segment. Les données descriptives sont par exemple contenues dans les informations de service spécifiées par la norme DVB-SI. Elles sont enregistrées dans la mémoire 9 du récepteur.

5

Après avoir décrit les éléments constituant l'invention selon un exemple préféré de réalisation, nous allons maintenant expliquer comment ceux-ci coopèrent, avec l'aide de la figure 3.

10

Dans un premier temps, l'utilisateur enregistre le journal télévisé dans la mémoire 9 de son récepteur 1 (étape 3.1). L'utilisateur lance ensuite l'exécution de son navigateur résident de son récepteur et demande la visualisation du document enregistré, ce document étant constitué d'au moins un reportage. Classiquement dans un journal télévisé, un reportage commence par une vidéo montrant le présentateur introduisant oralement le sujet qui va être traité. Au niveau visuel, aucune information pertinente n'est fournie. Au niveau sonore au contraire les paroles du présentateur sont essentielles à l'utilisateur pour déterminer le sujet du reportage. A l'étape 3.2, le navigateur sélectionne une ou plusieurs images extraites du corps du reportage. Ces images ne montrant pas le présentateur, leurs contenus visuels sont riches en enseignement. Selon la présente invention, le récepteur 1 va élaborer une nouvelle image 2.1 en redimensionnant l'image du présentateur 2.2 qui apparaît au début du reportage afin de libérer de la place sur l'écran pour afficher d'autres images 2.3, fixes celle-là, extraites à d'autres moments du reportage (étape 3.3). Un exemple de cette nouvelle image 2.1 dite « composite » est représenté par la figure 2. On aperçoit en haut à gauche de l'écran une réduction 2.2 de la vidéo avec le présentateur présent. En partie basse de l'écran, des imageries 2.3 (« thumbnail » en Anglais) automatiquement extraites du reportage et représentatives de son contenu visuel sont affichées. Dans l'exemple de la figure 2, douze imageries 2.3 sont affichées qui correspondent à douze moments

15

20

25

30

particuliers du reportage. D'autres dispositions sont également possibles, par exemple le présentateur apparaît au milieu de l'écran sans changement de taille, et les imajettes sont disposées en incrustation tout autour de lui. Il est préférable que la disposition des imajettes respecte l'ordre chronologique de leur apparition dans le reportage. Au besoin, elles
5 apparaissent avec un numéro ou en affichant le moment d'apparition à compter du début du reportage ou du document enregistré. Les moments d'apparition en minutes et secondes des images affichées sur la première ligne sont affichés dans l'exemple illustré par la figure 2. L'utilisateur a donc
10 à sa disposition des informations à la fois sonores et visuelles résumant le contenu du reportage et lui permettant de le guider dans son choix de visionnage du reportage complet.

Lorsque le présentateur n'apparaît plus à l'écran, il devient inutile d'afficher l'image composite 2.1 et la visualisation du reportage continue de
15 façon classique, sans traitement d'image au niveau du récepteur. Si l'utilisateur au vu de l'image composite 2.1 ne désire pas visualiser ce reportage, il appuie sur la touche →, ce qui termine la visualisation du reportage. Le navigateur affiche alors le reportage suivant, l'écran prenant la même apparence que celle présentée à la figure 2.

20 Une première façon d'extraire des imajettes consiste à analyser les reportages pour en déterminer les différents plan séquences. Un plan séquence (ou « shot » en Anglais) est une séquence ininterrompue tournée par une caméra ou un caméscope en une seule prise de vues. Les imajettes affichées sont les premières images des plans séquences du
25 reportages. D'autres techniques sont utilisables. Par exemple, le navigateur peut choisir de prendre l'image médiane dans le plan séquence, c'est-à-dire l'image qui est au milieu du plan. Une autre technique consiste à extraire les images correspondant à un minimum d'activité. La mesure d'activité est une technique connue de l'Homme du métier, elle est généralement basée sur le
30 mouvement dans les images. Il est également possible de sélectionner les imajettes contenant des visages en gros plans. D'autres techniques

reposant sur l'analyse sémantique du contenu vidéo peuvent être également utilisées.

Selon un mode préféré de réalisation, les imasettes sont une photo inanimée extraite du reportage. Ce mode est facile à produire avec des
5 moyens simples d'incrustation. En variante, une imasette est un plan séquence sélectionné dans le reportage. De ce fait, au début du reportage et pendant que le présentateur introduit le sujet, les plans séquences apparaissent les uns après les autres donnant ainsi une impression de mouvement. Le plan séquence disparaît de l'écran lorsqu'il se termine,
10 Quelques secondes plus tard, un autre plan séquence apparaît à un autre endroit ou à un autre. L'apparition et la disparition des images animées ajoutent un côté ludique à l'apparence d'écran. Des tests ont montré que trois ou quatre visualisations de plans séquences offrent un bon compromis entre l'occupation de l'écran et la quantité d'information visuelle offerte à
15 l'utilisateur.

La première partie d'un journal télévisé est généralement constituée d'un sommaire au cours duquel le présentateur présente les principaux reportages qui vont être développés dans le reste du journal. Selon une
20 variante, le récepteur 1 modifie l'apparence des séquences vidéo du sommaire en rajoutant des imasettes afin que l'utilisateur perçoive plus d'informations sur les sujets que les informations données uniquement par le présentateur. Selon cette variante, un reportage est découpé en deux : la séquence vidéo du sommaire où le présentateur introduit le sujet du
25 reportage, et le corps du reportage qui apparaît ensuite au cours du journal, et d'où sont extraits les imasettes.

La figure 4 présente un organigramme des étapes d'analyse d'un document vidéo enregistré permettant de générer une liste des liens entre le sommaire et les reportages et d'afficher les imasettes, pour la mise en
30 œuvre de perfectionnement.

A l'étape 4.1, le document vidéo est enregistré dans le récepteur audiovisuel 1. Il est par exemple reçu d'un réseau de diffusion. A l'étape 4.2, le document vidéo est segmenté en reportages. De nombreuses méthodes de la littérature proposent des solutions à ce problème. Par exemple, une technique possible consiste, suite à un découpage en plans de la vidéo à repérer les plans récurrents montrant une image déterminée, typiquement le présentateur du journal télévisé. Les différents plans présents entre les deux apparitions du présentateur sont regroupés au sein d'un même segment vidéo caractérisant un reportage. Dans un autre exemple, la vidéo est découpée en plans, puis, les plans sont regroupés en scènes en se basant sur la ressemblance visuelle des contenus des plans et leur proximité temporelle. On peut observer la composante audio en analysant l'ambiance sonore (un reportage sur un match de football, sur un documentaire écologique ou sur la dernière intervention d'un homme politique présentent des caractéristiques différentes), ou la continuité de la voix d'un même journaliste commentant son reportage. Les différents plans présents entre les deux apparitions du présentateur sont regroupés au sein d'un même segment vidéo caractérisant un reportage. Nous obtenons donc en sortie de ce module une liste de reportages ainsi qu'une liste de plans contenant le présentateur.

A l'étape 4.3, le navigateur analyse le document pour déterminer le sommaire. Typiquement, le sommaire se situe au début, plus rarement à la fin du document. Le sommaire peut être un seul plan séquence montrant un présentateur de face. Eventuellement, une musique déterminée est détectable en fond sonore du commentaire. La recherche des images du sommaire peut s'opérer suivant différentes hypothèses, par exemple :

- les images du sommaire contiennent des vignettes rectangulaires généralement fixes en haut à gauche ou en bas à droite du présentateur,
- le sommaire est diffusé au début du journal avant l'apparition du présentateur sous forme d'une succession de clips vidéo.

D'autres techniques de reconnaissance visuelle existent. Par exemple, une sélection de l'ensemble des plans précédant le premier reportage est d'abord effectuée. Sur l'ensemble de ces plans, une détection de vignettes est mise en œuvre sur l'ensemble des images de chaque plan.

5 Cette détection peut s'appuyer sur le fait que les vignettes apparaissent dans une partie de l'image, le reste de l'image demeurant quasi fixe. On peut utiliser par exemple un algorithme basé sur l'utilisation des images résultant de la différence entre images successives associée à une détection de forme rectangulaire par morphologie mathématique.

10 Si le sommaire est diffusé en début de journal et que les images de reportages n'apparaissent pas en médaillon dans l'image mais en plein écran, il est possible d'utiliser une méthode de détection s'appuyant sur le fait que :

- le présentateur donne le résumé oral accompagnant le sommaire, ce qui peut permettre d'utiliser une segmentation en locuteur puis de sélectionner les images où le présentateur parle mais n'est pas visible ;

15

- le sommaire est généralement compris entre le générique et la première apparition du présentateur ; dans ce cas là, la détection du générique s'appuyant sur une détection des plages de musique ou sur une connaissance a priori du générique permet de retrouver le sommaire comme étant tout ce qui n'est pas générique.

20

Une fois le sommaire détecté, l'étape suivante 4.4 consiste à détecter les similitudes existant entre chaque portion du sommaire et chaque reportage. La méthode consiste ici à venir comparer les images du sommaire avec les images des reportages. Des techniques classiques d'extraction d'attributs caractéristiques de l'image associée à une mesure de distance permettent de fournir une mesure de similarité permettant de remonter à l'association des images sommaires et reportages. Dans de nombreux cas, on peut considérer que le sommaire est constitué de résumés vidéo des reportages associés, l'invention pourra avantageusement

25

30

utiliser la composante temporelle pour rendre le processus de mise en correspondance d'images plus robuste. Par exemple, l'ordre de présentation des reportages dans le sommaire est généralement le même que celui des reportages du journal. La mesure de similarité est ainsi renforcée si les deux
5 images présentant des similitudes se présentent dans le même ordre, les unes étant dans le sommaire, les autres dans les reportages.

Une fois cette étape réalisée, des couples formés d'une partie du sommaire et d'un reportage sont élaborés. Grâce à une interface utilisateur adaptée, ces associations vont maintenant être exploitées afin de donner
10 plus d'information visuelle aux images du sommaire.

La figure 5 illustre le découpage d'un document et la détermination des similitudes entre des parties de sommaire et des reportages sur un journal télévisé. A la suite de la segmentation, le programme de traitement a déterminé que le sommaire comporte cinq parties et que le journal comporte
15 cinq reportages. En analysant les similitudes entre les parties de sommaire et les reportages, le programme de traitement a pu établir quatre similitudes : la première partie du sommaire et le premier reportage, la seconde partie du sommaire et le second reportage, la troisième partie du sommaire et le quatrième reportage et la quatrième partie du sommaire avec
20 le cinquième reportage. Dans l'exemple illustré par la figure 5, le troisième reportage n'a pas été annoncé dans le sommaire, ce reportage ne peut donc pas être indexé par une partie du sommaire, il vu au moment de la visualisation intégrale du journal enregistré.

25 A l'étape 4.5, l'utilisateur a demandé de visualiser le document, le sommaire est affiché dans un premier temps. Le navigateur affiche une image composite 2.1 en insérant dans les images de parties du sommaire, des images extraites du reportage associé. Les images extraites du reportage associé sont incrustées dans une fenêtre d'écran en haut à droite
30 de l'écran. Les techniques d'incrustation étant bien connues de l'homme du métier, il n'est pas utile de les expliciter ici. De cette manière, l'utilisateur

perçoit en plus du contenu sonore produit par le présentateur un contenu visuel lui informant de l'objet du reportage associé. Le même exemple d'image que celui montré à la figure 2 est ainsi généré.

L'utilisateur perçoit des informations à la fois sonores et visuelles
5 résumant le contenu du reportage et lui permettant de la guider dans son choix de voir un reportage déterminé. Pendant la visualisation de cette partie du sommaire associé à ce reportage, le récepteur 1 émet le son du sommaire. Si l'utilisateur appuie sur la touche « OK » de la télécommande 8 au cours de la visualisation de cette partie de sommaire, le récepteur 1
10 extrait de la mémoire le reportage associé à cette partie et affiche en plein écran la vidéo de ce reportage en commençant par le début. A la fin de la vidéo, la visualisation revient automatiquement à la partie suivante du sommaire.

Selon un perfectionnement, l'utilisateur appuie sur la touche
15 sommaire, la visualisation de la vidéo s'arrête, et la première imagerie du reportage correspondante est mise graphiquement en évidence. L'utilisateur peut alors naviguer sur les imageries à l'aide des touches de direction. Après en avoir sélectionné une, en appuyant sur « OK », le récepteur 1 extrait de la mémoire la séquence du reportage qui commence par l'imagerie
20 sélectionnée et l'affiche à l'écran. De même que précédemment, à la fin du reportage, la visualisation revient automatiquement à la partie suivante du sommaire.

Selon un perfectionnement, les imageries de l'image composite 2.1 servent d'index pour accéder à la partie du reportage les contenant. Lorsque
25 l'image composite 2.1 est affichée, l'utilisateur appuie sur la touche « STOP » ce qui fige l'image de l'écran, l'image du présentateur étant alors mise graphiquement en évidence. En utilisant les touches de navigation, l'utilisateur navigue sur les imageries et en met une en évidence. Puis, en appuyant sur la touche « OK », l'utilisateur déclenche la visualisation du
30 plan séquence du reportage contenant l'image ainsi sélectionnée. Lorsque

la fin du plan séquence est atteinte, l'écran affiche de nouveau l'image composite 2.1 au moment où sa visualisation a été arrêtée.

Selon un perfectionnement, le document comporte en association des données audiovisuelles, des informations textuelles. Typiquement ces informations sont diffusées pour les personnes mal entendant et affichées en bas de l'écran sur requête de l'utilisateur. Ces informations textuelles contenues dans des métadonnées, sont associées à des reportages. Lorsque l'utilisateur affiche l'image composite 2.1 créée par le récepteur 1, une zone de l'écran affiche au moins une partie du texte contenu dans ces informations supplémentaires. En variante, le récepteur 1 dispose d'un moyen de reconnaissance vocale permettant d'analyser la parole du présentateur au cours du reportage et de la convertir en texte. Le texte transcrit à partir des signaux audio est affiché pendant la visualisation du sommaire.

Selon un autre perfectionnement, la visualisation d'une image composite 2.1 comme présentée par la figure 2 s'effectue au cours d'une durée déterminée. En effet, au début de la partie du sommaire correspondant à un reportage, le présentateur introduit le sujet mais à la fin de cette partie, l'utilisateur est bien au courant du reportage et n'a plus besoin d'être informé par les imajettes. De sorte que, ces imajettes peuvent être enlevées et l'image du présentateur peut apparaître seul en plein écran. Des tests ont montré qu'au bout de dix secondes, le sujet est bien introduit et les imajettes apportent peu d'informations en plus de ce qui a été dit.

Bien que la présente invention ait été décrite en référence aux modes de réalisation particuliers illustrés, celle-ci n'est nullement limitée par ces modes de réalisation, mais ne l'est que par les revendications annexées. On notera que des changements ou des modifications pourront être apportés par l'Homme du métier.

Revendications

1. Procédé de création d'un nouveau document audiovisuel à partir d'un document audiovisuel enregistré dans un appareil, le document enregistré comportant une séquence vidéo (2.2) présentant une caractéristique particulière, tel que l'image d'un présentateur ; caractérisé en ce qu'il comporte les étapes suivantes :

- création d'une séquence vidéo composite (2.1) composée des images extraites de la séquence vidéo enregistrée combinées à au moins une image extraite du document à un moment n'appartenant pas à la séquence vidéo enregistré.

2. Procédé de création selon la revendication 1, le document comportant un sommaire composé d'une pluralité de séquences vidéo (2.2) présentant ladite caractéristique particulière et une pluralité de reportages dont certain au moins est associé à une séquence vidéo (2.2) ; caractérisé en ce qu'il comporte au sein de l'appareil les étapes suivantes :

- découpage du document (4.2) enregistré en une pluralité de reportages,

- découpage du sommaire (4.3) en séquences vidéo et association de chaque séquence vidéo avec le reportage correspondant (4.4)

- création d'autant de séquences vidéo composite (2.1) que de séquences vidéo extraites du document enregistré, chaque séquence vidéo composite étant composée d'images extraites de la séquence vidéo enregistrée combinées à au moins une image extraite du reportage associé.

3. Procédé de création selon la revendication 2 ; caractérisé en ce qu'il comporte en outre une étape d'introduction d'une commande lors de l'affichage de la séquence vidéo composite déclenchant la visualisation plein écran du reportage associé.

4. Procédé de création selon l'une quelconque des revendications 2 et 3 ; caractérisé en ce qu'il comporte une étape d'introduction d'une commande d'interruption de l'affichage de la séquence vidéo composite, déclenchant une nouvelle étape d'affichage de la séquence vidéo composite
5 correspondant au reportage suivant.

5. Procédé de création selon l'une quelconque des revendications précédentes ; caractérisé en ce qu'il comporte une étape d'introduction d'une commande de sélection d'une image (2.3) extraite du document à un
10 moment n'appartenant pas à la séquence vidéo enregistré, déclenchant une étape de visualisation du plan séquence dont est extraite ladite image sélectionnée.

6. Procédé de création selon l'une quelconque des revendications
15 précédentes ; caractérisé en ce qu'il comporte une étape d'affichage d'une séquence vidéo composite (2.1) d'une durée déterminée au début de la séquence vidéo (2.2) présentant une caractéristique particulière.

7. Procédé de création selon l'une quelconque des revendications
20 précédentes ; caractérisé en ce que l'étape de création comporte le rajout d'informations textuelles diffusées dans des informations de services et associées au document.

8. Procédé de création selon l'une quelconque des revendications
25 précédentes ; caractérisé en ce que l'étape de création comporte le rajout d'une indication horaire associée à un image extraite du document à un moment n'appartenant pas à la séquence vidéo enregistré, l'indication horaire identifiant le moment d'apparition au sein du document de cette image.

9. Récepteur audiovisuel (1) comprenant une unité centrale (3), un moyen de reproduction (10) d'au moins une partie d'un document audiovisuel enregistré dans un moyen de stockage (9), le document comportant au moins une première séquence vidéo présentant une caractéristique particulière, tel que l'image d'un présentateur ; caractérisé en ce qu'il comporte un moyen de création (3, 12) d'une séquence vidéo composite (2.1) composée des images (2.2) extraites de la séquence vidéo enregistrée combinées à au moins une image (2.3) extraite du document à un moment n'appartenant pas à la séquence vidéo enregistrée.

10

10. Récepteur audiovisuel (1) selon la revendication 9, le document comportant un sommaire composé d'une pluralité de séquences vidéo (2.2) présentant ladite caractéristique particulière et une pluralité de reportages dont certain au moins est associé à une séquence vidéo (2.2) ; caractérisé en ce qu'il comporte :

15

- un moyen de découpage (3, 12) en une pluralité de reportages du document enregistré dans le moyen de stockage (9),

20

- un moyen de découpage du sommaire (4.3) en séquences vidéo et un moyen d'association de chaque séquence vidéo avec le reportage correspondant (4.4), le moyen de création créant autant de séquences vidéo composite (2.1) que de séquences vidéo extraites du document enregistré, chaque séquence vidéo composite étant composée d'images extraites de la séquence vidéo enregistrée combinées à au moins une image extraite du reportage associé.

25

11. Récepteur audiovisuel (1) selon la revendication 10 ; caractérisé en ce qu'il comporte un moyen d'introduction (7,8) d'une commande déclenchant lors de l'affichage de la séquence vidéo composite la visualisation plein écran du reportage associé.

30

12. Récepteur audiovisuel (1) selon l'une quelconque des revendications 10 et 11 ; caractérisé en ce qu'il comporte un moyen d'introduction (7,8) d'une commande interrompant l'affichage de la séquence vidéo composite, et déclenchant l'affichage de la séquence vidéo composite
5 correspondant au reportage suivant.

13. Récepteur audiovisuel (1) selon l'une quelconque des revendications 9 à 12 ; caractérisé en ce qu'il comporte un moyen d'introduction (7,8) d'une commande de sélection d'une image (2.3) extraite
10 du document à un moment n'appartenant pas à la séquence vidéo enregistré, déclenchant l'affichage du plan séquence dont est extraite ladite image sélectionnée.

14. Récepteur audiovisuel (1) selon l'une quelconque des revendications 9 à 13 ; caractérisé en ce que le moyen de reproduction (10) affiche la séquence vidéo composite (2.1) au cours d'une durée déterminée au début de la séquence vidéo (2.2) présentant une caractéristique particulière.

15. Récepteur audiovisuel (1) selon l'une quelconque des revendications 9 à 14 ; caractérisé en ce que le moyen de création (3,12) rajoute à la séquence vidéo composite des informations textuelles diffusées dans des informations de services et associées au document.

16. Récepteur audiovisuel (1) selon l'une quelconque des revendications 9 à 15 ; caractérisé en ce que le moyen de création (3,12) rajoute à la séquence composite une indication horaire associée à une image extraite du document à un moment n'appartenant pas à la séquence vidéo enregistré, l'indication horaire identifiant le moment d'apparition au
30 sein du document de cette image.

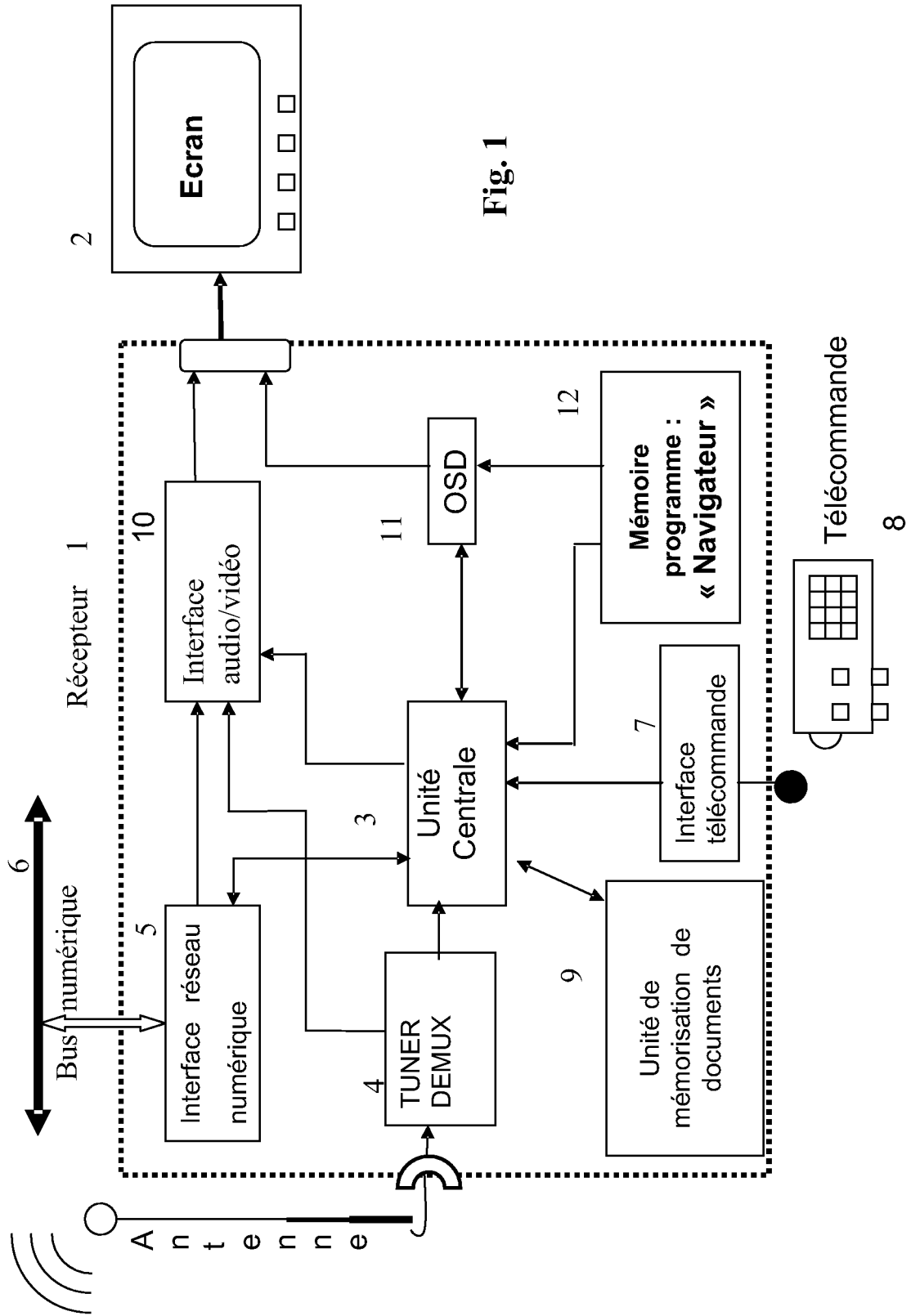


Fig. 1

Image composite 2.1

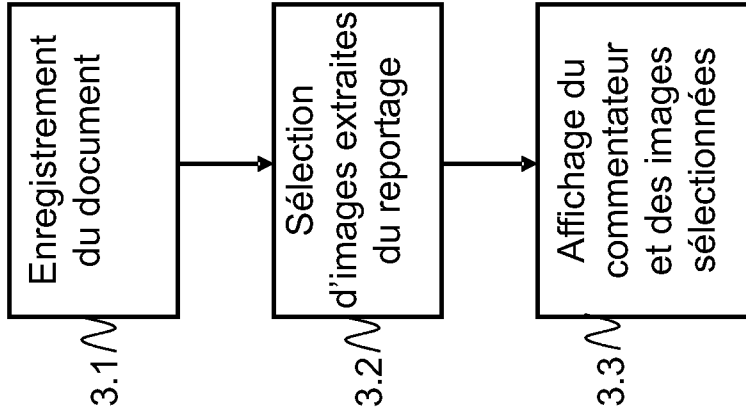
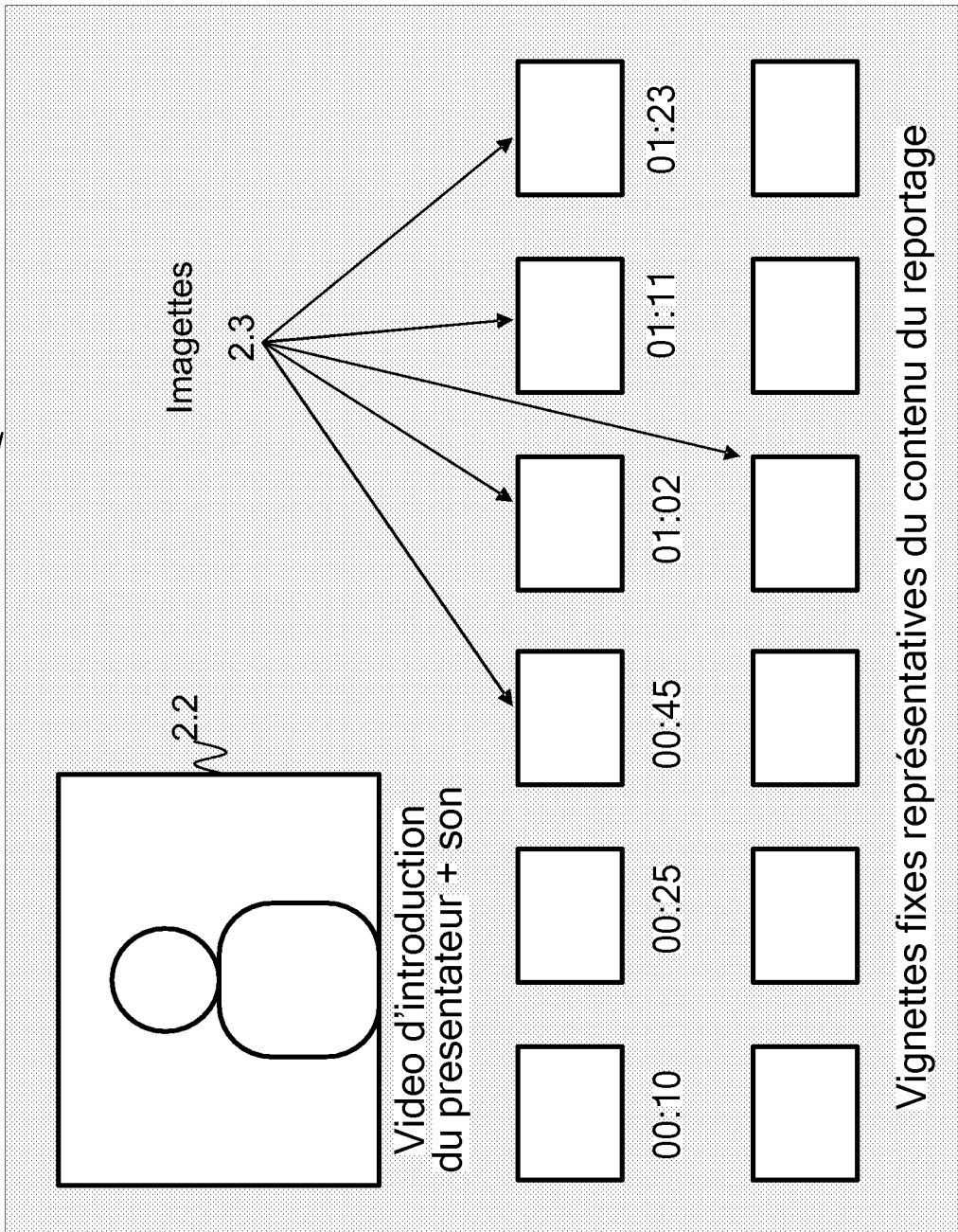


Fig. 3

Fig. 2

Fig. 4

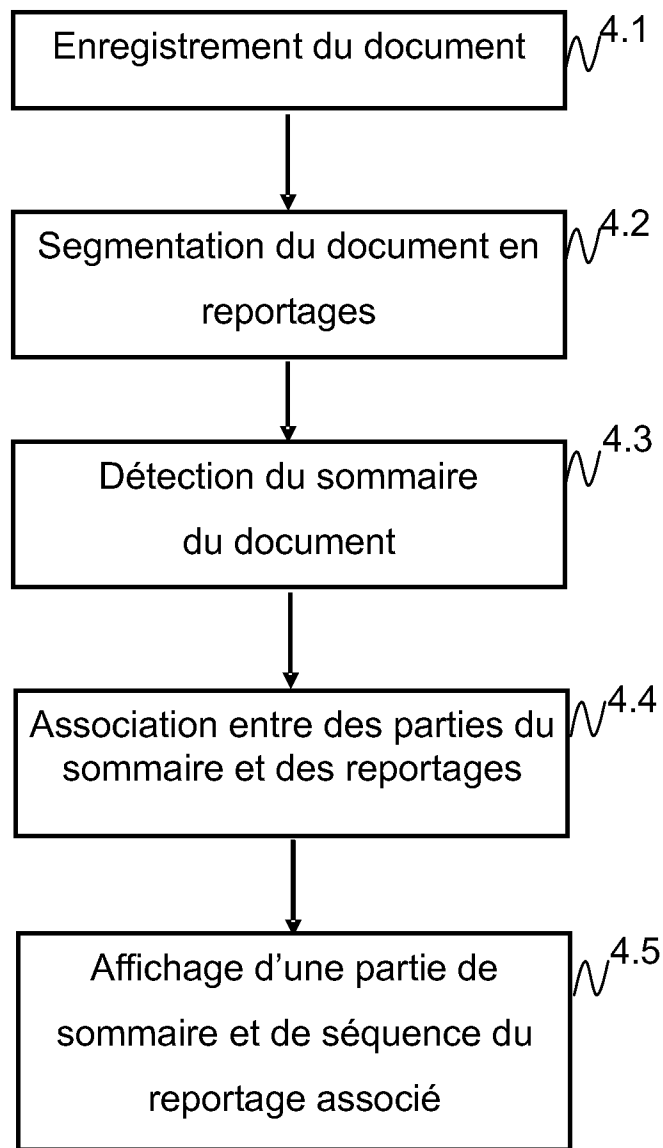
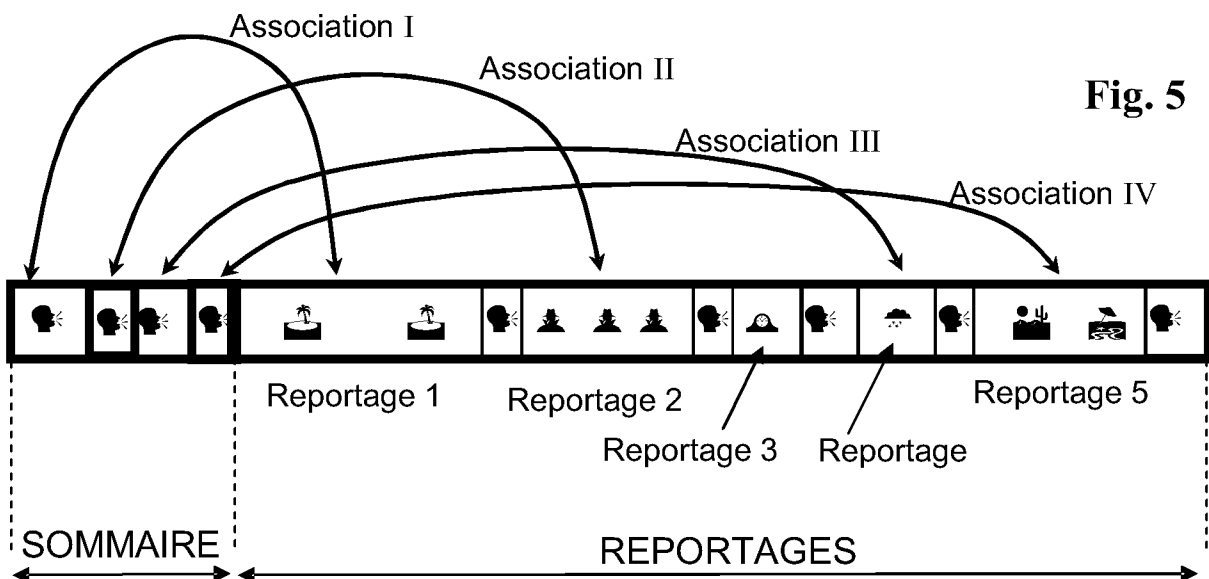


Fig. 5



INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2007/051650

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N5/76 H04N5/445 G11B27/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04N G11B G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2004/197088 A1 (FERMAN AHMET MUFIT [US] ET AL) 7 October 2004 (2004-10-07) abstract paragraphs [0029], [0067], [0073] figure 16	1,10
Y	-----	2-8, 10-16
Y	US 6 404 925 B1 (FOOTE JONATHAN T [US] ET AL) 11 June 2002 (2002-06-11) abstract column 1, lines 30-45 column 2, lines 1-40 claim 1	2-8, 10-16
A	----- -/--	1-10

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search

4 June 2007

Date of mailing of the international search report

13/06/2007

Name and mailing address of the ISA/
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Lauri, Lauro

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2007/051650

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2005/193408 A1 (SULL SANGHOON [KR] ET AL) 1 September 2005 (2005-09-01) abstract paragraphs [0093], [0233], [0243], [0281], [0297], [0340], [0342], [0356]	2-8, 10-16
A	-----	1-10
A	US 2003/108334 A1 (NEVENKA DIMITROVA [US] ET AL) 12 June 2003 (2003-06-12) abstract paragraphs [0003], [0012] claim 1 -----	1-16

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2007/051650

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2004197088	A1	07-10-2004	NONE
US 6404925	B1	11-06-2002	JP 2000298498 A 24-10-2000
US 2005193408	A1	01-09-2005	US 2005193425 A1 01-09-2005 US 2005204385 A1 15-09-2005
US 2003108334	A1	12-06-2003	AU 2002351026 A1 17-06-2003 CN 1599904 A 23-03-2005 EP 1485821 A2 15-12-2004 WO 03049430 A2 12-06-2003 JP 2005512212 T 28-04-2005

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°

PCT/EP2007/051650

A. CLASSEMENT DE L'OBJET DE LA DEMANDE INV. H04N5/76 H04N5/445 G11B27/00		
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB		
B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE		
Documentation minimale consultée (système de classification suivi des symboles de classement) H04N G11B G06F		
Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche		
Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés) EPO-Internal		
C. DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	US 2004/197088 A1 (FERMAN AHMET MUFIT [US] ET AL) 7 octobre 2004 (2004-10-07) abrégé alinéas [0029], [0067], [0073] figure 16	1, 10
Y	-----	2-8, 10-16
Y	US 6 404 925 B1 (FOOTE JONATHAN T [US] ET AL) 11 juin 2002 (2002-06-11) abrégé colonne 1, ligne 30-45 colonne 2, ligne 1-40 revendication 1	2-8, 10-16
A	-----	1-10
	-/--	
<input checked="" type="checkbox"/> Voir la suite du cadre C pour la fin de la liste des documents <input checked="" type="checkbox"/> Les documents de familles de brevets sont indiqués en annexe		
* Catégories spéciales de documents cités:		
A document définissant l'état général de la technique, non considéré comme particulièrement pertinent		*T* document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention
E document antérieur, mais publié à la date de dépôt international ou après cette date		*X* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément
L document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)		*Y* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier
O document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens		*&* document qui fait partie de la même famille de brevets
P document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée		
Date à laquelle la recherche internationale a été effectivement achevée 4 juin 2007		Date d'expédition du présent rapport de recherche internationale 13/06/2007
Nom et adresse postale de l'administration chargée de la recherche internationale Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Fonctionnaire autorisé Lauri, Lauro

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°

PCT/EP2007/051650

C(suite), DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
Y	US 2005/193408 A1 (SULL SANGHOON [KR] ET AL) 1 septembre 2005 (2005-09-01) abrégé alinéas [0093], [0233], [0243], [0281], [0297], [0340], [0342], [0356]	2-8, 10-16
A	-----	1-10
A	US 2003/108334 A1 (NEVENKA DIMITROVA [US] ET AL) 12 juin 2003 (2003-06-12) abrégé alinéas [0003], [0012] revendication 1 -----	1-16

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale n°

PCT/EP2007/051650

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 2004197088 A1	07-10-2004	AUCUN	
US 6404925 B1	11-06-2002	JP 2000298498 A	24-10-2000
US 2005193408 A1	01-09-2005	US 2005193425 A1 US 2005204385 A1	01-09-2005 15-09-2005
US 2003108334 A1	12-06-2003	AU 2002351026 A1 CN 1599904 A EP 1485821 A2 WO 03049430 A2 JP 2005512212 T	17-06-2003 23-03-2005 15-12-2004 12-06-2003 28-04-2005