

①2 DEMANDE DE BREVET D'INVENTION

A1

②2 Date de dépôt : 30.08.13.

③0 Priorité : 06.09.12 RU 2012138174; 06.03.13 IB
PCT2013051782.

④3 Date de mise à la disposition du public de la
demande : 07.03.14 Bulletin 14/10.

⑤6 Liste des documents cités dans le rapport de
recherche préliminaire : *Ce dernier n'a pas été
établi à la date de publication de la demande.*

⑥0 Références à d'autres documents nationaux
apparentés :

⑦1 Demandeur(s) : S.I.SV.EL SOCIETA' ITALIANA PER
LO SVILUPPO DELL'ELETTRONICA S.P.A. — IT.

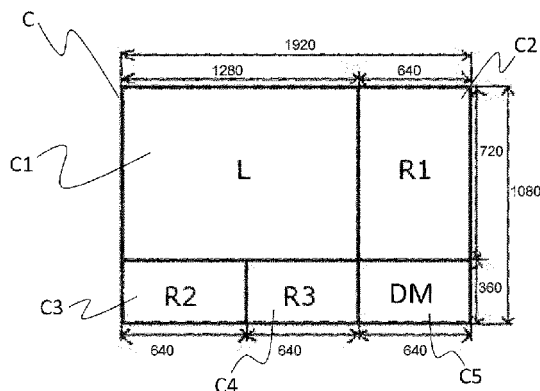
⑦2 Inventeur(s) : D'AMATO PAOLO, BALLOCCA
GIOVANNI, BUSHLANOV FEDOR et POLYAKOV
ALEXEY.

⑦3 Titulaire(s) : S.I.SV.EL SOCIETA' ITALIANA PER LO
SVILUPPO DELL'ELETTRONICA S.P.A..

⑦4 Mandataire(s) : CABINET PLASSERAUD.

⑤4 PROCÉDE POUR GÉNÉRER, TRANSMETTRE ET RECEVOIR DES IMAGES STÉRÉOSCOPIQUES, ET
DISPOSITIFS CONNEXES.

⑤7 La présente invention concerne un procédé pour gé-
nérer un flux vidéo stéréoscopique (101) comprenant des
images composites (C) qui comprennent des informations
sur des images droite (R) et gauche (L), et une carte de pro-
fondeur. Des pixels sont sélectionnés à partir des images
droite (R) et gauche (L), puis entrés dans une image com-
posite (C) du flux vidéo stéréoscopique. Le procédé com-
prend l'entrée de tous les pixels des images droite (R) et
gauche (L) dans l'image composite (C), une desdites deux
images reste inchangée, l'autre divisée en régions (R1, R2,
R3) comprenant une pluralité de pixels. Les pixels de la ou
des carte(s) de profondeur sont entrés dans la région de
l'image composite qui n'est pas occupée par des pixels des
images droite et gauche. L'invention concerne un procédé
pour reconstruire les images droite et gauche, depuis une
image composite, et des dispositifs permettant l'implémen-
tation desdits procédés.



PROCÉDÉ POUR GÉNÉRER, TRANSMETTRE ET RECEVOIR DES IMAGES
STÉRÉOSCOPIQUES, ET DISPOSITIFS CONNEXES

La présente invention concerne la génération, le stockage, la transmission, la réception et la reproduction de flux vidéo stéréoscopiques, à savoir des flux vidéo qui, lorsqu'ils sont traités de façon appropriée dans un dispositif de visualisation, produisent des séquences d'images qui sont perçues comme étant tridimensionnelles
5 par un téléspectateur.

Comme l'on sait, la perception de tridimensionnalité peut être obtenue en reproduisant deux images, une pour l'œil droit du téléspectateur et l'autre pour l'œil gauche du téléspectateur.

Un flux vidéo stéréoscopique transporte donc des informations sur deux
10 séquences d'images, correspondant aux perspectives droite et gauche d'un objet ou d'une scène. Un tel flux peut également transporter des informations supplémentaires.

La demande de brevet international PCT/IB2010/055918, publiée le 30 juin 2011 sous la référence WO2011/077343A1, décrit un procédé de multiplexage d'image gauche/droite et un procédé de démultiplexage (ainsi que des dispositifs
15 connexes) qui permettent de préserver l'équilibre entre des résolutions horizontale et verticale, offrant ainsi des avantages par rapport à des techniques connues telles que « côte à côte » et « haut et bas ».

Selon ledit procédé de multiplexage, les pixels de la première image (par exemple l'image gauche) sont entrés dans l'image composite inchangés, alors que la
20 seconde image est divisée en régions dont les pixels sont arrangés dans des zones libres de l'image composite, comme cela est représenté sur la figure 1, qui représente le cas dans lequel deux images dites 720p sont entrées dans une image conteneur 1080p.

Dans la réception, l'image divisée en régions est reconstruite et puis envoyée
25 à l'affichage. Par exemple, des affichages sont connus qui fonctionnent conformément au principe dit « image alternative », à savoir représentant les deux images L et R (L pour « Left » ou gauche, R pour « Right » ou droite) en succession temporelle. Pour la vision stéréoscopique, des lunettes dites « actives » doivent être portées, à savoir des lunettes qui, synchronisées avec la succession d'images L et R,

obscurcissent un verre et maintiennent le verre de l'autre œil ouvert, pour que chaque œil puisse seulement voir l'image prévue pour lui.

On sait que la vision stéréoscopique par l'intermédiaire de tels affichages peut s'avérer être agaçante pour certains téléspectateurs, auxquels il serait souhaitable
5 d'offrir la possibilité de varier (réduire) la profondeur des images afin de l'adapter à leurs préférences subjectives et à la taille de l'écran. Pour ce faire, il est nécessaire de fournir, à l'intérieur de l'affichage, une synthèse d'images intermédiaires, entre celles en train d'être transmises, qui seront alors affichées à la place des images réellement transmises. Une telle reconstruction peut être réalisée en utilisant des
10 techniques connues, si une ou plusieurs cartes de profondeur associées aux images transmises sont disponibles.

En outre, des affichages dits auto-stéréoscopiques, qui ne nécessitent pas l'utilisation de lunettes, ont récemment commencé à apparaître sur le marché. De même, de tels affichages réalisent une synthèse d'images non transmises, et donc
15 nécessitent au moins une carte de profondeur fournissant les informations nécessaires pour une telle synthèse.

Il est ainsi devenu nécessaire d'introduire un nouveau format pour générer, transporter et reconstruire des flux stéréoscopiques, lequel format peut être utilisé pour des dispositifs classiques de réception et de reproduction 2D et pour des
20 dispositifs actuels de réception et de reproduction 3D stéréoscopiques à deux vues (avec ou sans réglage de profondeur), ainsi que pour des dispositifs auto-stéréoscopiques futurs utilisant plus de deux vues, tout en préservant en même temps la compatibilité du format avec les infrastructures de distribution et de production de flux vidéo et les dispositifs actuellement utilisés.

25 L'objet de la présente invention est donc de proposer un procédé pour générer, transmettre et recevoir des images stéréoscopiques, et des dispositifs connexes, dont l'objectif est de satisfaire les conditions décrites ci-dessus.

L'invention concerne un procédé et un dispositif pour multiplexer les deux images relatives aux perspectives droite et gauche (ci-après appelées image droite et
30 image gauche), ainsi qu'une ou plusieurs cartes de profondeur, à l'intérieur d'une image composite unique.

Plus particulièrement, l'invention concerne un procédé pour générer un flux vidéo stéréoscopique comprenant des images composites C, lesdites images

- composites C comprenant des informations sur une image droite R et une image gauche L d'un contenu tridimensionnel vidéo, dans lequel des pixels de ladite image droite R et des pixels de ladite image gauche L sont sélectionnés, et lesdits pixels sélectionnés sont entrés dans une image composite C dudit flux vidéo stéréoscopique,
- 5 le procédé étant caractérisé en ce que tous les pixels de ladite image droite R et tous les pixels de ladite image gauche L sont entrés dans ladite image composite C en laissant une desdites deux images inchangées, en divisant l'autre en un nombre de régions R1, R2, R3 possédant une superficie totale égale à celle de ladite autre image R, et en entrant lesdites régions dans ladite image composite C,
- 10 dans lequel ladite image composite C possède des dimensions plus importantes que ce qui est nécessaire pour entrer tous les pixels de ladite image gauche L et de ladite image droite R, et dans lequel, dans les pixels de l'image composite C restant après ladite entrée, au moins une carte de profondeur ou de disparité DM est entrée qui est relative à la profondeur ou la disparité des pixels dudit contenu vidéo tridimensionnel,
- 15 lesdites cartes visant à reconstruire, en réception, des images non transmises dans ledit flux stéréoscopique.

Dans divers modes de réalisation du procédé selon l'invention, on peut éventuellement avoir recours en outre à l'une et/ou à l'autre des dispositions suivantes :

- 20 - ladite au moins une carte de profondeur ou de disparité DM est codée sous forme d'image à niveaux de gris ;
- le contenu d'informations vidéo de ladite au moins une carte de profondeur ou de disparité est transporté par un seul signal de luminance, sans utiliser de signal de chrominance ;
- 25 - ladite au moins une carte de profondeur ou de disparité DM possède une résolution inférieure à celle d'une version d'origine de celle-ci, la résolution de cette dernière étant égale à celle desdites images gauche L et image droite R ;
- ladite au moins une carte de profondeur ou de disparité DM est obtenue en soumettant ladite carte de profondeur d'origine à un sous-échantillonnage 4 à 1. ;
- 30 - ladite au moins une carte de profondeur ou de disparité DM est une carte de profondeur ou de disparité associée à l'une des deux images droite R ou gauche L, ou associée à un point de vue intermédiaire, entre L et R ; - ladite au moins une carte de profondeur ou de disparité DM comprend une carte de profondeur ou de disparité

associée à une image droite R et une carte de profondeur ou de disparité associée à une image gauche L ;

- lesdites cartes de profondeur ou de disparité associées à une image droite R et une image gauche L sont entrées dans lesdits pixels restants de l'image composite C au moyen de techniques de mise en paquet d'images ;
- si ledit nombre de régions est trois, lesdites régions sont obtenues par l'intermédiaire des étapes de :
 - division de ladite autre image R en deux parties possédant la même dimension horizontale R1, R2, R3 ;
 - division d'une desdites deux parties R2, R3 en deux parties possédant la même dimension verticale R2, R3 ;
- des informations de signalisation identifiant le type de flux vidéo généré sont entrées sous forme de métadonnées dans ladite image composite C ou ledit flux vidéo stéréoscopique ;
- lesdites informations de signalisation sont adaptées afin de permettre de distinguer entre au moins deux des types suivant d'images composites :
 - image composite du type à format mosaïque sans cartes de profondeur ;
 - image composite du type à format mosaïque avec une carte de profondeur ;
 - image composite du type à format mosaïque avec deux cartes de profondeur en configuration côte à côte ;
 - image composite du type à format mosaïque avec deux cartes de profondeur en configuration haut et bas ;
 - image composite du type à format mosaïque avec deux cartes de profondeur en configuration à format mosaïque.

L'invention concerne également un procédé et un dispositif pour démultiplexer ladite image composite, à savoir pour extraire à partir de celle-ci les images droite et gauche et la ou les carte(s) de profondeur entrées par le dispositif de multiplexage.

Plus particulièrement, l'invention concerne un dispositif pour générer un flux vidéo stéréoscopique comprenant des images composites C, lesdites images composites C comprenant des informations sur une image droite R, une image

gauche L, caractérisé en ce qu'il comprend des moyens pour implémenter les étapes du procédé selon l'invention.

L'invention concerne également un procédé pour reconstruire au moins une paire d'images d'un flux vidéo stéréoscopique commençant à partir d'une image composite C, ladite image composite C comprenant des informations sur une image droite R, une image gauche L, le procédé comprenant les étapes de :

- génération d'une première image desdites images droite R et gauche L en copiant un seul groupe de pixels contigus à partir d'une première région C1 de ladite image composite,
- 10 - génération de l'autre image desdites images droite R et gauche L en copiant d'autres groupes de pixels contigus à partir d'un nombre de régions distinctes C2, C3, C4 de ladite image composite, ledit nombre de régions distinctes étant différent de ladite première région C1 ;
- génération d'au moins une carte de profondeur ou de disparité DM en copiant au moins un groupe de pixels contigus à partir d'une région supplémentaire C5 de ladite image composite, différente de ladite première région et dudit nombre de régions distinctes.
- 15

Dans divers modes de réalisation du procédé selon l'invention, on peut éventuellement avoir recours en outre à l'une et/ou à l'autre des dispositions suivantes :

- ledit nombre de régions est trois :
 - une C2 desdites régions de l'image composite C possède la même dimension verticale que ladite première région C1 et la moitié de sa dimension horizontale ;
 - 25 - les deux autres C3, C4 desdites régions de l'image composite C possèdent des dimensions horizontale et verticale égales, et la moitié de la dimension verticale de ladite première région C1.
- ladite au moins une carte de profondeur ou de disparité DM est générée en commençant à partir d'une image à niveaux de gris dérivée d'un signal de luminance contenu dans les pixels contigus de ladite région supplémentaire C5 ;
- 30 - le procédé comprend une étape d'augmentation des dimensions horizontale et verticale de ladite au moins une carte de profondeur ou de disparité DM jusqu'à une dimension égale à celle desdites images droite R et gauche L ;

- le procédé comprend une étape d'obtention, à partir de ladite image composite C ou à partir du flux vidéo, d'informations de signalisation adaptées pour reconnaître le type de flux vidéo généré ;

- lesdites informations de signalisation sont adaptées afin de distinguer entre au moins deux des types suivants d'image composites :

- image composite du type à format mosaïque sans cartes de profondeur ;
- image composite du type à format mosaïque avec une carte de profondeur ;
- image composite du type à format mosaïque avec deux cartes de profondeur en configuration côte à côte ;
- image composite du type à format mosaïque avec deux cartes de profondeur en configuration haut et bas ;
- image composite du type à format mosaïque avec deux cartes de profondeur en configuration à format mosaïque.

L'invention concerne également un dispositif pour reconstruire au moins une paire d'images d'un flux vidéo stéréoscopique commençant à partir d'une image composite C, ladite image composite C comprenant des informations sur une image droite R, une image gauche L, le dispositif comprenant :

- des moyens pour générer une première image desdites images droite R et gauche L en copiant un seul groupe de pixels contigus à partir d'une première région C1 de ladite image composite,
- des moyens pour générer l'autre image desdites images droite R et gauche L en copiant d'autres groupes de pixels contigus à partir d'un nombre de régions distinctes C2, C3, C4 de ladite image composite, ledit nombre de régions distinctes étant différent de ladite première région C1 ;
- des moyens pour générer au moins une carte de profondeur ou de disparité DM en copiant au moins un groupe de pixels contigus à partir d'une région supplémentaire C5 de ladite image composite, différent de ladite première région et dudit nombre de régions distinctes.

Dans divers modes de réalisation du dispositif selon l'invention, on peut éventuellement avoir recours en outre à l'une et/ou à l'autre des dispositions suivantes :

- si ledit nombre de régions est trois :

- une C2 desdites régions de l'image composite C possède la même dimension verticale que ladite première région C1 et la moitié de sa dimension horizontale ;
- les deux autres C3, C4 desdites régions de l'image composite C possèdent des dimensions horizontale et verticale égales, et la moitié de la dimension verticale de ladite première région C1 ;
- lesdits moyens pour générer au moins une carte de profondeur ou de disparité DM utilisent une image à niveaux de gris dérivée d'un signal de luminance contenu dans les pixels contigus de ladite région supplémentaire C5 ;
- le dispositif comprend des moyens pour augmenter les dimensions horizontale et verticale de ladite au moins une carte de profondeur ou de disparité DM jusqu'à une dimension égale à celle desdites images droite R et gauche L ;
- le dispositif comprend des moyens adaptés pour reconnaître le type de flux vidéo reçu en fonction d'informations de signalisation identifiant ledit type de flux, contenues dans ladite image composite C ou dans ledit flux vidéo ;
- lesdites informations de signalisation permettent de distinguer entre au moins deux des types suivants d'image composites :
 - image composite du type à format mosaïque sans cartes de profondeur ;
 - image composite du type à format mosaïque avec une carte de profondeur ;
 - image composite du type à format mosaïque avec deux cartes de profondeur en configuration côte à côte ;
 - image composite du type à format mosaïque avec deux cartes de profondeur en configuration haut et bas ;
 - image composite du type à format mosaïque avec deux cartes de profondeur en configuration à format mosaïque ;
- le dispositif comprend des moyens qui, en fonction desdites informations utiles pour distinguer un type d'image composite, sont adaptés pour produire :
 - seulement ladite première image desdites images droite R et gauche L ; ou
 - ladite première et ladite seconde desdites images droite R et gauche L ; ou
 - ladite première et ladite seconde desdites images droite R et gauche L et ladite au moins une carte de profondeur ou de disparité DM ;

- le dispositif comprend des moyens adaptés pour réaliser une ou plusieurs des opérations suivantes, en fonction desdites informations de signalisation :
 - la reconnaissance du type d'image reçu, dans le but de reconstruire correctement les deux images droite et gauche du contenu vidéo tridimensionnel ;
 - la reconnaissance de la présence d'une ou deux cartes de profondeur ou de disparité et du type de configuration de celles-ci ;
 - dans le cas de deux cartes de profondeur ou de disparité, l'obtention de chacune des deux cartes ;
 - la réalisation, sur lesdites cartes de profondeur ou de disparité, d'opérations adaptées pour mettre les dimensions des cartes à des valeurs égales à celles des images du contenu vidéo ;
 - le dispositif comprend des moyens conçus afin de générer des images supplémentaires correspondant à des vues supplémentaires en commençant à partir desdites images droite R et gauche L et en utilisant lesdites cartes de profondeur.
 - le dispositif comprend des moyens adaptés pour afficher deux séquences d'images, au moins l'une d'entre elles comprenant des images synthétisées en commençant à partir d'au moins une des vues transmises et à partir d'au moins une carte de profondeur ;
 - le dispositif comprend des moyens adaptés pour fournir au téléspectateur la possibilité de choisir des séquences d'images concernant des points de vue plus ou moins proches, afin de varier la perception de profondeur ;
 - le dispositif comprend un affichage auto-stéréoscopique, et comprenant des moyens adaptés pour utiliser lesdites images supplémentaires correspondant à des vues supplémentaires afin de permettre à des téléspectateurs positionnés à des points différents dans l'espace de voir des séquences d'images différentes.
- L'invention concerne également un flux vidéo stéréoscopique caractérisé en ce qu'il comprend au moins une image composite C générée au moyen du procédé selon l'invention.
- Comme on peut le voir sur la figure 1a, concernant la demande de brevet international susmentionnée (format dit « mosaïque »), dans l'image composite, il y a une région non utilisée (C5) dont les dimensions sont la moitié, à la fois horizontalement et verticalement, de celles des deux images L et R. Selon un mode

de réalisation possible de l'invention, au moins une carte de profondeur (DM) peut être entrée dans ladite région non utilisée, comme cela est représenté sur la figure 1b.

Une carte de profondeur relative à une image x doit être entendue comme étant une image à niveaux de gris dans laquelle chaque pixel possède une valeur de
5 luminance qui est proportionnelle à la profondeur, à savoir la coordonnée « z », du pixel lui-même, où, par convention, il est supposé que la valeur $z = 0$ correspond à la position sur l'écran, et des valeurs positives de z correspondent à des pixels positionnés derrière l'écran, alors que des valeurs négatives correspondent à des pixels positionnés devant l'écran. Comme la région non utilisée de l'image
10 composite possède des dimensions horizontale et verticale qui sont la moitié des dimensions des images L et R, dans un mode de réalisation de la présente invention, il est possible d'entrer, dans une telle région, une carte de profondeur (relative à une des deux images L et R) possédant des résolutions horizontale et verticale égales à la moitié de l'image correspondante. On a observé qu'une telle perte de résolution n'est
15 pas préjudiciable car, étant donné l'inexactitude avec laquelle des cartes de profondeur peuvent généralement être calculées ou mesurées, il est préférable de soumettre des cartes à résolution complète à des opérations de sous-échantillonnage en réalisant des interpolations entre les valeurs de pixel, car de telles opérations peuvent réduire la composante de bruit, entraînant des images reconstruites de plus
20 haute qualité.

Selon d'autres modes de réalisation de l'invention, il est possible d'entrer deux cartes de profondeur dans ladite région non utilisée (C5).

La demande de brevet international susmentionnée décrit également d'autres formes de multiplexage et de démultiplexage des images stéréoscopiques L et R,
25 auxquelles le procédé de la présente invention peut être appliqué également, bien que moins efficacement car l'espace laissé disponible pour entrer la carte de profondeur est plus petit. Par conséquent, il peut y avoir une réduction supplémentaire de la résolution de ladite carte. Bien qu'étant toujours au sein des principes généraux de la présente invention, de telles autres implémentations ne seront pas décrites dans les
30 présentes.

Un objet particulier de la présente invention est de proposer un procédé pour générer, transmettre et recevoir des images stéréoscopiques, et des dispositifs

connexes, tels qu'ils sont présentés dans les revendications jointes, qui font partie intégrante de la présente description.

Des objets et avantages supplémentaires de la présente invention deviendront plus évidents à partir de la description suivante de quelques modes de réalisation de celle-ci, qui sont fournis à titre d'exemple non limitatif en référence aux dessins joints, sur lesquels :

la figure 1a représente l'image composite dans le format d'art antérieur (format mosaïque) ;

la figure 1b représente un exemple d'une image composite selon la présente invention ;

la figure 2 représente un schéma de principe d'un dispositif pour multiplexer l'image droite, l'image gauche et une carte de profondeur en une image composite ;

la figure 3 est un organigramme d'un procédé exécuté par le dispositif de la figure 2 ;

la figure 4 représente une forme possible de désassemblage d'une image destinée à être entrée dans une image composite ;

la figure 5 représente un schéma de principe d'un dispositif pour extraire l'image gauche, l'image droite et une carte de profondeur à partir de l'image composite ;

la figure 6 est un organigramme d'un procédé exécuté par le dispositif de la figure 5.

La figure 2 représente un schéma de principe d'un dispositif 100 pour générer un flux vidéo stéréoscopique 101 avec au moins une carte de profondeur, conformément aux variantes de l'invention.

Sur la figure 2, le dispositif 100 reçoit deux séquences d'images 102 et 103, par exemple deux flux vidéo respectivement prévus pour l'œil gauche (L) et pour l'œil droit (R), ainsi qu'une séquence de cartes de profondeur 106 concernant le contenu tridimensionnel associé au flux vidéo stéréoscopique.

La carte de profondeur de la séquence 106 peut être associée à une des deux images droite et gauche appartenant respectivement aux séquences 102 et 103, ou elle peut être créée sous forme d'interpolation entre les cartes de profondeur pour les images droite et gauche, à savoir relatives à un point de vue intermédiaire de la scène.

Dans ce premier mode de réalisation, qui va être décrit ci-dessous, la carte de profondeur est générée par l'intermédiaire d'un algorithme parmi ceux déjà connus dans l'état de l'art, qui sont fondés, par exemple, sur une comparaison entre une image droite et une image gauche, et qui retourne une matrice (à savoir la carte de
5 profondeur), dont la taille est égale aux pixels d'une des deux images comparées, et dont les éléments possèdent une valeur qui est proportionnelle à la profondeur de chaque pixel de ladite image. Une autre technique de génération de carte de profondeur est fondée sur la mesure de la distance de l'objet dans la scène à partir de la paire de caméras vidéo qui réalisent des prises de vue de la scène : cette distance
10 peut être facilement mesurée au moyen d'un laser. Dans le cas de flux vidéo artificiels générés à l'aide d'ordinateurs électroniques, les caméras vidéo sont virtuelles, car elles consistent en deux points de vue d'une certaine scène artificiellement créée par un ordinateur. Dans un tel cas, les cartes de profondeur sont générées par l'ordinateur et sont très précises.

15 À titre de possibilité autre que l'exemple de la figure 2, les cartes de profondeur de la séquence 106 peuvent être générées à l'intérieur du dispositif 100. Dans ce cas, le dispositif 100, au lieu de recevoir la séquence de cartes de profondeur à partir de l'extérieur, comprend un module approprié (comme cela est représenté sur le dessin), dans lequel sont entrées les images L et R des séquences 102 et 103, et qui
20 calcule ensuite les cartes de profondeur correspondantes.

Le dispositif 100 permet d'implémenter un procédé pour multiplexer deux images des deux séquences 102 et 103 et la carte de profondeur de la séquence 106.

Afin d'implémenter le procédé pour multiplexer les images droite et gauche et la carte de profondeur, le dispositif 100 comprend un module de
25 désassemblage 104 pour diviser une image d'entrée (l'image droite dans l'exemple de la figure 1b) en une pluralité de sous-images, chacune correspondant à une région de l'image reçue, un module de sous-échantillonnage et de filtrage 107 pour traiter la carte de profondeur, et un module d'assemblage 105 capable d'entrer les pixels d'images reçues, y compris la carte de profondeur, dans une image composite unique
30 destinée à être fournie en sortie. Si aucun traitement de la séquence 106 n'est nécessaire, le module 107 peut être omis. Ceci peut être le cas, par exemple, lorsque la carte de profondeur est générée par laser et possède, dès le début, une résolution plus basse que celle des images L et R.

Un exemple d'un procédé de multiplexage implémenté par le dispositif 100 va à présent être décrit en faisant référence à la figure 3.

Le procédé commence à l'étape 200. Ensuite (étape 201), une des deux images entrées (droite ou gauche) est divisée en une pluralité de régions, comme cela est représenté sur la figure 4. Dans l'exemple de la figure 4, l'image désassemblée est une image R d'un flux vidéo 720p, à savoir un format progressif avec une résolution de 1280 x 720 pixels.

L'image R de la figure 4 provient du flux vidéo 103 qui transporte les images prévues pour l'œil droit, et est désassemblée en trois régions R1, R2 et R3, de préférence de forme rectangulaire.

Le désassemblage de l'image R est obtenu en la divisant en deux parties de la même taille et ensuite en sous-divisant une de ces parties en deux parties de la même taille.

La région R1 possède une taille de 640 x 720 pixels et est obtenue en prenant tous les 640 premiers pixels de chaque rangée. La région R2 possède une taille de 640 x 360 pixels et est obtenue en prenant les pixels de 641 à 1280 des 360 premières rangées. La région R3 possède une taille de 640 x 360 pixels et est obtenue en prenant les pixels restants de l'image R, à savoir les pixels de 641 à 1280 des 360 dernières rangées.

Dans l'exemple de la figure 2, l'étape de désassemblage de l'image R est réalisée par le module 104, qui reçoit une image entrée R (dans ce cas l'image R) et produit trois sous-images (à savoir trois groupes de pixels) correspondant aux trois régions R1, R2 et R3.

Ensuite (étapes 202, 203 et 204) l'image composite C est construite, qui comprend les informations concernant les images droite et gauche et la carte de profondeur reçue ; dans l'exemple décrit dans les présentes, ladite image composite C est une image du flux vidéo stéréoscopique produit, et donc elle est également appelée image conteneur.

Tout d'abord (étape 202), l'image entrée reçue par le dispositif 100 et non désassemblée par le dispositif 105 (l'image gauche L dans l'exemple de la figure 2) est entrée inchangée dans une zone non divisée à l'intérieur d'une image conteneur, qui est dimensionnée de manière telle à comprendre tous les pixels des deux images entrées. Par exemple, si les images entrées possèdent une taille de 1280 x 720 pixels,

alors une image conteneur appropriée pour contenir les deux sera une image de 1920 x 1080 pixels, par exemple une image d'un flux vidéo du type 1080p (format progressif avec 1920 x 1080 pixels).

Dans l'exemple de la figure 1, l'image gauche L est entrée dans l'image conteneur C et positionnée dans le coin gauche supérieur. Ceci est obtenu en copiant les 1280 x 720 pixels de l'image L dans une zone C1 constituée des 1280 premiers pixels des 720 premières rangées de l'image conteneur C.

Dans l'étape suivante 203, l'image désassemblée dans l'étape 201 par le module 104 est entrée dans l'image conteneur. Ceci est réalisé par le module 105 en copiant les pixels de l'image désassemblée dans l'image conteneur C dans les zones de celle-ci qui n'ont pas été occupées par l'image L, à savoir des zones externes à la zone C1.

Afin d'atteindre la meilleure compression possible et de réduire la génération d'artefacts lors de la décompression du flux vidéo, les pixels des sous-images produites par le module 104 sont copiés en préservant les relations spatiales respectives. Autrement dit, les régions R1, R2 et R3 sont copiées dans des zones respectives de l'image C sans subir de déformation, exclusivement au moyen d'opérations de translations.

Un exemple de l'image conteneur C produite par le module 105 est représenté sur la figure 1b.

La région R1 est copiée dans les 640 derniers pixels des 720 premières rangées (zone C2), à savoir à côté de l'image L copiée auparavant.

Les régions R2 et R3 sont copiées sous la zone C1, à savoir respectivement dans les zones C3 et C4, qui comprennent respectivement les 640 premiers pixels et les 640 pixels suivants des 360 dernières rangées.

Les opérations pour entrer les images L et R dans l'image conteneur n'impliquent aucune modification de l'équilibrage entre les résolutions horizontale et verticale.

La technique décrite ci-dessus pour entrer des images L et R dans l'image conteneur C sera ci-après définie comme étant de type à format mosaïque.

Dans les pixels libres de l'image C, à savoir dans la zone C5, le module 105 entre, sous forme d'image, la carte de profondeur (DM) concernant la paire

stéréoscopique L et R (étape 204). Avant l'étape 204, la carte de profondeur DM peut être sous-échantillonnée, filtrée ou traitée davantage par le module 107.

La carte de profondeur est de préférence codée sous forme d'image à niveaux de gris, dont le contenu d'informations peut donc être transporté par le signal de
 5 luminance seul ; des chrominances ne sont pas utilisées et peuvent être, par exemple, nulles ; ceci permet d'obtenir une compression efficace de l'image conteneur C.

Dans un mode de réalisation préféré, la carte de profondeur DM possède une résolution de 640 x 360 pixels, correspondant à un sous-échantillonnage 4 à 1 (ou décimation) de la carte de profondeur d'origine possédant une résolution
 10 de 1280 x 720 pixels, correspondant à celle des images L et R. Chaque pixel de la carte sous-échantillonnée DM correspond à une région de pixel 2 x 2 de la carte d'origine. L'opération de sous-échantillonnage est typiquement réalisée en utilisant des procédures qui sont connues en elles-mêmes dans l'état de l'art.

L'image C ainsi obtenue est ensuite compressée et transmise à, ou
 15 sauvegardée sur, un support de stockage (par exemple un DVD). Dans ce but, des moyens de compression sont prévus qui sont adaptés pour compresser une image ou un signal vidéo, conjointement avec des moyens pour enregistrer et/ou transmettre l'image ou signal vidéo compressé.

La figure 5 représente un schéma de principe d'un récepteur 1100 qui
 20 décompresses l'image conteneur reçue (si elle est compressée), reconstruit les deux images droite et gauche, et les met à la disposition d'un dispositif de visualisation (par exemple un poste de télévision) permettant la réalisation de contenus 3D. Le récepteur 1100 peut être un boîtier décodeur ou un récepteur construit dans un poste de télévision.

25 Les mêmes remarques réalisées pour le récepteur 1100 sont également applicables à un lecteur d'image stocké (par exemple un lecteur DVD) qui lit une image conteneur (éventuellement compressée) et la traite afin d'obtenir une paire d'images correspondant aux images droite et gauche entrées dans l'image conteneur (éventuellement compressée) lue par le lecteur.

30 En référence à la figure 5, le récepteur reçoit (par l'intermédiaire de câble ou d'antenne) un flux vidéo stéréoscopique compressé 1101 et le décompresses au moyen d'un module de décompression 1102, obtenant ainsi un flux vidéo comprenant une séquence d'images C' correspondant aux images C. Dans le cas d'un

canal idéal ou si des images conteneurs sont lues à partir d'une mémoire de masse ou un support de données (Blu-ray, CD, DVD), les images C' correspondent aux images conteneurs C transportant les informations sur les images droite et gauche et la carte de profondeur, à l'exception des éventuels artefacts introduits par le procédé de compression.

Ces images C' sont alors fournies à un module de reconstruction 1103, qui exécute une reconstruction d'image et un procédé d'extraction de carte de profondeur, comme cela est décrit ci-dessous en référence à la figure 6.

Il est évident que, si le flux vidéo n'est pas compressé, le module de décompression 1102 peut être omis et le signal vidéo peut être fourni directement au module de reconstruction 1103.

Le procédé de reconstruction commence à l'étape 1300, lorsque l'image conteneur décompressée C' est reçue.

Le module de reconstruction 1103 extrait (étape 1301) l'image gauche L en copiant les 720 x 1080 premiers pixels contigus de l'image décompressée dans une nouvelle image qui est plus petite que l'image conteneur, par exemple une image d'un flux 720p. L'image L ainsi reconstruite est envoyée à la sortie du récepteur 1100 (étape 1302).

Le terme « pixels contigus » fait référence à des pixels d'une image inchangée appartenant à une zone non divisée de l'image.

Ensuite, le procédé permet d'extraire l'image droite R à partir de l'image conteneur C'.

L'étape d'extraction de l'image droite (voir également figure 4) commence en copiant (étape 1303) la zone R1 présente dans l'image C'. De façon plus détaillée, les pixels des 640 colonnes de R1 sont copiés dans les 640 premières colonnes correspondantes de la nouvelle image qui représente l'image reconstruite Rout. Ensuite, R2 est extraite (étape 1304). À partir de l'image décompressée C' (qui, comme susmentionné, correspond à l'image C de la figure 1b), les pixels de la zone C3 (correspondant à la région source R2) sont sélectionnés. À ce stade, les 640 colonnes de pixels sont copiées dans les colonnes libres adjacentes à celles juste copiées à partir de R1.

En ce qui concerne R3 (étape 1305), les pixels de la région C4 sont extraits à partir de l'image C' et sont copiés dans la dernière colonne libre dans le coin gauche inférieur de l'image reconstruite.

À ce stade, l'image droite Rout a été complètement reconstruite et peut être
5 produite (étape 1306).

Enfin, le module de reconstruction 1103 extrait (étape 1307) la carte de profondeur en copiant, dans une zone de mémoire, les valeurs de luminance des 640 x 320 derniers pixels de l'image conteneur décompressée C', correspondant à la zone C5. Le contenu de ladite zone de mémoire est envoyé au récepteur 1100
10 (étape 1302) et sera utilisé par l'affichage pour générer des images interpolées non transmises dans le flux vidéo stéréoscopique. Le procédé pour reconstruire les images droite et gauche et la carte de profondeur contenues dans l'image conteneur C' est ainsi terminé (étape 1309). Ledit procédé est répété pour chaque image du flux vidéo reçu par le récepteur 1100, pour que la sortie consiste de deux
15 flux vidéo 1104 et 1105 pour l'image droite et pour l'image gauche, respectivement, et d'un flux vidéo 1106 correspondant à la carte de profondeur.

Le procédé décrit ci-dessus pour reconstruire les images droite et gauche et la carte de profondeur pour la synthèse d'image est fondé sur l'hypothèse que le démultiplexeur 1100 sait comment l'image conteneur C a été construite et peut ainsi
20 extraire les images droite et gauche et la carte de profondeur de synthèse.

Naturellement, ceci est possible si le procédé de multiplexage est standardisé.

Afin de prendre en compte le fait que l'image conteneur peut être générée selon un procédé quelconque qui utilisent la solution qui est l'objet des revendications jointes, le démultiplexeur utilise de préférence des informations de
25 signalisation contenues sous forme de métadonnées dans une région prédéfinie de l'image composite ou dans le flux vidéo, qui identifient le type de flux vidéo généré pour connaître la manière de décompacter le contenu de l'image composite et la manière de reconstruire les images droite et gauche et la carte de profondeur pour la synthèse d'images stéréoscopiques supplémentaires.

30 Après avoir décodé la signalisation, le démultiplexeur connaîtra la position de l'image inchangée (par exemple l'image gauche dans les exemples décrits ci-dessus), ainsi que les positions des régions dans lesquelles l'autre image a été désassemblée

(par exemple l'image droite dans les exemples décrits ci-dessus) et la position de la carte de profondeur.

Avec ces informations, le démultiplexeur peut extraire l'image inchangée (par exemple l'image gauche) et la carte de profondeur et reconstruire l'image désassemblée (par exemple l'image droite).

Bien que la présente invention ait été illustrée jusqu'à présent en faisant référence à certains modes de réalisation préférés et avantageux, il est évident qu'elle n'est pas limitée à de tels modes de réalisation et que de nombreux changements peuvent lui être apportés par l'homme du métier souhaitant combiner en une image composite deux images relatives à deux perspectives différentes (droite et gauche) d'un objet ou d'une scène et la carte de profondeur associée.

Dans une variante possible, par exemple, au lieu d'entrer, dans l'image composite C, la carte de profondeur relative à une des deux images, une carte dite de « disparité » ou de « déplacement » est entrée. Selon des hypothèses appropriées (les prises de vue sont effectuées avec des caméras vidéo équipées d'optiques identiques), une telle carte peut être facilement dérivée de la carte de profondeur, à laquelle elle peut être facilement associée. Si les deux images droite et gauche sont affichées superposées sur le même affichage et des lunettes ne sont pas utilisées pour les séparer, on peut facilement réaliser que, afin d'obtenir une image à partir de l'autre, il est nécessaire de déplacer les objets d'une certaine quantité. Plus précisément, afin d'obtenir l'image droite en commençant à partir de l'image gauche, il est nécessaire de déplacer les objets situés derrière l'écran vers la droite d'une quantité qui augmente avec la profondeur à laquelle de tels objets sont situés. Les objets qui sont situés exactement sur l'écran n'ont pas besoin d'être déplacés, alors que les objets situés devant l'écran doivent être déplacés vers la gauche d'une quantité qui augmente en fonction de la distance à partir de l'écran.

Dans les conditions mentionnées auparavant, entre la profondeur P et la disparité D, une relation du type suivant existe :

$$D = I * P / (P + P_0)$$

où I est la distance interoculaire et P_0 est la distance du téléspectateur par rapport à l'écran. Il faut noter que, pour P se rapprochant de l'infini, D se rapprochera de I , et pour $P = 0$ (objets situés sur l'écran), D sera égal à 0.

5 Naturellement, afin de reconstruire une image intermédiaire, entre les images gauche et droite, il est possible d'adopter la même procédure décrite ci-dessus, mais les valeurs de disparité devront être multipliées par un coefficient c entre 0 et 1, qui est une fonction de la distance du point de vue intermédiaire par rapport au point de vue de l'image de référence (la gauche dans ce cas).

10 Il faut noter que, lorsque l'image droite est reconstruite en commençant à partir de la gauche conformément à la description ci-dessus, ou lorsqu'une image intermédiaire est reconstruite, certaines zones sont laissées non couvertes, qui correspondent aux pixels d'objets présents dans l'image droite mais non dans l'image gauche, car ils sont cachés par d'autres objets devant eux (appelés « occlusions »).

15 Afin de réaliser une reconstruction complète d'une image intermédiaire, il serait donc nécessaire d'avoir à disponibilité les images droite et gauche ainsi que les cartes de profondeur ou de disparité. De cette manière, en fait, les zones vides (occultées) peuvent être remplies en prenant les pixels correspondants à partir de l'autre image et en les déplaçant d'une quantité égale à la disparité relative multipliée par le coefficient $1 - c$.

20 Comme on peut l'entendre à partir de la description ci-dessus, une autre variante possible de l'invention peut nécessiter l'entrée de deux cartes de profondeur ou de disparité, au lieu d'une. De telles cartes, faisant respectivement référence à l'image gauche et à l'image droite, peuvent être entrées dans le même espace où une seule carte a été entrée dans le cas précédent, en utilisant des techniques connues de mise en paquet d'images telles que, par exemple, côte à côte ou haut et bas. Dans ce
25 premier cas, la résolution horizontale des deux cartes est réduite de moitié, alors que, dans ce dernier cas, la résolution verticale est réduite de moitié. Il est également possible d'utiliser une variante supplémentaire de la technique de mise en paquet d'images définie ci-dessus par « format mosaïque ».

30 Les procédures pour entrer les deux cartes côté génération et pour extraire les deux cartes côté réception peuvent être facilement dérivées de celles décrites en faisant référence au cas à carte unique, avec des variations évidentes bien connues de l'homme du métier.

Naturellement, la signalisation présente dans le flux vidéo doit également être capable de discerner la présence d'une ou de deux cartes. Par conséquent, ladite signalisation doit contenir des informations adaptées pour permettre la distinction entre au moins deux des types suivants d'images composites :

- 5 1) image composite du type à format mosaïque sans carte de profondeur ou de disparité (cas de la figure 1a) ;
- 2) image composite du type à format mosaïque avec une carte de profondeur ou de disparité (case de la figure 1b) ;
- et éventuellement également :
- 10 3) image composite du type à format mosaïque avec deux cartes de profondeur ou de disparité en configuration haut et bas ;
- 4) image composite du type à format mosaïque avec deux cartes de profondeur ou de disparité en configuration côte à côte ;
- 5) image composite du type à format mosaïque avec deux cartes de
- 15 profondeur ou de disparité en configuration à format mosaïque.

Le récepteur comprend de préférence un ou plusieurs blocs de traitement adaptés pour réaliser une ou plusieurs des opérations suivantes, en fonction des informations de signalisation :

- la reconnaissance du type d'image reçu, dans le but de reconstruire
- 20 correctement les deux images droite et gauche du contenu vidéo tridimensionnel, comme cela est décrit ci-dessus ;
- la reconnaissance de la présence d'une ou deux cartes de profondeur ou de disparité et du type de configuration de celles-ci ;
- s'il y a deux cartes de profondeur ou de disparité, l'obtention de chacune des
- 25 deux cartes ;
- la réalisation, sur les cartes de profondeur ou de disparité, d'opérations adaptées pour mettre les dimensions des cartes à des valeurs égales à celles des images du contenu vidéo. Ces opérations peuvent être, par exemple, d'un type inverse à un sous-échantillonnage, par exemple des opérations d'interpolation.
- 30 D'autres variantes peuvent concerner l'implémentation physique de l'invention. Par exemple, les modules électroniques qui implémentent les dispositifs décrits ci-dessus, en particulier le dispositif 100 et le récepteur 1100, peuvent être sous-divisés et distribués de façon variée ; en outre, ils peuvent être prévus sous

forme de modules matériels ou sous forme d'algorithmes logiciels implémentés par un processeur, en particulier un processeur vidéo équipé de zones de mémoire appropriées pour stocker temporairement les images entrées reçues. Ces modules peuvent donc exécuter en parallèle ou en séries une ou plusieurs des étapes de traitement vidéo des procédés de multiplexage et démultiplexage d'image selon la présente invention. Il est également évident que, bien que les modes de réalisation préférés fassent référence au multiplexage de deux flux vidéo 720p en un flux vidéo 1080p, d'autres formats peuvent être utilisés également.

Il est évident que l'ordre dans lequel les procédures de multiplexage et de démultiplexage représentées sur les figures 3 et 6 sont exécutées est purement illustratif : il peut être modifié pour une raison quelconque, sans changer l'essence du procédé.

L'invention n'est pas non plus limitée à un type particulier d'arrangement de l'image composite, car différentes solutions pour générer l'image composite peuvent offrir des avantages et/ou désavantages spécifiques.

L'invention, avec toutes ses variantes, propose un format universel pour générer, transporter et reproduire des contenus 3D sur un type quelconque d'affichage actuel ou futur.

Dans le cas d'un dispositif de reproduction 2D, le processeur vidéo du dispositif de reproduction éliminera simplement les images R et les cartes de profondeur (DM ou DM1 et DM2) qui peuvent être présentes à la sortie du récepteur 1100 et affichera, suivant la mise à échelle, seulement la séquence d'images L sur un dispositif de visualisation associé.

Ceci s'applique également au cas d'un dispositif de reproduction 3D dans lequel l'utilisateur a activé le mode d'affichage 2D.

Un dispositif de reproduction 3D dans lequel le mode d'affichage 3D a été activé peut présenter deux comportements différents, suivant le fait que la profondeur de la scène peut être réglée (réduite) ou non. Dans ce premier cas, le processeur vidéo utilisera les deux séquences d'images L et R pour générer l'effet tridimensionnel. Dans ce dernier cas, le processeur vidéo utilisera les cartes de profondeur (une ou deux) incluses dans les images composites C' associées à chaque paire d'images stéréoscopiques R et L pour générer des vues intermédiaires, entre L

et R, obtenant ainsi des images tridimensionnelles possédant une profondeur variable, inférieure à celle pouvant être obtenue à partir de L et R.

Ce dernier cas est représenté par des lecteurs auto-stéréoscopiques, qui nécessitent un nombre très important de vues (quelques dizaines) pour générer l'effet tridimensionnel pour des téléspectateurs positionnés à des points différents dans l'espace devant l'affichage. Dans ce cas, le processeur vidéo utilisera les cartes de profondeur (une ou deux) incluses dans les images composites C', conjointement avec les images L et R elles-mêmes, pour synthétiser une série d'autres images. Devant l'affichage, il y a un nombre de verres ou de barrières, de sorte que, à un point quelconque dans l'espace où la vision stéréoscopique est possible, le téléspectateur perçoive juste une paire desdites images.

Ainsi, le processeur vidéo du dispositif de reproduction peut comprendre des moyens adaptés pour envoyer, à l'affichage, deux séquences d'images, dont au moins une est constituée d'images synthétisées en commençant à partir d'au moins une des vues transmises et à partir d'au moins une carte de profondeur. Dans ce cas, il comprend également de préférence des moyens adaptés pour fournir au téléspectateur la possibilité de choisir des séquences d'images relatives à des points de vue plus ou moins proches, afin de varier la perception de profondeur.

Le processeur vidéo du dispositif de reproduction peut également comprendre des moyens adaptés pour générer des images supplémentaires correspondant à des vues supplémentaires, pour que les téléspectateurs positionnés à des points différents dans l'espace puissent voir différentes séquences d'images par l'intermédiaire d'un affichage auto-stéréoscopique associé.

Aucun des formats proposés jusqu'à présent n'offre une telle flexibilité et ampleur d'utilisation, tout en même temps, en garantissant toujours une très bonne qualité de reproduction en ce qui concerne l'équilibre des résolutions horizontale et verticale et d'attribution de résolution correcte aux images stéréoscopiques et aux cartes de profondeur associées.

Les opérations de reconstruction décrites ci-dessus peuvent avoir lieu partiellement dans le dispositif récepteur et partiellement dans le dispositif d'affichage.

La présente invention peut avantageusement être au moins partiellement réalisée par l'intermédiaire de programmes d'ordinateur comprenant des moyens de

codage pour implémenter une ou plusieurs étapes des procédés décrits ci-dessus, lorsque de tels programmes sont exécutés par un ordinateur. Il faut donc entendre que la portée de protection s'étend auxdits programmes d'ordinateur ainsi qu'à des moyens lisibles par ordinateur qui comprennent des messages enregistrés, lesdits

5 moyens lisibles par ordinateur comprenant des moyens de codage de programme pour implémenter une ou plusieurs étapes des procédés décrits ci-dessus, lorsque lesdits programmes sont exécutés par un ordinateur. L'exemple de mode de réalisation décrit ci-dessus peut être soumis à des variations sans s'éloigner de la portée de protection de la présente invention, y compris toutes les conceptions

10 équivalentes connues de l'homme du métier.

Les éléments et caractéristiques présentés dans les divers modes de réalisation préférés peuvent être associés les uns aux autres sans toutefois s'éloigner de la portée de protection de la présente invention.

À partir de la description ci-dessus, l'homme du métier pourra produire

15 l'objet de l'invention sans introduire de détails d'implémentation supplémentaires.

REVENDICATIONS

1. Procédé pour générer un flux vidéo stéréoscopique (101) comprenant des images composites (C), lesdites images composites (C) comprenant des informations sur une image droite (R) et une image gauche (L) d'un contenu tridimensionnel vidéo, dans lequel :
 - 5 des pixels de ladite image droite (R) et des pixels de ladite image gauche (L) sont sélectionnés, et
lesdits pixels sélectionnés sont entrés dans une image composite (C) dudit flux vidéo stéréoscopique,
le procédé étant caractérisé en ce que tous les pixels de ladite image droite (R) et tous
 - 10 les pixels de ladite image gauche (L) sont entrés dans ladite image composite (C) en laissant une desdites deux images inchangées, en divisant l'autre en un nombre de régions (R1, R2, R3) possédant une superficie totale égale à celle de ladite autre image (R), et en entrant lesdites régions dans ladite image composite (C),
dans lequel ladite image composite (C) possède des dimensions plus importantes que
 - 15 ce qui est nécessaire pour entrer tous les pixels de ladite image gauche (L) et de ladite image droite (R),
et dans lequel, dans les pixels de l'image composite (C) restant après ladite entrée, au moins une carte de profondeur ou de disparité (DM) est entrée qui est relative à la profondeur ou la disparité des pixels dudit contenu vidéo tridimensionnel, lesdites
 - 20 cartes visant à reconstruire, en réception, des images non transmises dans ledit flux stéréoscopique.
2. Procédé selon la revendication 1, dans lequel ladite au moins une carte de profondeur ou de disparité (DM) est codée sous forme d'image à niveaux de gris.
- 25 3. Procédé selon la revendication 2, dans lequel le contenu d'informations vidéo de ladite au moins une carte de profondeur ou de disparité est transporté par un seul signal de luminance, sans utiliser de signal de chrominance.
- 30 4. Procédé selon l'une quelconque des revendications précédentes, dans lequel ladite au moins une carte de profondeur ou de disparité (DM) possède une résolution

inférieure à celle d'une version d'origine de celle-ci, la résolution de cette dernière étant égale à celle desdites images gauche (L) et image droite (R).

5 5. Procédé selon la revendication 4, dans lequel ladite au moins une carte de profondeur ou de disparité (DM) est obtenue en soumettant ladite carte de profondeur d'origine à un sous-échantillonnage 4 à 1.

10 6. Procédé selon l'une quelconque des revendications précédentes, dans lequel ladite au moins une carte de profondeur ou de disparité (DM) est une carte de profondeur ou de disparité associée à l'une des deux images droite (R) ou gauche (L), ou associée à un point de vue intermédiaire, entre L et R.

15 7. Procédé selon l'une quelconque des revendications précédentes, dans lequel ladite au moins une carte de profondeur ou de disparité (DM) comprend une carte de profondeur ou de disparité associée à une image droite (R) et une carte de profondeur ou de disparité associée à une image gauche (L).

20 8. Procédé selon la revendication 7, dans lequel lesdites cartes de profondeur ou de disparité associées à une image droite (R) et une image gauche (L) sont entrées dans lesdits pixels restants de l'image composite (C) au moyen de techniques de mise en paquet d'images.

25 9. Procédé selon l'une quelconque des revendications précédentes, dans lequel, si ledit nombre de régions est trois, lesdites régions sont obtenues par l'intermédiaire des étapes de :

- division de ladite autre image (R) en deux parties possédant la même dimension horizontale (R1, R2, R3) ;
- division d'une desdites deux parties (R2, R3) en deux parties possédant la même dimension verticale (R2, R3).

30

10. Procédé selon l'une quelconque des revendications précédentes, dans lequel des informations de signalisation identifiant le type de flux vidéo généré sont entrées

sous forme de métadonnées dans ladite image composite (C) ou ledit flux vidéo stéréoscopique.

11. Procédé selon la revendication 10, dans lequel lesdites informations de signalisation sont adaptées afin de permettre de distinguer entre au moins deux des types suivant d'images composites :

- image composite du type à format mosaïque sans cartes de profondeur ;
- image composite du type à format mosaïque avec une carte de profondeur ;
- image composite du type à format mosaïque avec deux cartes de profondeur en configuration côte à côte ;
- image composite du type à format mosaïque avec deux cartes de profondeur en configuration haut et bas ;
- image composite du type à format mosaïque avec deux cartes de profondeur en configuration à format mosaïque.

15

12. Dispositif pour générer un flux vidéo stéréoscopique (101) comprenant des images composites (C), lesdites images composites (C) comprenant des informations sur une image droite (R), une image gauche (L), caractérisé en ce qu'il comprend des moyens pour implémenter les étapes du procédé selon l'une quelconque des revendications précédentes.

20

13. Procédé pour reconstruire au moins une paire d'images d'un flux vidéo stéréoscopique commençant à partir d'une image composite (C), ladite image composite (C) comprenant des informations sur une image droite (R), une image gauche (L), le procédé comprenant les étapes de :

25

- génération d'une première image desdites images droite (R) et gauche (L) en copiant un seul groupe de pixels contigus à partir d'une première région (C1) de ladite image composite,
- génération de l'autre image desdites images droite (R) et gauche (L) en copiant d'autres groupes de pixels contigus à partir d'un nombre de régions distinctes (C2, C3, C4) de ladite image composite, ledit nombre de régions distinctes étant différent de ladite première région (C1) ;

30

- génération d'au moins une carte de profondeur ou de disparité (DM) en copiant au moins un groupe de pixels contigus à partir d'une région supplémentaire (C5) de ladite image composite, différente de ladite première région et dudit nombre de régions distinctes.

5

14. Procédé selon la revendication 13, dans lequel, si ledit nombre de régions est trois :

- une (C2) desdites régions de l'image composite (C) possède la même dimension verticale que ladite première région (C1) et la moitié de sa dimension horizontale ;

10 - les deux autres (C3, C4) desdites régions de l'image composite (C) possèdent des dimensions horizontale et verticale égales, et la moitié de la dimension verticale de ladite première région (C1).

15. Procédé selon la revendication 13, dans lequel ladite au moins une carte de
15 profondeur ou de disparité (DM) est générée en commençant à partir d'une image à niveaux de gris dérivée d'un signal de luminance contenu dans les pixels contigus de ladite région supplémentaire (C5).

16. Procédé selon la revendication 15, comprenant une étape d'augmentation des
20 dimensions horizontale et verticale de ladite au moins une carte de profondeur ou de disparité (DM) jusqu'à une dimension égale à celle desdites images droite (R) et gauche (L).

17. Procédé selon la revendication 13, comprenant une étape d'obtention, à partir de
25 ladite image composite (C) ou à partir du flux vidéo, d'informations de signalisation adaptées pour reconnaître le type de flux vidéo généré.

18. Procédé selon la revendication 17, dans lequel lesdites informations de
30 signalisation sont adaptées afin de distinguer entre au moins deux des types suivants d'image composites :

- image composite du type à format mosaïque sans cartes de profondeur ;
- image composite du type à format mosaïque avec une carte de profondeur ;

- image composite du type à format mosaïque avec deux cartes de profondeur en configuration côte à côte ;

- image composite du type à format mosaïque avec deux cartes de profondeur en configuration haut et bas ;

- 5 - image composite du type à format mosaïque avec deux cartes de profondeur en configuration à format mosaïque.

19. Dispositif pour reconstruire au moins une paire d'images d'un flux vidéo stéréoscopique commençant à partir d'une image composite (C), ladite image composite (C) comprenant des informations sur une image droite (R), une image gauche (L), le dispositif comprenant :

- des moyens pour générer une première image desdites images droite (R) et gauche (L) en copiant un seul groupe de pixels contigus à partir d'une première région (C1) de ladite image composite,

- 15 - des moyens pour générer l'autre image desdites images droite (R) et gauche (L) en copiant d'autres groupes de pixels contigus à partir d'un nombre de régions distinctes (C2, C3, C4) de ladite image composite, ledit nombre de régions distinctes étant différent de ladite première région (C1) ;

- des moyens pour générer au moins une carte de profondeur ou de disparité (DM) en copiant au moins un groupe de pixels contigus à partir d'une région supplémentaire (C5) de ladite image composite, différent de ladite première région et dudit nombre de régions distinctes.

20. Dispositif selon la revendication 19, dans lequel, si ledit nombre de régions est trois :

- une (C2) desdites régions de l'image composite (C) possède la même dimension verticale que ladite première région (C1) et la moitié de sa dimension horizontale ;

- les deux autres (C3, C4) desdites régions de l'image composite (C) possèdent des dimensions horizontale et verticale égales, et la moitié de la dimension verticale de ladite première région (C1).

21. Dispositif selon la revendication 19, dans lequel lesdits moyens pour générer au moins une carte de profondeur ou de disparité (DM) utilisent une image à niveaux de

gris dérivée d'un signal de luminance contenu dans les pixels contigus de ladite région supplémentaire (C5).

22. Dispositif selon la revendication 21, comprenant des moyens pour augmenter les dimensions horizontale et verticale de ladite au moins une carte de profondeur ou de disparité (DM) jusqu'à une dimension égale à celle desdites images droite (R) et gauche (L).

23. Dispositif selon la revendication 19, comprenant des moyens adaptés pour reconnaître le type de flux vidéo reçu en fonction d'informations de signalisation identifiant ledit type de flux, contenues dans ladite image composite (C) ou dans ledit flux vidéo.

24. Dispositif selon la revendication 23, dans lequel lesdites informations de signalisation permettent de distinguer entre au moins deux des types suivants d'image composites :

- image composite du type à format mosaïque sans cartes de profondeur ;
- image composite du type à format mosaïque avec une carte de profondeur ;
- image composite du type à format mosaïque avec deux cartes de profondeur en configuration côte à côte ;
- image composite du type à format mosaïque avec deux cartes de profondeur en configuration haut et bas ;
- image composite du type à format mosaïque avec deux cartes de profondeur en configuration à format mosaïque.

25

25. Dispositif selon la revendication 24, comprenant des moyens qui, en fonction desdites informations utiles pour distinguer un type d'image composite, sont adaptés pour produire :

- seulement ladite première image desdites images droite (R) et gauche (L) ; ou
- ladite première et ladite seconde desdites images droite (R) et gauche (L) ; ou
- ladite première et ladite seconde desdites images droite (R) et gauche (L) et ladite au moins une carte de profondeur ou de disparité (DM).

26. Dispositif selon la revendication 23, comprenant des moyens adaptés pour réaliser une ou plusieurs des opérations suivantes, en fonction desdites informations de signalisation :

- la reconnaissance du type d'image reçu, dans le but de reconstruire correctement les
5 deux images droite et gauche du contenu vidéo tridimensionnel ;
- la reconnaissance de la présence d'une ou deux cartes de profondeur ou de disparité et du type de configuration de celles-ci ;
- dans le cas de deux cartes de profondeur ou de disparité, l'obtention de chacune des deux cartes ;
- 10 - la réalisation, sur lesdites cartes de profondeur ou de disparité, d'opérations adaptées pour mettre les dimensions des cartes à des valeurs égales à celles des images du contenu vidéo.

27. Dispositif selon la revendication 19, comprenant des moyens conçus afin de
15 générer des images supplémentaires correspondant à des vues supplémentaires en commençant à partir desdites images droite (R) et gauche (L) et en utilisant lesdites cartes de profondeur.

28. Dispositif selon la revendication 27, comprenant des moyens adaptés pour
20 afficher deux séquences d'images, au moins l'une d'entre elles comprenant des images synthétisées en commençant à partir d'au moins une des vues transmise et à partir d'au moins une carte de profondeur.

29. Dispositif selon la revendication 28, comprenant des moyens adaptés pour fournir
25 au téléspectateur la possibilité de choisir des séquences d'images concernant des points de vue plus ou moins proches, afin de varier la perception de profondeur.

30. Dispositif selon la revendication 27, comprenant un affichage auto-
stéréoscopique, et comprenant des moyens adaptés pour utiliser lesdites images
30 supplémentaires correspondant à des vues supplémentaires afin de permettre à des téléspectateurs positionnés à des points différents dans l'espace de voir des séquences d'images différentes.

31. Flux vidéo stéréoscopique (1101) caractérisé en ce qu'il comprend au moins une image composite (C) générée au moyen du procédé selon l'une quelconque des revendications 1 à 11.

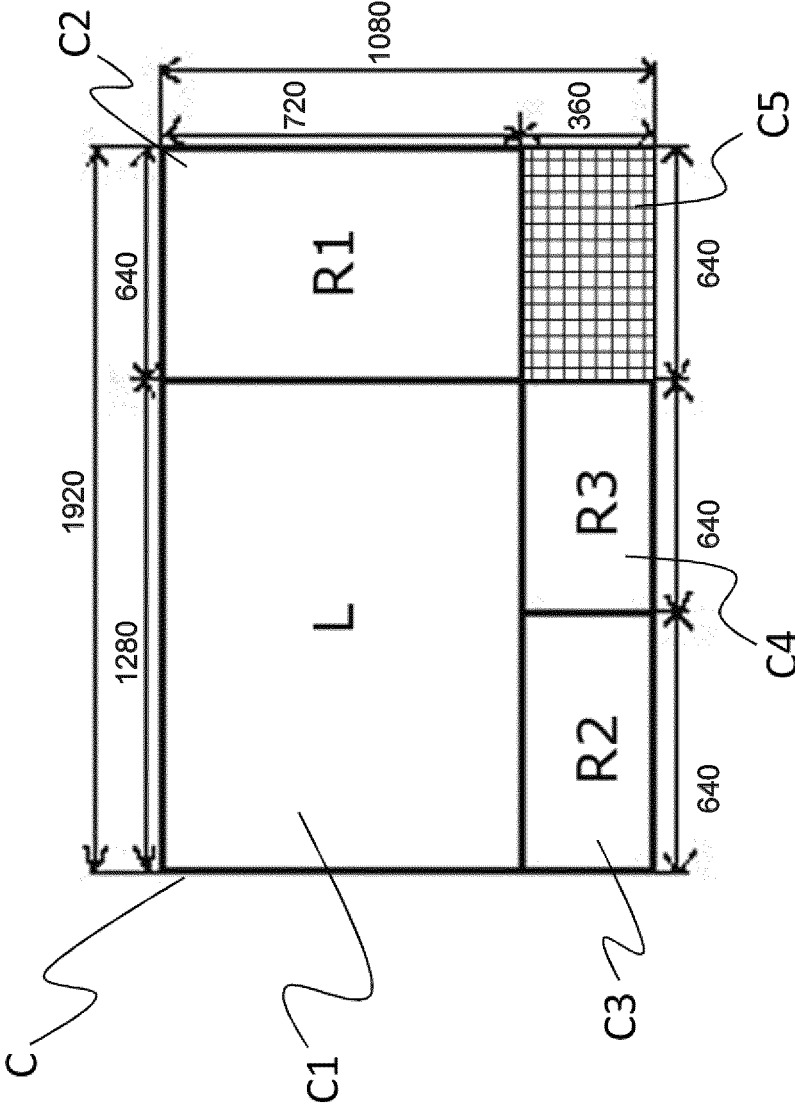


Fig. 1a

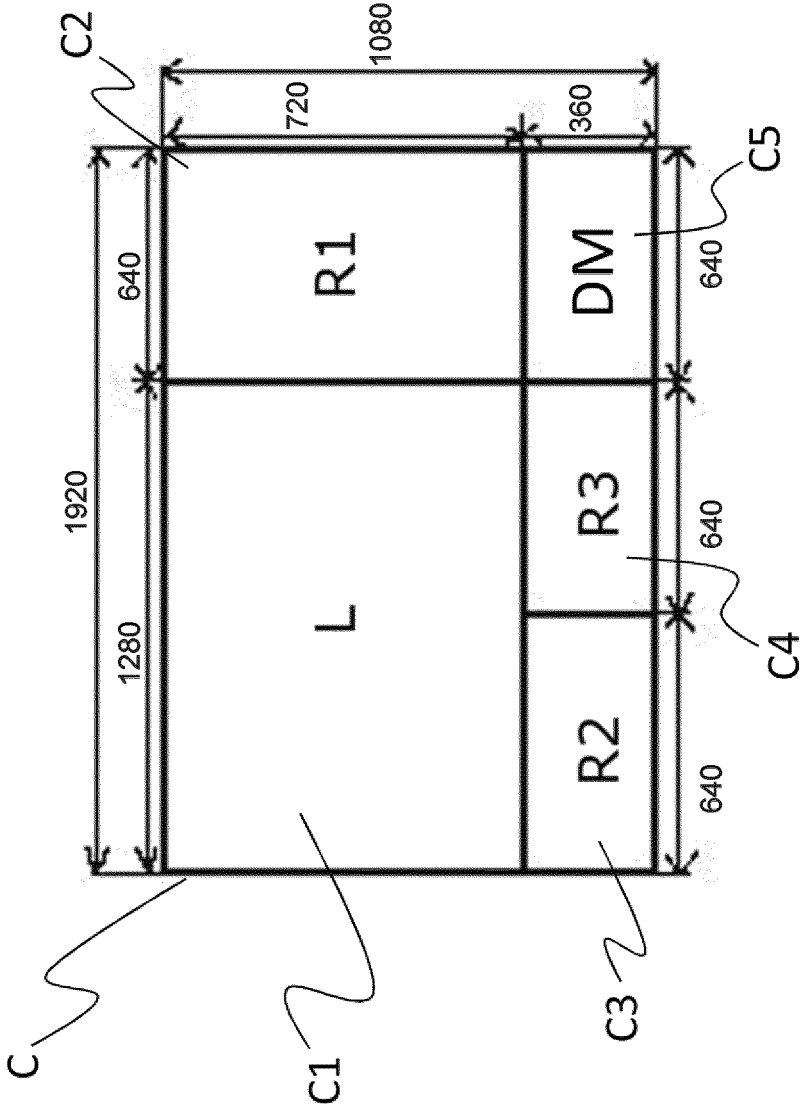


Fig. 1b

3/7

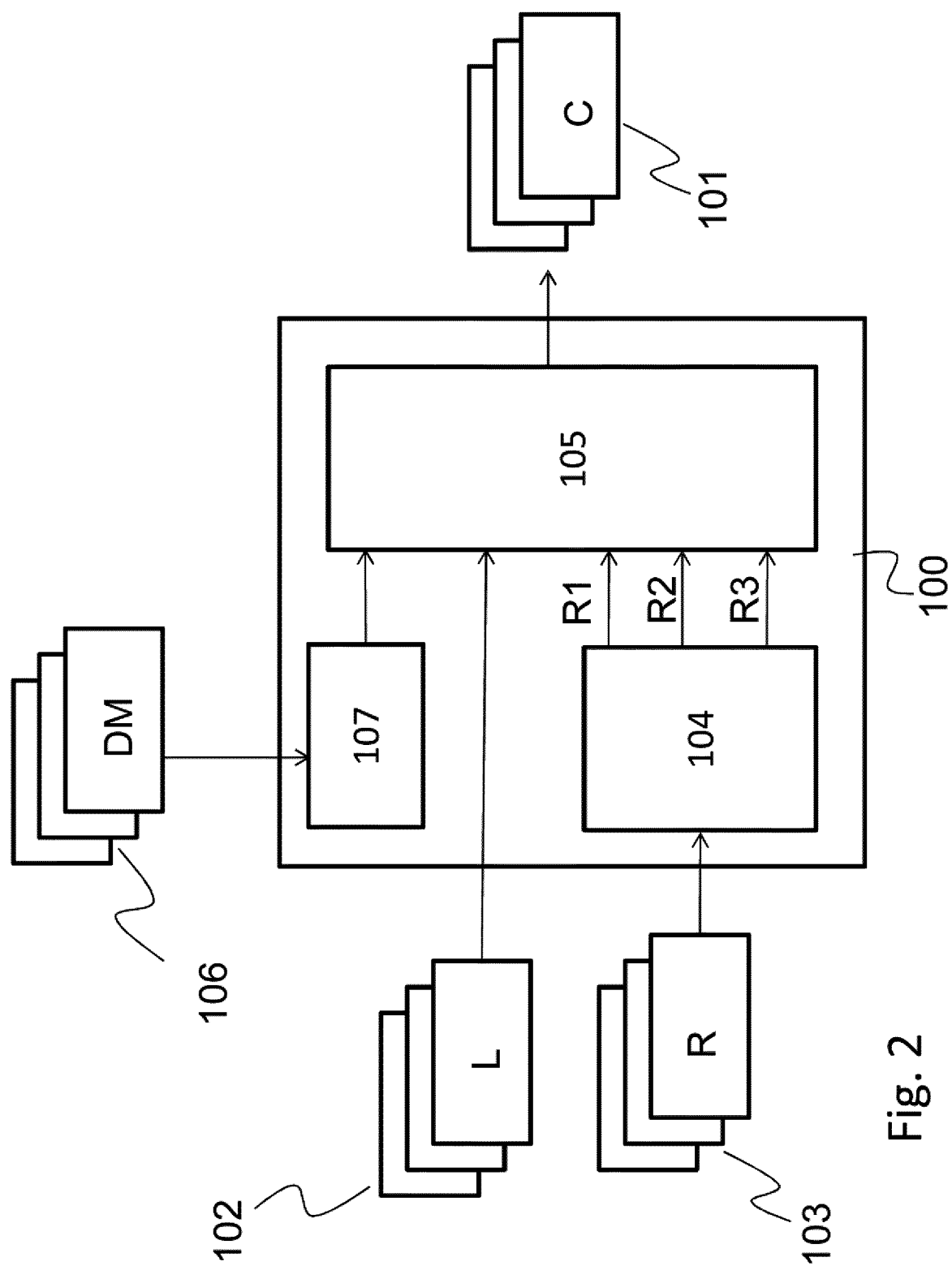


Fig. 2

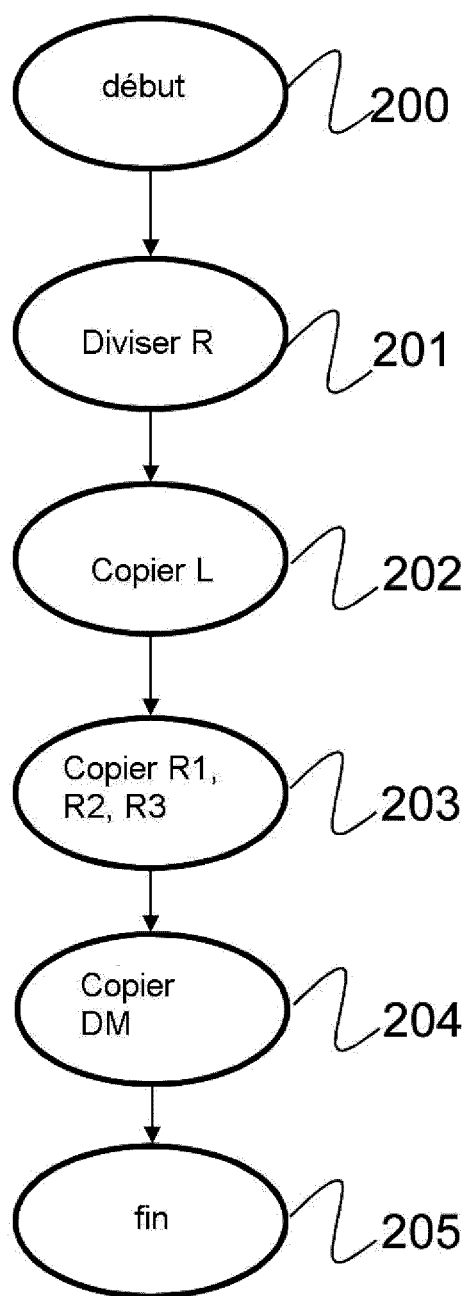


Fig. 3

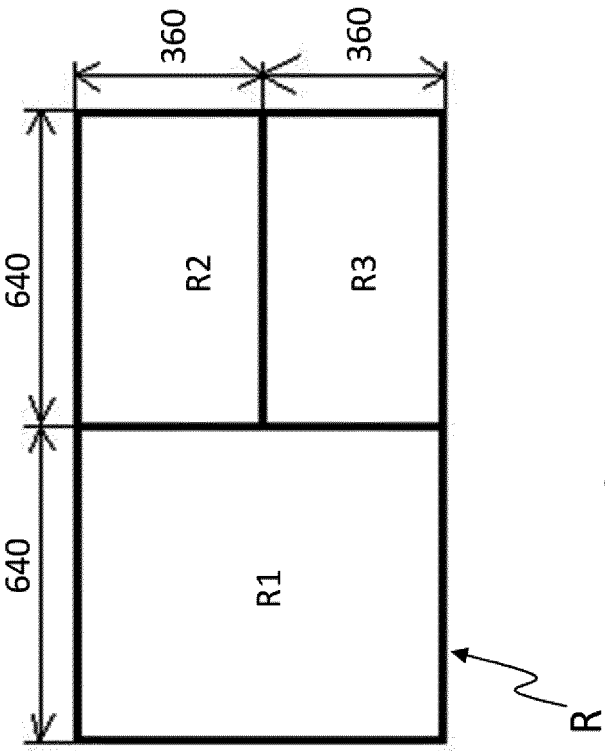


Fig. 4

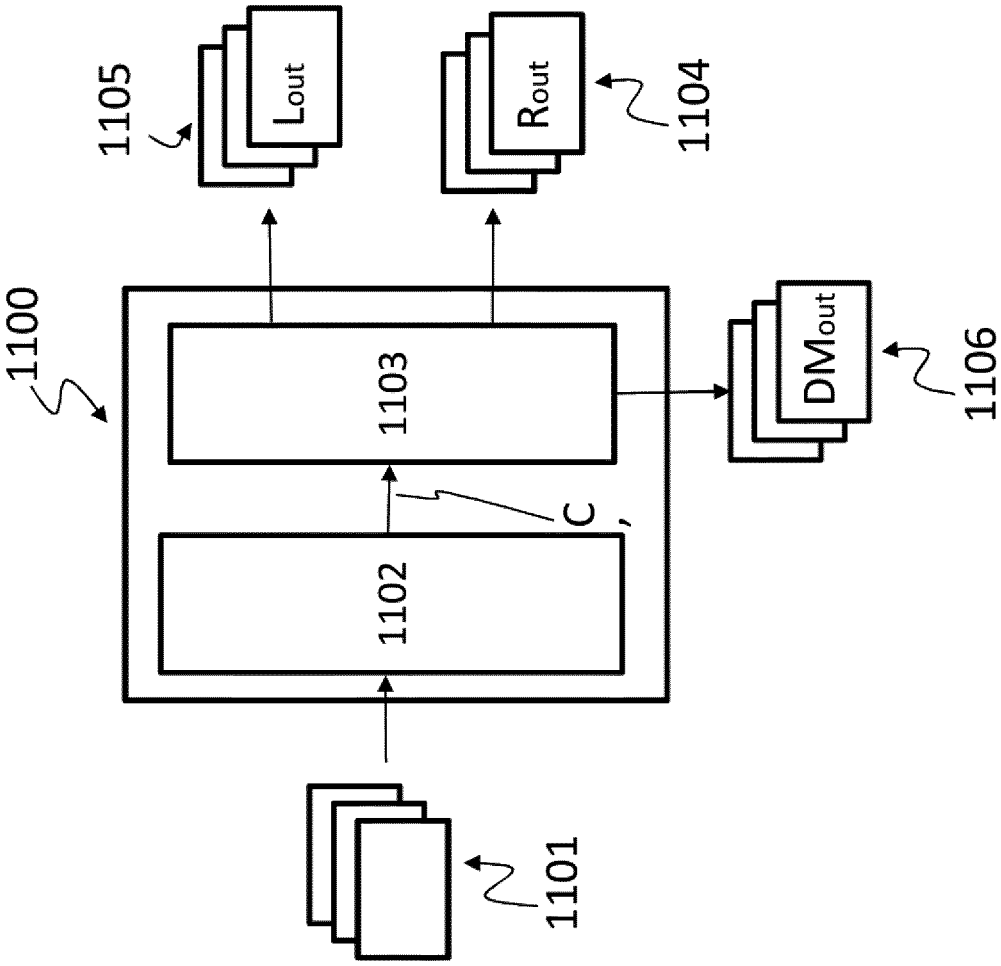


Fig. 5

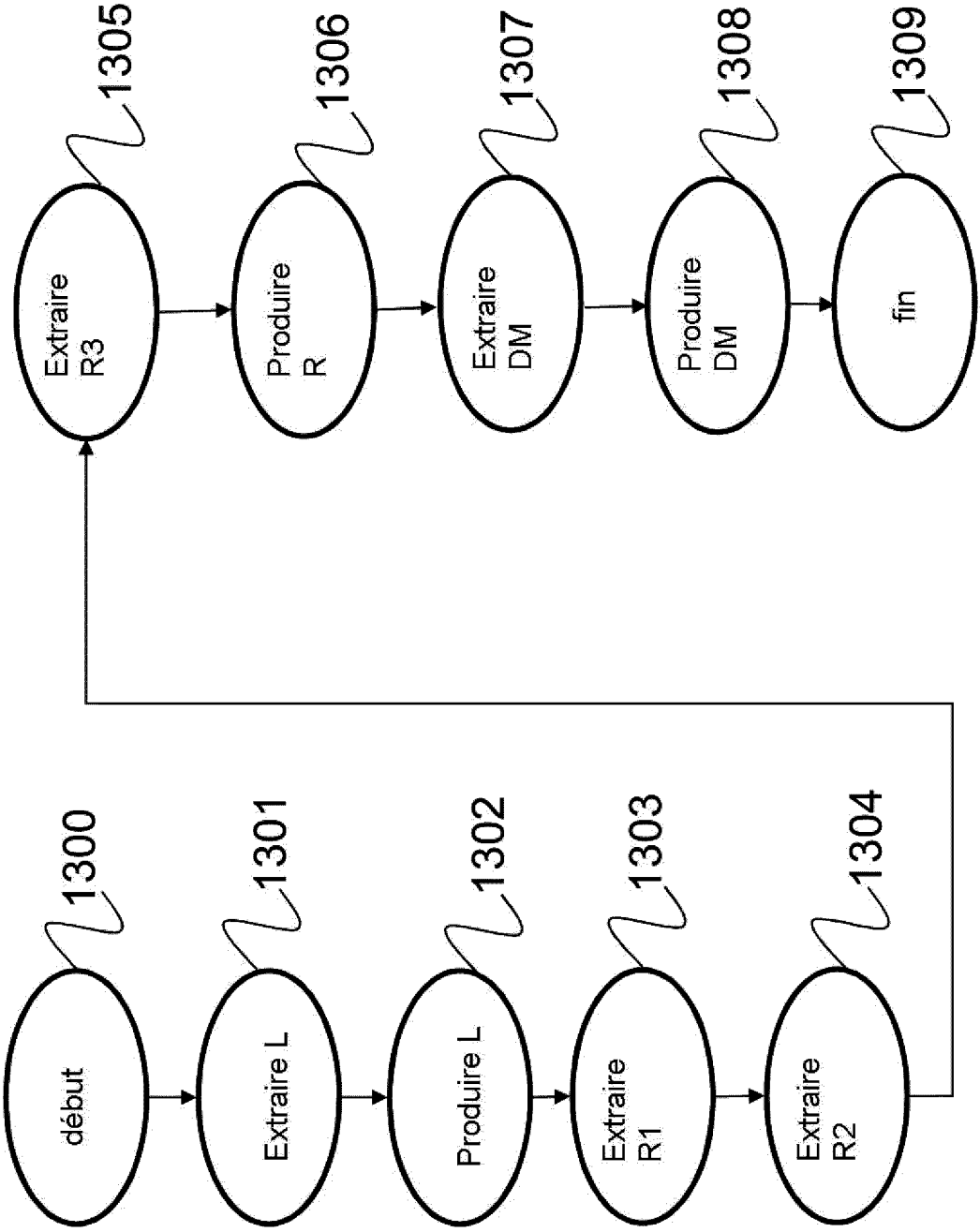


Fig. 6