

⑫

BREVET D'INVENTION

B1

⑤④ procédé et dispositif de capture vidéo, procédé et dispositif de composition d'une vidéo et procédé d'analyse de vidéo.

②② Date de dépôt : 05.01.22.

③⑦ Priorité :

⑥⑦ Références à d'autres documents nationaux
apparentés :

☐ Demande(s) d'extension :

⑦① Demandeur(s) : *ROKNI Tomaj* — FR.

④③ Date de mise à la disposition du public
de la demande : 07.07.23 Bulletin 23/27.

④⑤ Date de la mise à disposition du public du
brevet d'invention : 04.10.24 Bulletin 24/40.

⑤⑥ Liste des documents cités dans le rapport de
recherche :

Se reporter à la fin du présent fascicule

⑦② Inventeur(s) : *ROKNI Tomaj*.

⑦③ Titulaire(s) : *ROKNI Tomaj*.

⑦④ Mandataire(s) : *CABINET BEAU DE LOMENIE*.



Description

Titre de l'invention : procédé et dispositif de capture vidéo, procédé et dispositif de composition d'une vidéo et procédé d'analyse de vidéo.

Domaine technique

[0001] L'invention concerne le traitement d'une vidéo d'une scène à partir de plusieurs vidéos représentatives d'une partie de la scène.

Technique antérieure

[0002] La diffusion et la captation d'événements, tels que par exemple, mais non limitativement, des concerts, des événements sportifs requiert bien souvent le déploiement de moyens coûteux et complexes à mettre en œuvre. En effet, il est bien souvent nécessaire de positionner de nombreuses caméras permettant la capture de l'événement, les caméras étant positionnées généralement tout autour de la scène à filmer. Non seulement la captation de la scène est rendue complexe par le positionnement des nombreuses caméras, mais également la composition et la restitution de la scène par les diffuseurs.

[0003] La complexité de tels systèmes peut amener les spectateurs à utiliser des dispositifs plus simples à manipuler tels par exemple, des téléphones cellulaires pour la captation. Cependant la recomposition d'une vidéo à partir de téléphones positionnés autour de la scène reste complexe notamment en raison de la mobilité de tels téléphones et donc des positions respectives des téléphones vis à vis de la scène.

Exposé de l'invention

[0004] La présente invention a pour but de remédier à tout ou partie des inconvénients de l'art antérieur, notamment ceux exposés ci-avant, en proposant une solution qui permette, en comparaison avec les solutions de l'état de la technique, de recomposer une vidéo unique à partir de plusieurs vidéos d'une même scène tout en évitant les solutions complexes de gestion de la mobilité des terminaux de capture et l'intervention d'un utilisateur.

[0005] A cet effet, l'invention concerne un procédé de composition d'une vidéo comprenant:

[0006] - la réception d'un au moins un premier dispositif de capture vidéo d'une première vidéo représentative d'une première partie d'une scène et de métadonnées associées à au moins une image fixe de la première partie de ladite scène, la métadonnée étant représentative d'une position géographique du premier dispositif de capture au moment de la capture de l'image,

[0007] - la réception d'un au moins un second dispositif de capture vidéo d'une deuxième vidéo représentative d'au moins une deuxième partie de ladite scène et de métadonnées

associées à au moins une image fixe de la seconde partie de la scène, la métadonnée étant représentative de la position géographique du second dispositif de capture au moment de la capture de l'image, les premières et deuxième vidéos étant synchronisées,

- [0008] - la détermination de points d'intérêt fixes dans chacune des vidéos,
- [0009] - la composition d'une vidéo composite représentative de ladite scène à partir des première et au moins une deuxième vidéos par mise en correspondance desdits points d'intérêt,
- [0010] - la mise à disposition de la vidéo représentative de la scène,
- [0011] la détection d'un mouvement d'au moins un des deux dispositifs de capture sur la base des métadonnées reçues déclenchant une nouvelle détermination des points d'intérêt fixes dans la vidéo capturée par le dispositif en mouvement puis la composition d'une nouvelle vidéo composite.
- [0012] Ainsi, le procédé peut avantageusement détecter que les dispositifs de capture ont bougé, ou au moins l'un d'entre eux et automatiquement recalculer les points d'intérêt. De cette manière, la composition des deux vidéos génère une vidéo de qualité constante à partir des deux vidéos. Aucune intervention humaine n'est nécessaire ni pour repositionner les dispositifs de capture ni pour demander un nouveau calcul des points d'intérêt tant que le procédé peut calculer ou recalculer les points d'intérêt ou tant que la variation des métadonnées reste inférieure à un seuil.
- [0013] Dans des modes particuliers de mise en œuvre, la détermination de points d'intérêt fixes dans chacune des vidéos comprend :
- [0014] - la transmission à au moins un dispositif d'affichage vidéo, des première et au moins deuxième vidéos,
- [0015] - la réception du au moins un dispositif d'affichage vidéo, des points d'intérêt relatifs à chacune des parties de la scène.
- [0016] Dans des modes particuliers de mise en œuvre, la mise à disposition comprend l'un ou l'autre ou plusieurs parmi :
- [0017] - l'enregistrement de la vidéo composite,
- [0018] - l'affichage de la vidéo composite,
- [0019] - la transmission de la vidéo composite à au moins un dispositif d'analyse.,
- [0020] De manière avantageuse, la vidéo recomposée est destinée à être diffusée, par exemple à un groupe de diffusion. Elle peut être diffusée en direct, le procédé étant rapide et ne provoquant qu'un très léger décalage avec l'événement. La vidéo peut aussi être enregistrée et diffusée en décalé, à la demande.
- [0021] Dans des modes particuliers de mise en œuvre, lorsque ladite vidéo est transmise à au moins un dispositif d'analyse, il comprend en outre
- [0022] - la réception, suite à la transmission de données d'enrichissement de la scène,

- [0023] - la synchronisation des données d'enrichissement avec la vidéo composite ;
- [0024] - la mise à disposition de la vidéo composite enrichie des données d'enrichissement.
- [0025] Ainsi, outre la recomposition automatique de la vidéo à partir des deux vidéos, le procédé peut avantageusement proposer à un utilisateur final un contenu enrichi de données additionnelles relatives au contenu capturé.
- [0026] Selon un troisième aspect, l'invention concerne un procédé d'analyse de vidéo comprenant
- [0027] - La réception, d'une vidéo composite représentative d'une scène obtenue par un procédé de composition d'une vidéo selon l'invention,
- [0028] - L'ajout de données d'enrichissement sur ladite vidéo composite pour annoter ladite vidéo composite,
- [0029] - La mise à disposition de la vidéo composite annotée et des données d'enrichissement.
- [0030] De manière préférée, les données d'enrichissement sont obtenues par un positionnement des données d'enrichissement par un utilisateur à travers une interface graphique.
- [0031] Selon un autre mode de réalisation, le procédé de composition analyse la vidéo composite (par exemple en utilisant de l'intelligence artificielle) et compare les données d'enrichissement reçues avec les données analysées et en cas d'incohérence, soumet lesdites incohérences à l'utilisateur au travers de l'interface graphique pour validation finale.
- [0032] L'invention concerne également un dispositif de capture vidéo comprenant des moyens de capture vidéo, une interface de communication avec au moins un serveur et au moins un processeur apte à mettre en œuvre un procédé de capture vidéo selon l'invention.
- [0033] L'invention concerne également un serveur comprenant
- [0034] - au moins une première interface de communication avec un dispositif de capture vidéo,
- [0035] - au moins une seconde interface de communication avec un dispositif d'affichage vidéo,
- [0036] - au moins un processeur apte à mettre en œuvre un procédé de composition d'une vidéo selon l'invention.
- [0037] L'invention concerne également un système comprenant au moins un premier et un second dispositif de capture vidéo selon l'invention, au moins un serveur selon l'invention et au moins un dispositif d'analyse apte à mettre en œuvre un procédé d'analyse selon l'invention.
- [0038] L'invention concerne également un programme d'ordinateur comportant des instructions pour la mise en œuvre d'un des procédés selon l'invention lorsque le

programme est exécuté par un ordinateur.

- [0039] L'invention concerne également un support d'enregistrement lisible par un ordinateur sur lequel est enregistré un programme d'ordinateur selon l'invention.

Brève description des dessins

- [0040] [Fig.1] La [Fig.1] représente un procédé de capture vidéo selon un mode de réalisation particulier de l'invention.
- [0041] [Fig.2] La [Fig.2] représente un procédé de composition d'une vidéo selon un mode de réalisation particulier de l'invention.
- [0042] [Fig.3] La [Fig.3] représente la position des points d'intérêt selon un mode particulier de réalisation de l'invention.
- [0043] [Fig.4] La [Fig.4] représente un procédé d'analyse de vidéo selon un mode de réalisation particulier de l'invention.
- [0044] [Fig.5] La [Fig.5] représente un dispositif de capture vidéo selon un mode de réalisation particulier de l'invention.
- [0045] [Fig.6] La [Fig.6] représente un serveur de composition d'une vidéo selon un mode de réalisation particulier de l'invention.
- [0046] [Fig.7] La [Fig.7] représente un dispositif d'analyse de vidéo selon un mode de réalisation particulier de l'invention.
- [0047] [Fig.8] La [Fig.8] représente un système selon un premier mode de réalisation particulier de l'invention.
- [0048] [Fig.9] La [Fig.9] représente la position de deux caméras autour d'une scène à capturer selon un mode de réalisation particulier de l'invention.
- [0049] [Fig.10] La [Fig.10] la position de trois caméras autour d'une scène à capturer selon un mode de réalisation particulier de l'invention.
- [0050] [Fig.11] La [Fig.11] représente un système selon un second mode de réalisation particulier de l'invention.

Description des modes de réalisation

- [0051] Dans les modes de réalisation décrits ci-après, la scène capturée est une scène représentative d'un match de football et donc plus précisément des événements se passant sur un terrain de football lors d'un match. Ainsi, lorsqu'il est fait référence à une première partie de la scène et à une seconde partie de la scène, il est fait particulièrement référence à une première partie du terrain de football, par exemple un peu plus qu'un demi-terrain pour une première partie et un peu plus qu'un demi terrain pour une seconde partie. Bien entendu, ceci est un exemple de scène capturée et ne saurait aucunement constituer une limitation de l'invention à la captation d'un match de football. De même, d'autres sports peuvent être envisagés mais également d'autres événements tels des concerts ou autres scènes de la vie.

- [0052] Par ailleurs, la description des figures 1 à 9 fait référence principalement à des modes de réalisation basés sur la capture de deux parties de la même scène. Il est à noter que le fait de parler de deux parties de la scène n'est pas limitatif à deux et que l'invention ne se limite pas à deux captures mais à au moins deux. La [Fig.10] illustre par exemple le positionnement possible de trois moyens de capture.
- [0053] La [Fig.1] représente sous forme d'ordinogramme, les principales étapes d'un procédé de capture vidéo selon un mode de réalisation préféré de l'invention. Ledit procédé comprend des étapes E1a et E1b lors desquelles deux vidéos sont capturées. Plus précisément, lors de l'étape E1a, une vidéo d'une partie de ladite scène est capturée et lors de l'étape E1b, une vidéo d'au moins une autre partie de la dite scène est capturée. Les deux étapes E1a et E1b sont effectuées de manière simultanée. Les deux parties capturées contiennent au moins une portion commune de la scène, cette portion commune comprenant au moins deux points d'intérêt fixes communs.
- [0054] Les deux vidéos sont capturées de manière synchrone ou peuvent être synchronisées grâce à des étiquettes temporelles. Pour ce faire, préalablement au démarrage du procédé de capture, un procédé de synchronisation (non représenté sur les figures) est lancé de manière à associer à chaque vidéo capturée une étiquette temporelle (« timecode » en anglais). Ceci est par exemple avantageusement réalisé en utilisant le protocole NTP (« network Time Protocol » en anglais). Chacune des images capturées ayant son étiquette temporelle, il est possible de synchroniser les deux vidéos lors de leur réception par un dispositif distant, peu importe les délais de transmission de chacune des vidéos qui pourraient être causés par des perturbations réseau ou des vitesses de transmission différentes selon le réseau utilisé par chacun des dispositifs de capture.
- [0055] Dans d'autres modes de réalisation, la synchronisation des deux flux peut se faire grâce aux pistes sonores des flux. La fréquence d'échantillonnage du son étant beaucoup plus élevée que la fréquence des images, et les pistes sonores étant pratiquement identiques sur tous les flux car les micros sont très proches, caler les pistes sonores permet également de caler les étiquettes temporelles, donc les flux. Ce système est utilisé pour la synchronisation et le montage des captations multi caméras tout en évitant les effets d'échos.
- [0056] Ensuite, lors d'une étape E2a, E2b, des métadonnées sont obtenues, préférentiellement à partir de données du capteur GPS (pour « Global Positioning System » en anglais ou système de positionnement global en français) et/ou de données d'un accéléromètre et/ou d'un gyroscope et/ou d'un magnétomètre. Chaque vidéo d'une partie de la scène est capturée par des moyens de capture et ces moyens peuvent par exemple être associés ou collaborer avec un GPS, un accéléromètre, un gyroscope ou un magnétomètre.

- [0057] Selon un autre mode de réalisation, un processeur associé aux moyens de capture peut également, à l'aide d'un programme logiciel, obtenir lesdites métadonnées.
- [0058] Ainsi, lors de la capture de chaque partie de la scène, il est possible d'associer à chaque image capturée les métadonnées. Chaque image est donc associée à une étiquette temporelle et à des métadonnées. Bien entendu, certaines métadonnées varient peu telles la latitude et la longitude alors qu'en cas d'intempéries, certaines métadonnées vont varier plus, telles l'assiette, la boussole ou l'inclinaison. Ainsi, la longitude et la latitude peuvent être associées à la vidéo et transmises moins fréquemment que d'autres qui pourraient varier davantage au cours de la capture.
- [0059] Les métadonnées sont principalement relatives à la position géographique, pouvant inclure également l'orientation, d'un dispositif de capture mettant en œuvre le procédé et par exemple décrit en [Fig.4]. Ainsi, un changement dans les métadonnées permet d'identifier un déplacement ou un mouvement de la caméra. Un tel déplacement/mouvement, n'est pas nécessairement mais peut être, un déplacement opéré, volontairement ou involontairement, par une personne. Un tel déplacement/mouvement peut également être lié à la météorologie par exemple. S'il y a beaucoup de vent, il est possible que le dispositif de capture mettant en œuvre le procédé dévie de sa position initiale. Ceci est d'autant plus vrai si le dispositif de capture est attaché en hauteur (pour avoir une meilleure vision de la scène à capturer). De préférence, pour éviter des modifications importantes des métadonnées, les moyens de capture sont fixés de manière rigide et ancrés au sol. Par exemple, ils peuvent être fixés sur un ou plusieurs trépieds, démontables ou fixés au sol.
- [0060] Lors d'une étape E3a, E3b, chaque flux vidéo (comprenant les étiquettes temporelles) et les métadonnées associées à la capture sont transmis vers un serveur, représenté en [Fig.5].
- [0061] Selon un mode de réalisation préféré, les métadonnées sont transmises de manière périodique vers le serveur. De manière périodique signifie à des instants régulièrement espacés dans le temps ou à des instants irrégulièrement espacés dans le temps. Par exemple, il peut être envisagé que les métadonnées soient transmises toutes les 60s. La fréquence d'envoi des métadonnées peut également être programmée et varier selon les conditions météorologiques par exemple si les dispositifs de capture sont situés en extérieur. Lorsqu'il y a beaucoup de vent par exemple, les moyens de capture peuvent éventuellement bouger, selon leur localisation ou leur fixation au sol et ainsi il peut être choisi de transmettre les métadonnées avec une fréquence plus grande que lorsque le temps est calme. Chacun des dispositifs de capture peut également transmettre les métadonnées selon une fréquence différente puisque chacun des dispositifs de capture peut être sujet à des variations différentes selon son emplacement ou les conditions propres à son emplacement.

- [0062] La transmission peut se faire en utilisant le réseau de communication sans fil, comme le réseau cellulaire 2G, 3G, 4G, 5G et plus ou alors en utilisant une connexion WIFI. Chacun des dispositifs de capture peut également utiliser un réseau différent pour transmettre la vidéo et les métadonnées.
- [0063] Les images des flux vidéo sont préférentiellement transmises à un débit de 30 images par seconde. Dans d'autres modes de réalisation, elles peuvent être transmises à un rythme de 60, 120 ou 240 images par seconde. L'évolution des réseaux de communication et des standards vidéo peut permettre la transmission de vidéos de taille supérieure, par exemple Full HD, 4K, 8K en utilisant des réseaux de transmission comme la 5G et versions ultérieures, ou les réseaux WIFI de normes futures.
- [0064] Bien entendu, la vidéo peut être transmise de manière encodée (compressée) avant transmission en utilisant des algorithmes de codage tels MPEG-4, H.264 ou X.264 ou d'autres formats...
- [0065] La [Fig.2] représente un procédé de composition d'une vidéo composite selon un mode de réalisation particulier de l'invention. Le procédé est avantageusement mis en œuvre dans un serveur tel que représenté en [Fig.5].
- [0066] Lors d'une étape T1, la vidéo (comprenant les étiquettes temporelles) et les métadonnées représentatives d'une première partie de la scène, et d'au moins une seconde partie de la scène, sont reçues. Les vidéos ont été capturées avec une même étiquette temporelle, elles peuvent donc être synchronisées à l'aide de ces étiquettes temporelles. De même, les métadonnées étant associées à une image, chaque image étant associée à une étiquette temporelle, les métadonnées sont donc associées à la même étiquette que l'image à laquelle elles sont associées.
- [0067] Lors d'une étape T2, on détermine dans chacune des deux vidéos des points d'intérêt fixes dans la scène. La composition de la vidéo composite représentative de ladite scène est réalisée à partir des première et au moins une deuxième vidéos par mise en correspondance des points d'intérêt.
- [0068] Lors de cette étape la détermination des points d'intérêt est liée à la nature de la scène capturée. Par exemple, lorsque la scène capturée est un match de football, 4 points d'intérêt sont déterminées sur chaque partie de la scène.
- [0069] Selon un premier mode de réalisation, comme illustré en [Fig.3], les 4 points d'intérêt sont définis comme étant les points d'intersection entre
- [0070] - La ligne médiane et les deux lignes de touche :
- [0071] - les points Rlu et Rld pour la première partie de la vidéo représentant la partie droite de la scène en [Fig.3] droite et les points Lru et Lrd pour la seconde partie de la vidéo en [Fig.3] gauche représentant la partie gauche de la scène
- [0072] - La ligne de but et les deux lignes de touche :
- [0073] les points Rru et Rrd pour la première partie de la vidéo représentant la partie droite

de la scène en [Fig.3] droite et les points Llu et Lld pour la seconde partie de la vidéo en [Fig.3] gauche représentant la partie gauche de la scène

- [0074] Le procédé de composition et particulièrement la détection des points d'intérêt sont liés au type de scène filmée. Ainsi, le procédé de composition pourra être paramétré au préalable avec le type d'événement capturé.
- [0075] Selon un mode de réalisation préféré, les points d'intrêts sont déterminés en utilisant différentes fonctions et algorithmes de la librairie open-cv. Tout d'abord les masques de couleur (vert pour le terrain et blanc pour les lignes). La détection de contours peut être effectuée par le filtre de Canny (algorithme de canny), par la méthode des gradients ou du laplacien. Afin d'éliminer les faux-positifs restant, le procédé met en œuvre les librairies open-cv, numpy, et skimage.morphologie dans python (ou équivalent en langage C++) ou l'algorithme watershed (ou ligne de partages des eaux). La fonction morpho est calibrée pour ne garder que les éléments en ligne droite les plus longs (éliminant ainsi les buts comprenant les poteaux et la barre transversale, le rond central ou la surface de réparation pour l'exemple du football).
- [0076] Selon un mode particulier de réalisation la détermination, à l'étape T2, de points d'intérêt fixes dans chacune desdites vidéos peut comprendre :
- [0077] - la transmission à au moins un dispositif d'analyse vidéo, des au moins deux vidéos reçues représentatives chacune d'au moins une partie de la scène,
- [0078] - La réception du au moins un dispositif d'analyse vidéo, des points d'intérêt relatifs à chacune des parties de ladite scène.
- [0079] Le dispositif d'analyse vidéo est représenté en [Fig.7]
- [0080] Ceci peut éventuellement permettre la détermination des points d'intérêt par un autre dispositif et éventuellement permettre si besoin d'ajuster les points d'intérêt dans le cas où la détection automatique ne pourrait avoir lieu, par exemple par l'intervention d'un utilisateur.
- [0081] Par exemple, si seulement trois points d'intérêt peuvent être déterminés lors de l'étape T2 par le procédé, alors l'utilisateur peut être amené à positionner un quatrième point d'intérêt pour améliorer le procédé de composition, plus fiable avec quatre points d'intérêt qu'avec trois points d'intérêt.
- [0082] Selon un mode particulier de réalisation la détermination, à l'étape T2, de points d'intérêt fixes dans chacune desdites vidéos peut comprendre :
- [0083] - la transmission à au moins un dispositif d'analyse vidéo, des au moins deux vidéos reçues représentatives chacune d'au moins une partie de la scène, et des points d'intérêt,
- [0084] - La réception du au moins un dispositif d'analyse vidéo, de nouveaux points d'intérêt relatifs à chacune des parties de ladite scène.
- [0085] Ainsi, ceci peut avantageusement ajuster un ou plusieurs points d'intérêt lorsque la

détection automatique ne fonctionne pas correctement ou de manière suffisamment précise.

- [0086] Selon un autre mode particulier de réalisation la détermination, à l'étape T2, de points d'intérêt fixes dans chacune desdites vidéos peut comprendre :
- [0087] - la transmission à au moins un dispositif d'analyse vidéo, des au moins deux vidéos reçues représentatives chacune d'au moins une partie de la scène, et des points d'intérêt,
- [0088] - La réception du au moins un dispositif d'analyse vidéo, d'un message de validation de la position de au moins un des points d'intérêt.
- [0089] Ceci peut avantageusement permettre à un utilisateur ou un opérateur de confirmer ou non la détection automatique de la position des points d'intérêt.
- [0090] Selon un autre mode de réalisation, lors de la transmission à au moins un dispositif d'analyse vidéo, des au moins deux vidéos reçues représentatives chacune d'au moins une partie de la scène, les deux vidéos sont transmises à la même résolution (définition) et plus particulièrement à la résolution la plus faible de celle des deux vidéos lorsque les deux vidéos ne sont pas capturées selon la même définition.
- [0091] Selon un autre mode de réalisation préféré, outre la position des points d'intérêt, des données telles que la détection des limites du terrain peuvent être transmises et validées ou corrigées par le dispositif d'analyse vidéo. Plus généralement, des données de géométrie relatives à ladite scène peuvent être transmises au dispositif d'analyse vidéo et repositionnées par celui-ci ou validées par celui-ci. Ainsi, selon ce mode de réalisation l'étape T2 peut comprendre :
- [0092] - la transmission à au moins un dispositif d'analyse vidéo, des au moins deux vidéos reçues représentatives chacune d'au moins une partie de la scène, des points d'intérêt ou de données relatives à la géométrie de la partie de la scène,
- [0093] - La réception du au moins un dispositif d'analyse vidéo, d'un message de validation de la position de au moins un des points d'intérêt ou de la géométrie de la scène.
- [0094] Lors d'une étape T3, les métadonnées reçues avec chaque vidéo sont analysées pour étudier leur variation. Comme décrit dans la [Fig.1], les métadonnées sont représentatives d'une position géographique des moyens de capture, de l'assiette, de l'inclinaison ou de la boussole (l'inclinaison des moyens de capture par rapport au Nord). Les métadonnées sont transmises de manière périodique, à des intervalles de temps réguliers ou non, et analysées lors de leur réception en les comparant à des métadonnées de référence. Les métadonnées de référence dont des métadonnées transmises par les moyens de capture lors de leur initialisation ou mise en place (lors du verrouillage). Si cette comparaison détecte que la position géographique (par exemple la latitude ou la longitude), ou l'inclinaison, ou l'assiette ou la boussole ont varié et que la variation est supérieure à un seuil donné, alors les points d'intérêt sont

recalculés. Comme indiqué, ces variations sont dues à un déplacement des moyens de capture. De manière avantageuse, lorsque les métadonnées ne varient pas, par exemple parce que les moyens de capture ne sont pas accessibles facilement et ne peuvent donc pas être déplacés par inadvertance ou lorsque les conditions météorologiques sont bonnes (absence de vent par exemple), alors la comparaison peut se faire de manière plus espacée dans le temps et pas systématiquement à chaque réception de nouvelle métadonnée.

- [0095] Si le résultat de la comparaison, étape T4, est supérieur à un seuil, chaque seuil étant défini pour chaque type de métadonnée, alors les points d'intérêt sont redéfinis. Sinon les points d'intérêt ne sont pas redéfinis. Bien entendu, les points d'intérêt étant utilisés lors de l'étape T5, ils sont toujours définis au lancement du procédé lors de la réception des premières métadonnées, les étapes d'analyse et de comparaison T3 et T4 intervenant ensuite au cours de la capture pour vérifier la stabilité géographique des dispositifs de capture.
- [0096] Selon un autre mode de réalisation les métadonnées ne sont pas comparées à un seuil mais doivent être comprises dans une plage de tolérance centrée autour de la valeur de la première métadonnée reçue correspondante dite métadonnée de référence.
- [0097] Il est possible qu'un seul des dispositifs de capture ait bougé et que la variation des métadonnées associées soit suffisamment importante pour nécessiter un nouveau calcul des points d'intérêt. Dans cas, les points d'intérêt sont recalculés uniquement pour cette partie de la vidéo lors de l'étape T2.
- [0098] Selon un autre mode de réalisation, plusieurs seuils peuvent être définis.
- [0099] Le premier seuil S1 est défini comme étant un seuil au-dessous duquel la variation des métadonnées n'implique pas de recalcul de position des points d'intérêt.
- [0100] Le second seuil S2 est défini comme étant un seuil au-dessus duquel les au moins deux dispositifs de capture ne capturent plus de zone commune permettant la détermination d'un nombre suffisants de points d'intérêt fixes pour la composition de la vidéo composite.
- [0101] Si la variation des métadonnées est sous le premier seuil S1, alors il n'y a pas de recalcul des points d'intérêt car cette faible variation n'implique pas de déformations des symétries ni des parallélismes de l'image composite à construire.
- [0102] Si la variation des métadonnées est comprise entre ce premier seuil S1 et le second seuil S2, alors un recalcul des points d'intérêt est nécessaire comme indiqué ci-dessus.
- [0103] Si la variation des métadonnées est supérieure à ce second seuil S2, alors le repositionnement des moyens ou d'au moins l'un des moyens de capture, est nécessaire avant le recalcul des points d'intérêt.
- [0104] Lors d'une étape T5, une vidéo composite représentant la scène est composée à partir des deux vidéos représentant chacune au moins une partie de la scène et des points

d'intérêt.

- [0105] Lors de cette étape, en référence à la [Fig.3], il s'agit de créer une unique vidéo, appelée vidéo composite, dans laquelle une unique scène est créée. La scène re-composée à partir des deux demi-scènes des vidéos de gauche et vidéos de droite est représentée en dessous.
- [0106] Ainsi, lors de cette étape, les deux vidéos sont mises en correspondance, en faisant correspondre les points Lru et Rlu et les points Lrd et Rld. Ceci est illustré en [Fig.3] basse, les points d'intérêt Lru et Rlu sont confondus de même que les points d'intérêt Lrd et Rld.
- [0107] La composition est réalisée en utilisant des algorithmes d'assemblage graphique, « stitching » en anglais, et de coupage, « cropping » en anglais. Dans le cas d'une scène représentant un terrain de sport, les contraintes suivantes sont prises en compte pour la composition de la vidéo composite :
- [0108] - L'image composite doit avoir les lignes de touches haute et basse parallèles,
- [0109] - L'image composite doit avoir la ligne médiane perpendiculaire aux lignes de touches,
- [0110] - Les lignes longeant les buts sur l'image composite sont chirales,
- [0111] - La longueur de ligne de touche basse doit faire 2 fois la définition du flux de plus basse résolution, en nombre de pixels,
- [0112] - la hauteur de la ligne médiane est choisie en comparant les hauteurs des deux points d'intérêt commun et en conservant la plus grande, la seconde devra s'étirer pour y correspondre.
- [0113] - Dans le cas de 3 dispositifs de capture, la hauteur de la ligne médiane sur le dispositif de capture central sera le référent,
- [0114] - La longueur de chaque moitié des lignes de touches hautes est identique
- [0115] - La largeur de l'image composite est de 2 fois la largeur du flux (si 2 flux), sa hauteur est réduite pour éliminer les zones de vides (trous) créées par le stitching.
- [0116] Selon un mode de réalisation particulier, les algorithmes de coupage et collage utilisent des bibliothèques python telles que open-cv et des fonctions `cv2.getPerspectiveTransform` et `cv2.warpPerspective`. Un encodage est ensuite réalisé en format X.264 avec la fonction `cv2.VideoWriter` en précisant la taille de l'image (size) et le nombre d'images par seconde (fps). D'autres fonctions équivalentes peuvent être utilisées lorsque le langage de codage utilisé est le C++.
- [0117] Selon un mode de réalisation particulier, cette étape peut comprendre une sous-étape de validation. Cette sous-étape de validation comprend la transmission à un dispositif d'analyse vidéo, d'une version, par exemple basse définition, de la vidéo composite.
- [0118] Le terminal abonné peut recevoir, des instructions, par un utilisateur, pour ajuster certains points géométriques de la scène re-composée obtenus lors de la composition.

Par exemple, dans le cadre d'un terrain de sport, il peut s'agir de l'ajustement des 4 coins du terrain de sport. Plus précisément, selon un mode préféré, seuls les deux points d'intérêt définis par l'intersection de la ligne de touche supérieure (haut de l'image) et les lignes de but sont modifiables. Les instructions peuvent aussi comprendre un ajustement, telle une diminution de la hauteur de l'image, en éliminant par exemple des éléments inutiles. Les instructions peuvent aussi comprendre le remplacement de ces zones inutiles par des bandeaux publicitaires.

- [0119] Par zone inutile, on peut entendre par exemple, des zones à l'extérieur des limites du terrain.
- [0120] Lors de cette étape de composition T5, selon une variante, les différences entre les paramètres de couleur et luminosité, tels que par exemple le contraste, la luminosité, la saturation, l'exposition entre les deux vidéos sont également prises en compte afin de générer la vidéo composite. Ainsi, ces différences peuvent être lissées, des filtres peuvent également être appliqués pour homogénéiser, voire transformer, l'un ou l'autre flux pour avoir une vidéo composite homogène sur les deux parties ainsi rassemblées.
- [0121] Une fois la vidéo recomposée à partir des points d'intérêt et des deux vidéos, elle est mise à disposition, étape T6.
- [0122] Selon une première variante, la mise à disposition consiste à enregistrer la vidéo composite de manière par exemple à la diffuser ultérieurement.
- [0123] Selon une autre variante, la mise à disposition consiste en l'affichage.
- [0124] Selon une autre variante, la mise à disposition consiste à transmettre la vidéo composite à un ou plusieurs dispositifs en vue de son affichage ou en vue de son enregistrement pour une visualisation ou utilisation ultérieure. La vidéo peut être enregistrée selon des résolutions différentes et ce en fonction de la capacité des réseaux de communication et des terminaux auxquelles elle est destinée.
- [0125] Les trois variantes peuvent bien entendu être combinées. De même, des définitions différentes de la vidéo composite peuvent être mises à disposition selon la bande passante, l'utilisation finale de la vidéo et les capacités du dispositif final d'utilisation de la vidéo.
- [0126] Lorsque la mise à disposition consiste en l'affichage de la vidéo composite (pouvant d'ailleurs être précédée de l'enregistrement), selon un mode de réalisation particulier, la vidéo composite peut être enrichie par un utilisateur, ou de manière automatique par un programme.
- [0127] Dans un mode de réalisation, lorsque c'est un programme qui enrichie la vidéo composite, l'enrichissement peut être réalisé à l'aide d'un logiciel d'analyse d'image et de détection des actions préférentielles liées à l'événement diffusé. Par exemple, lors d'un match, il est possible de détecter les buts, les corners, et de compiler des statistiques, il est également possible de rajouter par exemple le nom des joueurs, le

nombre de possessions de ballons, le nombre de ballons touchés, perdus, récupérés, le temps de possessions de ballons par joueur ou par équipe, le nombre de tirs cadrés ou pas, la distance parcourue par chaque joueur, par match, par poste et ainsi de voir l'évolution des statistiques lors d'une saison complète.

- [0128] Lorsque c'est un utilisateur qui enrichie la vidéo composite, elle est transmise à au moins un dispositif d'analyse, par exemple un dispositif 3 tel qu'illustré en [Fig.7] et utilisé par l'utilisateur qui va enrichir la vidéo composite. Dans ce cas, le procédé de composition de la vidéo comprend
- [0129] - la réception, suite à la transmission, de données d'enrichissement de ladite scène,
- [0130] - la synchronisation des données d'enrichissement avec la vidéo composite ;
- [0131] - la mise à disposition de la vidéo composite enrichie des données d'enrichissement.
- [0132] Lorsque la vidéo est transmise pour enrichissement, elle peut avantageusement être transmise avec un taux de compression élevé de manière à rendre la latence très basse. Le taux de compression est adapté au débit disponible. Ainsi, l'enrichissement se fait presque en direct ou « live » en anglais.
- [0133] Un mode de réalisation, lorsque l'utilisateur enrichie la vidéo, est décrit ultérieurement en [Fig.4].
- [0134] La [Fig.3], brièvement décrite ci-dessus illustre les deux parties de la scène et leur composition par le procédé de composition.
- [0135] La vidéo de gauche représente la capture vidéo d'un premier moyen de capture qui capture la partie droite d'un terrain de sport, plus particulièrement de football. Sur cette vidéo sont positionnés 4 points d'intérêt Llu, Lru, Lrd et Lld. Ces quatre points d'intérêt sont fixes et représentent comme décrit ci-dessus par exemple des intersections de lignes sur un terrain.
- [0136] La zone hachurée sur les deux vidéos de droite et de gauche représente une zone commune ou de chevauchement de la scène, c'est-à-dire une partie de la scène capturée par les deux dispositifs de capture vidéo. Ainsi les points Lru, Rlu représentent des points identiques de la scène et les points Lrd et Rld représentent également un même point de la scène.
- [0137] Le procédé de composition vidéo va générer la vidéo composite, vue de dessous de la [Fig.3], comme décrit précédemment en regard de la [Fig.2].
- [0138] Sur la [Fig.3], les points Lld et Rrd ne coïncident pas avec les bords de l'image. Ceci représente une version non limitative de l'image composite. Dans d'autres modes de réalisation, il est envisagé que les points d'intérêt Lld et Lrd coïncident avec les bords de l'image composite, celle-ci représentant alors en longueur la longueur du terrain. Ainsi la [Fig.3] représente une parmi plusieurs images composites, ceci dépendant du zoom souhaité. On peut envisager ainsi d'avoir une image composite comprenant uniquement le terrain ou également les ou une partie des tribunes, ou par exemple une

zone tampon autour du terrain ou autour de certains bords du terrain.

- [0139] La [Fig.4] représente un procédé d'analyse selon un mode particulier de réalisation de l'invention dans lequel la vidéo composite mise à disposition lors de l'étape T5 en référence à la [Fig.2], est transmise à un dispositif utilisateur. Ce dispositif utilisateur peut être, de manière non limitative, un téléphone mobile, un ordinateur, une tablette, une télévision... La vidéo composite est reçue sur le dispositif utilisateur lors de l'étape U1. Elle est également décodée si elle a été transmise de manière codée pour être ensuite affichée sur un écran du dispositif de l'utilisateur. Lors d'une étape U2, l'utilisateur positionne des informations, dites informations d'enrichissement, sur la vidéo composite ou associe à la vidéo composite des informations d'enrichissement. Les informations d'enrichissement, dans l'exemple d'un match de football, peuvent par exemple correspondre à des informations permettant d'annoter le match sous forme de mots tels que « but », « corner », « touche ». Elles peuvent également identifier les joueurs, ajouter des commentaires personnels de l'utilisateur, ajouter des statistiques sur le nombre de passes réussies, le score.... Ces informations ainsi que la vidéo composite peuvent ensuite être affichées sur le terminal de l'utilisateur, transmises au serveur pour diffusion vers d'autres terminaux d'affichage ou enregistrées ou les deux.
- [0140] Pour ce faire, l'utilisateur utilise de manière avantageuse mais non limitative une interface utilisateur. Cette interface graphique peut par exemple lui permettre de venir sélectionner des événements prédéfinis correspondant au type de scène capturée. Par exemple, lorsque la scène est un match de football, les événements prédéfinis peuvent être « but », « corner », « touche ». L'utilisateur peut également venir associer des noms aux joueurs, demander la compilation de statistiques...
- [0141] Si l'événement est un concert alors il peut venir rajouter le titre de la chanson, il peut également positionner le nom des musiciens, l'année de la chanson, le titre de l'album...
- [0142] Ainsi, les informations d'enregistrement sont obtenues ou positionnées en fonction d'une analyse du contenu de la scène.
- [0143] Une fois les données d'enrichissement positionnées, lors d'une étape U3, la vidéo composite et les données d'enrichissement sont synchronisées si besoin.
- [0144] Finalement, lors d'une étape U4, la vidéo composite enrichie des informations d'enrichissement est mise à disposition.
- [0145] Selon une première variante, la mise à disposition consiste à enregistrer la vidéo composite enrichie de manière à la diffuser ultérieurement par exemple.
- [0146] Selon une autre variante, la mise à disposition consiste en l'affichage de la vidéo composite enrichie sur le terminal de l'utilisateur.
- [0147] Selon une autre variante, la mise à disposition consiste à transmettre la vidéo composite enrichie à un serveur ou à un ou plusieurs dispositifs en vue de son

affichage ou en vue de son enregistrement pour une visualisation ou utilisation ultérieure.

- [0148] Lorsque le serveur reçoit la vidéo enrichie, il peut également compiler ou avoir analysé la vidéo et comparer son analyse aux données d'enrichissement pour vérifier ces dernières. En cas de conflit lors de la vérification, une demande de validation peut être transmise au dispositif d'analyse vidéo et donc à l'utilisateur par l'intermédiaire de l'interface graphique par exemple.
- [0149] Les trois variantes peuvent, bien entendu, être combinées.
- [0150] Le procédé est mis en œuvre dans un dispositif d'analyse et par exemple un dispositif d'analyse tel que décrit en [Fig.6].
- [0151] La [Fig.5] représente un dispositif de capture configuré pour mettre en œuvre un procédé tel que décrit en [Fig.1].
- [0152] Le dispositif 1 comprend des moyens de capture 11. Les moyens de capture 11 sont de préférence une caméra telle que l'on peut retrouver dans un téléphone mobile. Ainsi, de préférence, le dispositif 1 est un téléphone mobile mais peut également être un ordinateur ou plus généralement un dispositif équipé de moyens de capture. Les moyens de capture capturent au moins une partie d'une scène vidéo selon un mode de réalisation préféré.
- [0153] Le dispositif 1 comprend également des moyens d'affichage 16. Ces moyens d'affichage sont optionnels mais peuvent permettre avantageusement de visualiser la scène capturée par les moyens de capture 11. Le dispositif 1 comprend également une interface de communication 14 adaptée à communiquer avec des dispositifs distants, tels que par exemple mais non limitativement avec un serveur 4. Cette interface de communication peut être de type cellulaire (3G, 4G, 5G...) ou WIFI ou les deux, voire d'autres protocoles. Le dispositif 1 comprend également une mémoire 15 contenant des données et instructions de programme et une mémoire 12 permettant d'enregistrer des données, telles par exemple des données vidéos. La mémoire 12 peut par exemple enregistrer de manière temporaire ou plus long terme les données capturées par les moyens de capture 11, avant transmission à l'interface de communication 14. Le dispositif 1 comprend un processeur 13 apte à coopérer avec les moyens précités du dispositif 1 et principalement configuré pour mettre en œuvre le procédé décrit en [Fig.1]. Le dispositif 1 comprend également un accéléromètre 18, un capteur GPS 17, un magnétomètre 19 et un gyroscope (électronique) 20. Le capteur GPS 17 permet de déterminer la position géographique du dispositif de capture et l'accéléromètre 18 permet de déterminer les mouvements de l'appareil (la vitesse de déplacement). Le gyroscope 20 permet de mesurer la rotation du dispositif de capture et le magnétomètre 19 fonctionne comme une boussole numérique. Ainsi, l'inclinaison, la boussole, et l'assiette du dispositif de capture sont mesurés. En outre un journal des positions du

dispositif de capture peut être tenu à jour.

- [0154] Avant tout lancement de la capture selon un mode préféré de réalisation de l'invention, le dispositif de capture est verrouillé, c'est-à-dire qu'il est positionné dans une position considérée comme une position de référence. Il s'agit donc d'un verrouillage physique qui peut être suivi par un verrouillage numérique, permettant d'avoir un rendu constant, à savoir par exemple obtenir une vidéo n'ayant pas de zone surexposée. Lors du verrouillage physique, le dispositif de capture transmet des métadonnées obtenues par le capteur GPS 17 et l'accéléromètre 18 qui constituent les métadonnées de référence utilisées par le procédé selon la [Fig.2] pour la comparaison avec les métadonnées transmises ensuite de manière périodique.
- [0155] La [Fig.6] décrit un dispositif 4 de composition de vidéo et de manière préférée un dispositif de type serveur. Ainsi ce serveur peut par exemple être situé à distance ou dans le « nuage informatique » bien connu en anglais sous le terme de « cloud ». Par ailleurs, selon certains modes de réalisation, les fonctions de ce dispositif peuvent être distribuées dans plusieurs serveurs. Par exemple le dispositif d'analyse 3 peut recevoir les données (vidéo composite et points d'intérêt) d'un premier serveur et transmettre la vidéo enrichie et/ou les données d'enrichissement à un autre serveur, physiquement différent du serveur qui lui a transmis la vidéo composite et éventuellement les points d'intérêt. De même un troisième serveur peut communiquer/coopérer avec l'un ou l'autre des serveurs pour transmettre la vidéo à une ou plusieurs résolutions supplémentaires.
- [0156] Le dispositif de communication de vidéo 4 comprend un processeur 47 configuré pour mettre en œuvre un procédé de composition vidéo tel que décrit en [Fig.2]. Le processeur 47 collabore avec une mémoire d'enregistrement 43 et une mémoire de programme 44. La mémoire d'enregistrement 43 peut par exemple enregistrer de manière temporaire ou plus long terme les données reçues ou transmises par les interfaces de communication 41a, 41b, 41c. Les interfaces de communication 41a, 41b, 41c sont de préférence adaptées à communiquer respectivement avec plusieurs dispositifs de capture 1, un ou plusieurs dispositifs d'analyse 3. Le nombre d'interfaces de communication est donné à titre illustratif. De telles interfaces de communication sont de manière préférée des interfaces de type cellulaire (3 G, 4G, 5G et futures...) ou WIFI norme 5, 6 et futures ou les deux.
- [0157] La [Fig.7] représente un dispositif d'analyse vidéo 3 selon un mode préféré de réalisation.
- [0158] Le dispositif 3 comprend un processeur 33 configuré pour mettre en œuvre un procédé d'analyse vidéo tel que décrit en [Fig.4]. La mémoire 35 comprend des données et instructions de programme du processeur 33 et la mémoire 32 est utilisée pour enregistrer notamment de manière temporaire ou permanente la vidéo reçue à

travers l'interface de communication 34 et des données d'enrichissement saisies par l'utilisateur par l'intermédiaire de l'interface graphique 37. L'interface de communication 34 est de manière préférée une interface de type cellulaire (3 G, 4G, 5G...) ou WIFI ou les deux.

- [0159] La [Fig.8] représente un système selon un mode préféré de réalisation de l'invention. Deux dispositifs de capture 1a et 1b sont disposés autour d'une scène 2, ici un match de football. Les deux dispositifs sont disposés de manière à filmer chacun au moins une partie du terrain de football, avec une zone de chevauchement, c'est-à-dire que chacun des dispositifs peut capturer plus d'un demi-terrain. Les deux dispositifs 1a et 1b sont configurés pour mettre en œuvre un procédé de capture selon l'invention et décrit en regard de la [Fig.1]. Les deux dispositifs 1a et 1b transmettent chacun à travers leurs interfaces de communication respectives, la vidéo capturée (incluant dans un mode de réalisation préféré les étiquettes temporelles) et les métadonnées au serveur 4. Le serveur 4 peut être composé d'un ou plusieurs dispositifs distincts. Le serveur 4 est configuré pour mettre en œuvre un procédé de composition de vidéo selon l'invention et plus précisément décrit en [Fig.2]. Le serveur 4 transmet la vidéo obtenue selon le procédé de composition vidéo à un dispositif 3 d'analyse vidéo qui enrichie la vidéo selon le procédé d'analyse décrit en [Fig.4] et retransmet la vidéo enrichie et/ou les données d'enrichissement et ou la validation de points d'intérêt ou la validation de données d'enrichissement au serveur 4. Le serveur 4 enregistre ensuite les données d'enregistrement et/ou la vidéo annotée ou les transmet au dispositif 5 pour affichage ou enregistrement ou au dispositif 3. Le dispositif 5 est un dispositif électronique tel que par exemple un téléphone mobile, un ordinateur, un téléviseur.
- [0160] Selon un mode de réalisation préféré, un tel système peut être utilisé selon le scénario suivant.
- [0161] Le dispositif d'analyse 3 reçoit des instructions par un utilisateur pour lancer une application. De préférence le dispositif d'analyse 3 est un téléphone mobile. L'application reçoit de l'utilisateur un paramétrage de l'application. Un tel paramétrage peut être par exemple une indication du type d'événement capturé, par exemple un match de sport, par exemple de football. L'application reçoit également des informations relatives au nombre de dispositifs de capture 1a, 1b utilisés pour la capture de l'événement. De même ces dispositifs de capture sont de préférence des téléphones mobiles. L'application crée un QR-code pour chaque dispositif de capture 1a, 1b.
- [0162] Chacun des dispositifs de capture 1a, 1b, scanne l'un des QR-code présents sur le dispositif d'analyse 3. Ceci permet de lier chaque dispositif de capture à l'un des flux. Ceci permet de lancer une application sur les dispositifs de capture ou de télécharger l'application si elle n'était pas présente puis de lancer la synchronisation de l'horloge

selon le protocole NTP. Selon des modes de réalisation avantageux, les dispositifs de capture peuvent également faire des tests de débit pour vérifier la bande passante adaptée en fonction des paramètres du réseau, du type de réseau disponible (2G, 4G, 5G, WIFI norme 5, 6...), de l'environnement.

- [0163] Les dispositifs de capture s'identifient également auprès du serveur 4 de façon à permettre au serveur 4 de détecter quel flux est associé à quel dispositif de capture et éventuellement de modifier ceci.
- [0164] Un opérateur verrouille ensuite la position des dispositifs de capture autour de la scène en vérifiant que chaque partie de la scène capturée par chacun des dispositifs de capture comprend au moins une zone de chevauchement dans laquelle deux points d'intérêt fixes peuvent être déterminés. Cette détermination peut être visuelle par l'opérateur mais cette détermination peut également se faire au niveau du serveur. Par exemple le serveur peut transmettre sur les dispositifs de capture des notifications pour orienter/déplacer les dispositifs de capture, tant qu'il n'a pas réussi à déterminer une première fois, de points d'intérêt communs. Une fois les points d'intérêt communs détectés, l'opérateur verrouille la position physique des dispositifs de capture. Ensuite, comme décrit précédemment, les points d'intérêt sont recalculés automatiquement en fonction de la variation des métadonnées. Seule une variation importante des métadonnées ne permettant plus la détection de points d'intérêt rendra nécessaire le repositionnement physique des dispositifs de capture.
- [0165] Lors du verrouillage, les métadonnées sont également enregistrées et constituent ainsi des métadonnées de référence de position des dispositifs de capture 1a, 1b. Ces métadonnées de référence sont utilisées lors de la comparaison avec les métadonnées reçues de manière périodique (régulièrement ou irrégulièrement) pour déterminer le besoin de recalcul de la position des points d'intérêt comme indiqué en référence à la [Fig.2].
- [0166] La [Fig.9] illustre le positionnement de deux dispositifs 1a et 1b de capture vidéo autour d'une scène de football. Les dispositifs 1a et 1b sont semblables au dispositif 1 décrit en [Fig.5]. Ils sont positionnés géographiquement proches de la ligne médiane du terrain. Le dispositif 1a est orienté vers la partie gauche du terrain et le dispositif 1b est orienté vers la partie droite du terrain, de manière à ce que leurs moyens de capture respectifs capturent plus de la moitié de la scène soit plus d'un demi-terrain. Le dispositif de capture 1a a un angle de capture α et le dispositif de capture 1b a un angle de capture β . La zone hachurée de la [Fig.8] montre la zone de chevauchement des deux captures vidéo opérées par le dispositif 1a et le dispositif 1b.
- [0167] La [Fig.10] illustre le positionnement de trois dispositifs 1c, 1d et 1e de capture vidéo autour d'une scène de football. Les dispositifs 1c, 1d et 1e sont semblables aux dispositifs 1a et 1b décrits en référence à la [Fig.9] mais les moyens de capture ont des

angles de capture différents. Le troisième dispositif de capture 1e peut être placé de telle manière que ses moyens de capture soient parallèles à la ligne de touche.

- [0168] Il peut être nécessaire de recourir à un nombre plus important de dispositifs de capture. Notamment lorsque deux dispositifs de capture ne sont pas suffisants pour capturer intégralement l'ensemble d'une scène, par exemple parce que leur angle de capture est insuffisant ou parce qu'ils ne peuvent pas être positionnés de manière à couvrir toute la scène. Certaines limitations techniques ou géographiques peuvent également provoquer l'ajout de nouveaux moyens de capture par exemple. Ici, les angles de capture respectifs γ et ε des dispositifs 1c et 1d étant inférieurs aux angles de capture des dispositifs 1a et 1b, un troisième dispositif 1e est ajouté dont l'angle de capture est μ . Ainsi le positionnement des trois dispositifs de capture et leurs angles permettent de couvrir l'intégralité du terrain.
- [0169] La [Fig.11] représente un système selon un second mode de réalisation de l'invention. Un tel système est particulièrement avantageux lorsque les dispositifs de capture 1a et 1b ainsi que 3 ne sont pas connectés à un réseau tel un réseau cellulaire (3G, 4G, 5G, 6G) ou WIFI. Dans ce cas, en effet, ils ne peuvent pas communiquer avec un serveur tel qu'illustré en [Fig.8] et donc transmettre les vidéos capturées pour re-composition à un serveur.
- [0170] Une telle configuration est également avantageuse lorsque la bande passante entre les dispositifs de capture 1a et 1b et le serveur 4 est faible et ne permet pas l'émission des flux vidéos et métadonnées au serveur 3 à une qualité suffisante.
- [0171] Dans ce mode de réalisation, les dispositifs de capture 1a, 1b et le dispositif d'analyse 3 communiquent par l'intermédiaire d'un réseau local établi entre eux. Ceci est notamment possible par le biais des réseaux WIFI nouvelle génération ou réseaux cellulaires 5G, 6G et 6E.
- [0172] Dans ce mode de réalisation, le dispositif d'analyse 3 synchronise les deux dispositifs de capture 1a et 1b en créant une horloge temporelle provisoire qu'il leur transmet. Les flux vidéo des deux dispositifs 1a et 1b peuvent ensuite être synchronisés par le dispositif 3.
- [0173] Le dispositif 3 reçoit les deux flux vidéo et construit la vidéo composite à partir des deux vidéos reçues et des métadonnées reçues. Pour ce faire, le dispositif 3 met en œuvre le procédé selon la [Fig.2]. Une fois recomposée, la vidéo est ensuite enrichie par le dispositif 3 selon les étapes U2 à U4 du procédé décrit en [Fig.4].
- [0174] La vidéo ainsi enrichie est transmise au serveur 4 lorsque le dispositif 3 est de nouveau connecté à un réseau cellulaire ou WIFI ou lui permettant de bénéficier de suffisamment de bande passante pour la transmission. Le dispositif 3 et le serveur 4 peuvent ainsi enrichir, analyser la vidéo composite. De la même manière que précédemment, le serveur 4 peut demander des validations des informations

d'enrichissement transmises par le dispositif 3 et compiler des statistiques.

[0175] Bien entendu par flux vidéo ou par vidéo, on entend également tout au long de la description précédente, l'audio associée à cette vidéo.

Revendications

[Revendication 1]

Procédé de composition d'une vidéo comprenant :

- la réception (T1) d'un au moins un premier dispositif de capture vidéo (1a) d'une première vidéo représentative d'une première partie d'une scène (2) et de métadonnées associées à au moins une image fixe de ladite première partie de ladite scène, ladite métadonnée étant représentative d'une position géographique du premier dispositif de capture (1a) au moment de la capture de l'image,
- la réception (T1) d'un au moins un second dispositif de capture vidéo (1b) d'une deuxième vidéo représentative d'au moins une deuxième partie de ladite scène (2) et de métadonnées associées à au moins une image fixe de ladite seconde partie de ladite scène, ladite métadonnée étant représentative de la position géographique du second dispositif de capture (1b) au moment de la capture de l'image, lesdites premières et deuxièmes vidéos étant synchronisées,
- la détermination (T2) de points d'intérêt fixes dans chacune desdites vidéos,
- ladite composition (T5) d'une vidéo composite représentative de ladite scène à partir desdites première et au moins une deuxième vidéos par mise en correspondance desdits points d'intérêt,
- la mise à disposition (T6) de ladite vidéo représentative de ladite scène,

la détection (T3, T4) d'un mouvement d'au moins un des deux dispositifs de capture sur la base des métadonnées reçues déclenchant une nouvelle détermination (T2) des points d'intérêt fixes dans ladite vidéo capturée par le dispositif en mouvement puis ladite composition (T5) d'une nouvelle vidéo composite.

[Revendication 2]

Procédé selon la revendication 1 caractérisé en ce que la détermination (T2) de points d'intérêt fixes dans chacune desdites vidéos comprend :

- la transmission à au moins un dispositif d'analyse vidéo (3), desdites première et au moins deuxième vidéos,
- la réception dudit au moins un dispositif d'analyse vidéo (3), desdits points d'intérêt relatifs à chacune desdites parties de ladite scène (2).

[Revendication 3]

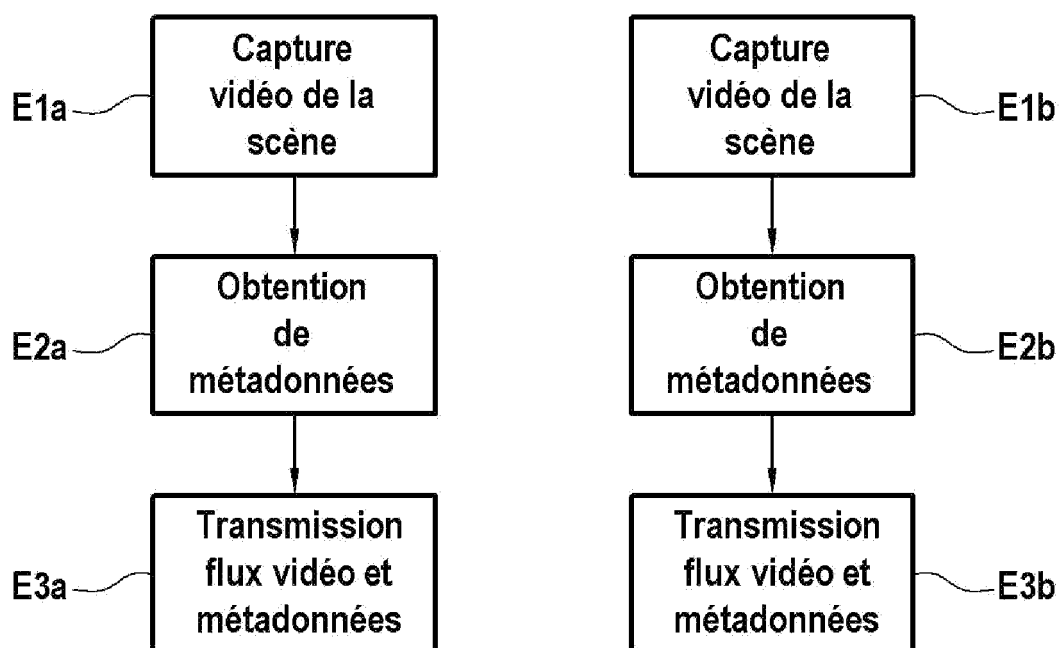
Procédé selon l'une des revendications 1 à 2 caractérisé en ce que la mise à disposition (T6) comprend l'un ou l'autre ou plusieurs parmi :

- l'enregistrement de ladite vidéo composite,
- l'affichage de ladite vidéo composite,

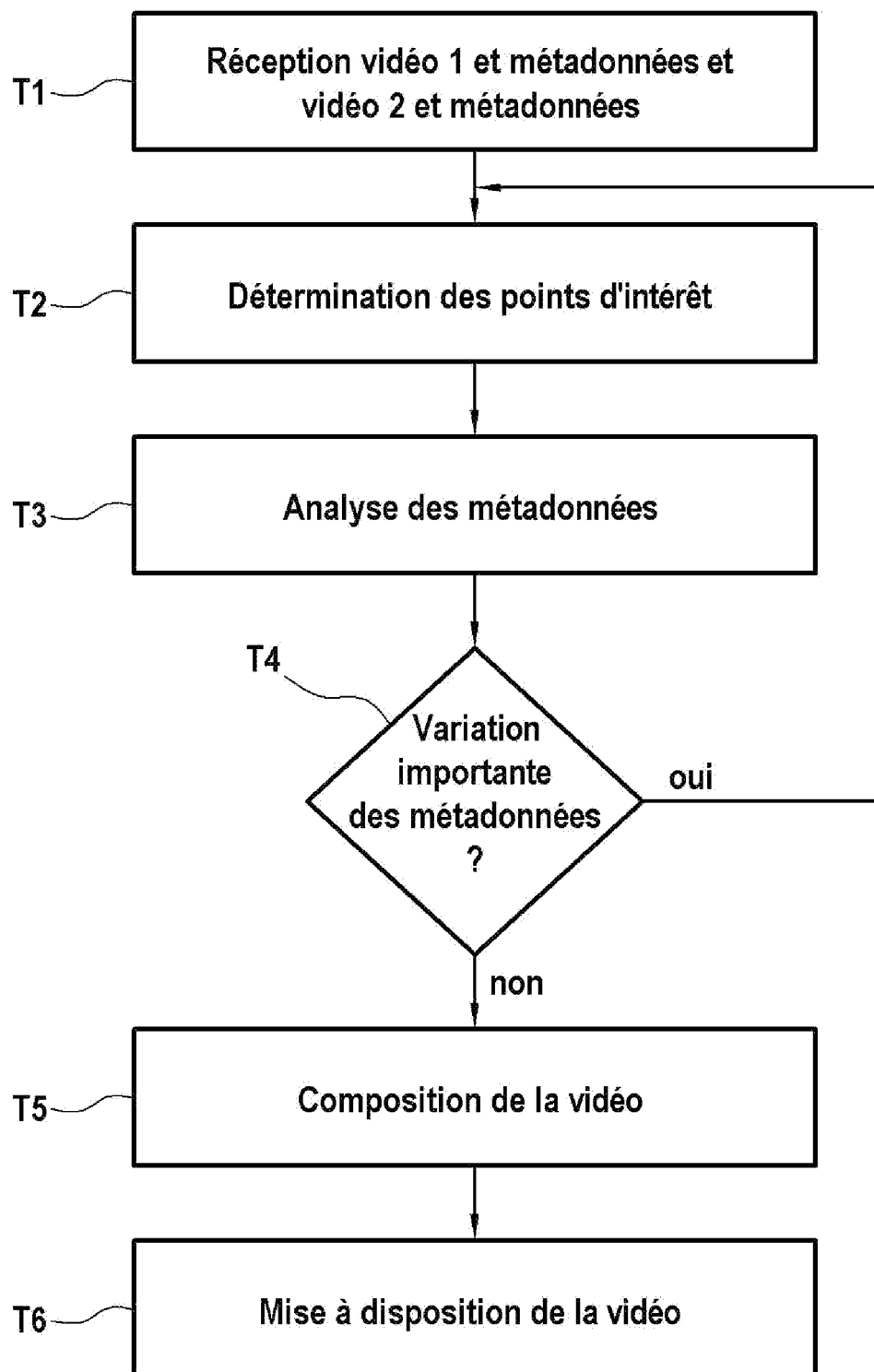
- la transmission de ladite vidéo composite à au moins un dispositif d'analyse (3).
- [Revendication 4] Procédé selon la revendication 3 caractérisé en ce que lorsque ladite vidéo composite est transmise à au moins un dispositif d'analyse (3), il comprend en outre
- la réception, suite à ladite transmission, de données d'enrichissement de ladite scène,
 - la synchronisation des données d'enrichissement avec ladite vidéo composite ;
 - la mise à disposition de ladite vidéo composite enrichie des données d'enrichissement.
- [Revendication 5] Procédé d'analyse de vidéo comprenant
- La réception (U1), d'une vidéo composite représentative d'une scène (2) obtenue par un procédé selon l'une des revendications 1 à 4 ,
 - L'ajout (U2) de données d'enrichissement sur ladite vidéo composite pour annoter ladite vidéo composite,
 - La mise à disposition (U4) de la vidéo composite annotée et des données d'enrichissement.
- [Revendication 6] Procédé d'analyse selon la revendication 5 caractérisé en ce que l'obtention des données d'enrichissement est obtenue par un positionnement des données d'enrichissement par un utilisateur à travers une interface graphique (37).
- [Revendication 7] Serveur comprenant :
- au moins une première interface de communication (41a) avec un dispositif de capture vidéo,
 - au moins une seconde interface de communication (41b) avec un dispositif d'affichage vidéo,
 - au moins un processeur (47) apte à mettre en œuvre un procédé selon l'une des revendications 1 à 4.
- [Revendication 8] Système comprenant au moins un premier (1a) et un second (1b) dispositif de capture vidéo, au moins un serveur (4) selon la revendication 7 et au moins un dispositif d'analyse (3, 5) apte à mettre en œuvre un procédé selon la revendication 5 ou 6.
- [Revendication 9] Programme d'ordinateur comportant des instructions pour l'exécution des étapes 6 du procédé de composition d'une vidéo selon l'une des revendications 1 à 4 ou du procédé d'analyse de vidéo selon l'une quelconque des revendications 5 ou 6 lorsque ledit programme est exécuté par un ordinateur.

[Revendication 10] Support d'enregistrement lisible par un ordinateur sur lequel est enregistré un programme d'ordinateur comprenant des instructions pour l'exécution des étapes du procédé de composition d'une vidéo selon l'une des revendications 1 à 4 ou du procédé d'analyse de vidéo selon l'une quelconque des revendications 5 ou 6.

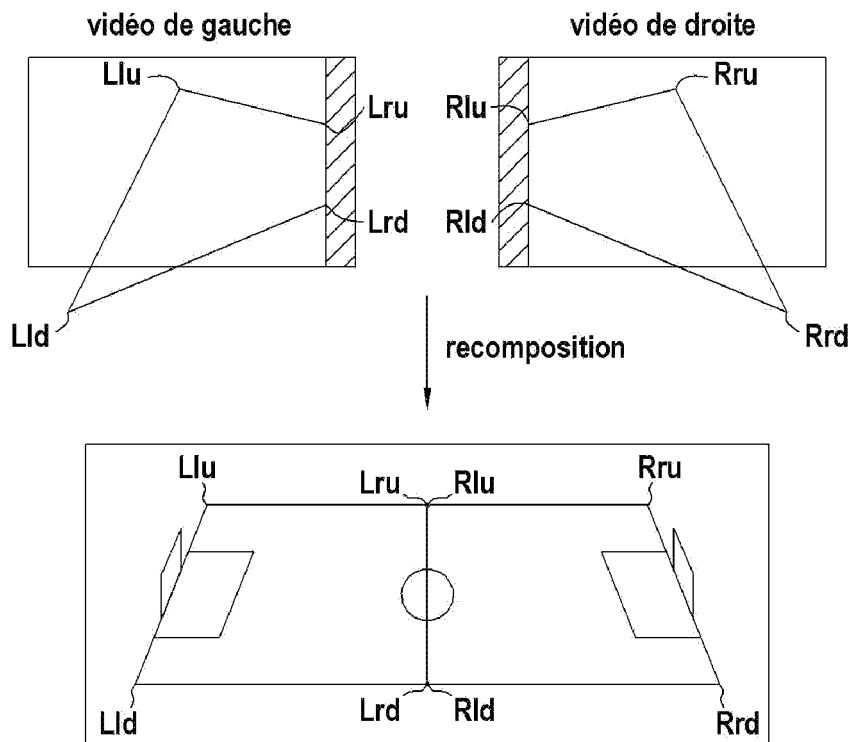
[Fig. 1]



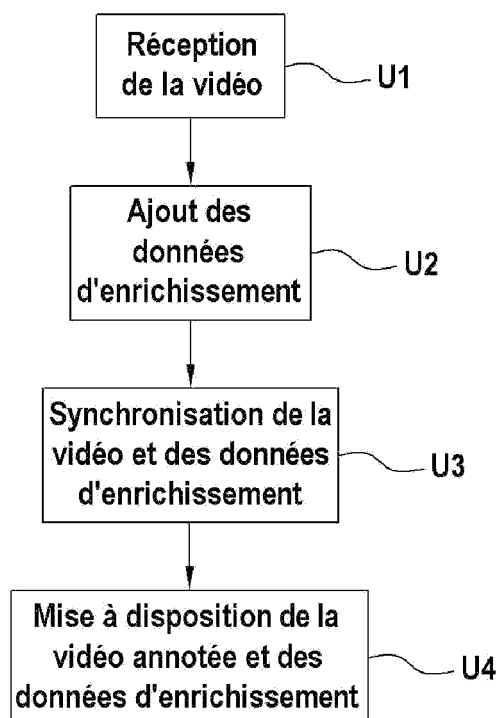
[Fig. 2]



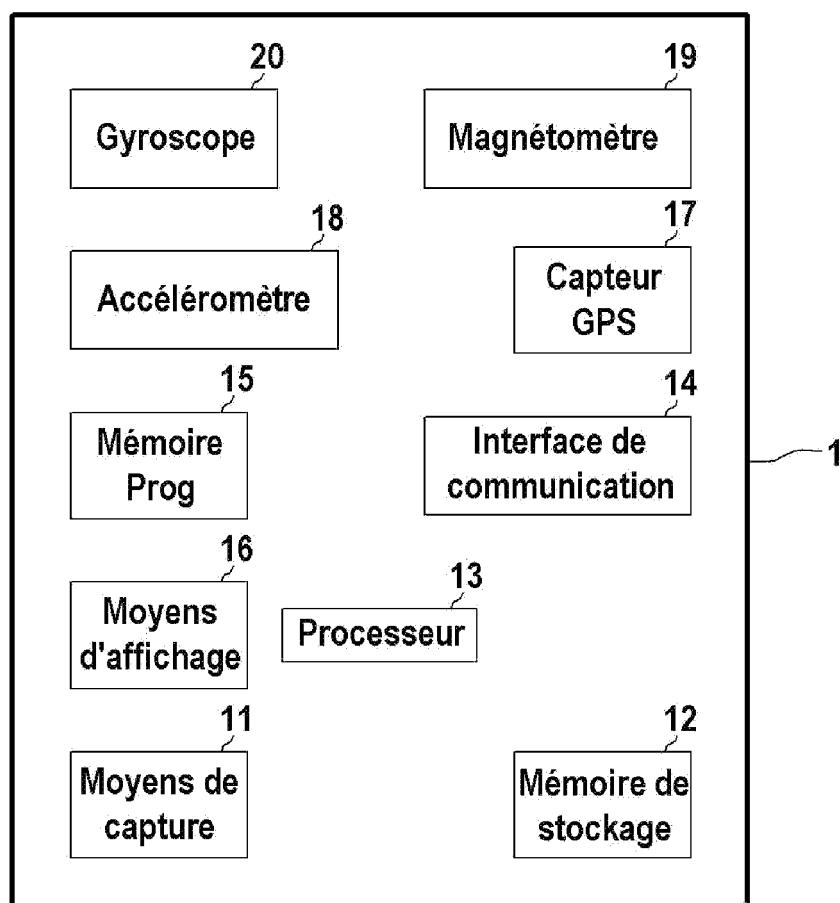
[Fig. 3]



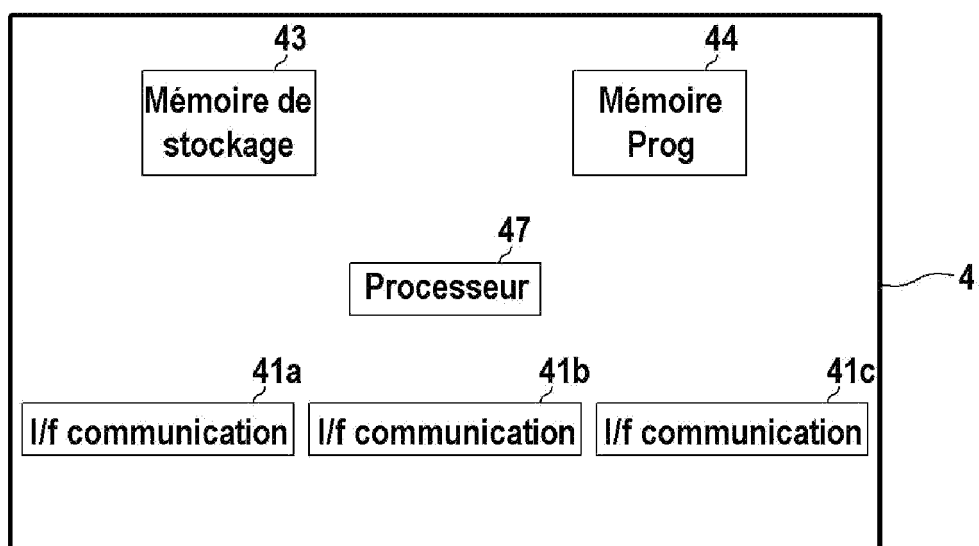
[Fig. 4]



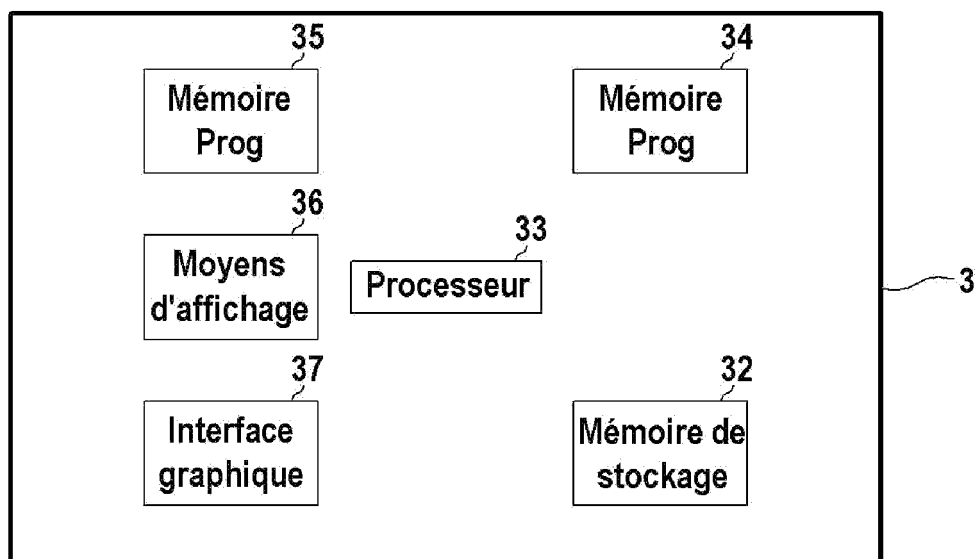
[Fig. 5]



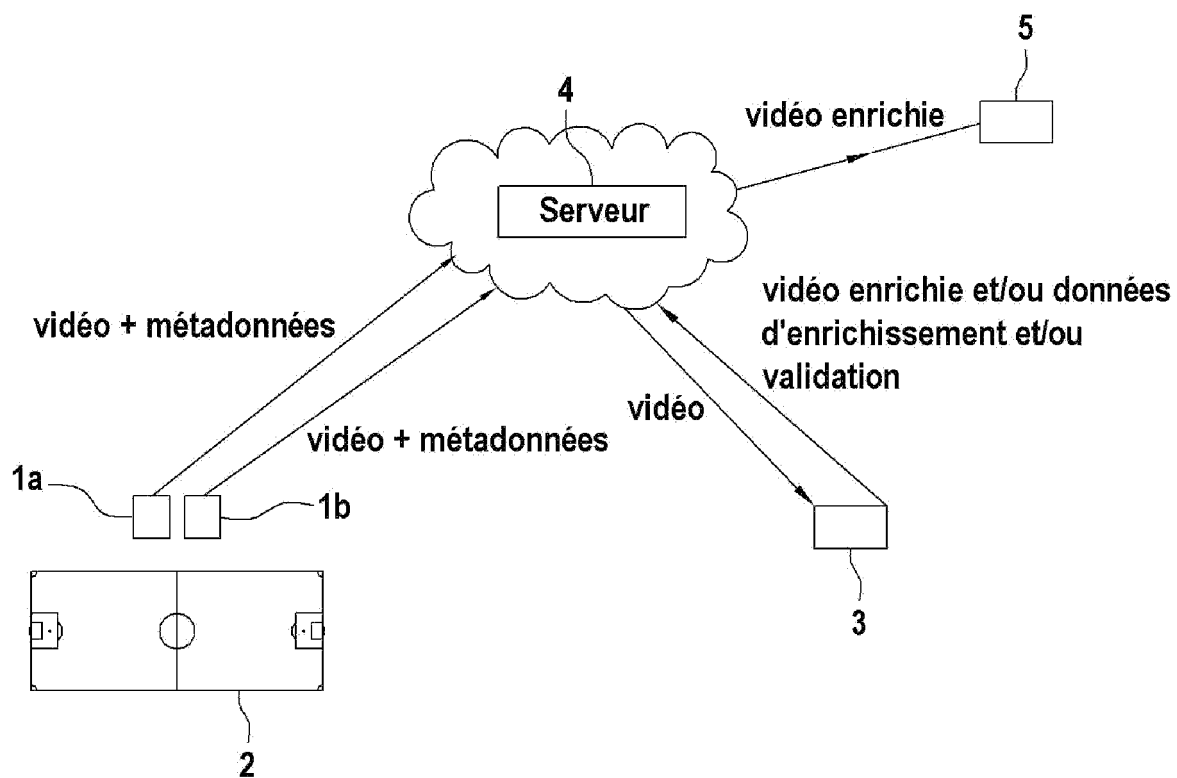
[Fig. 6]



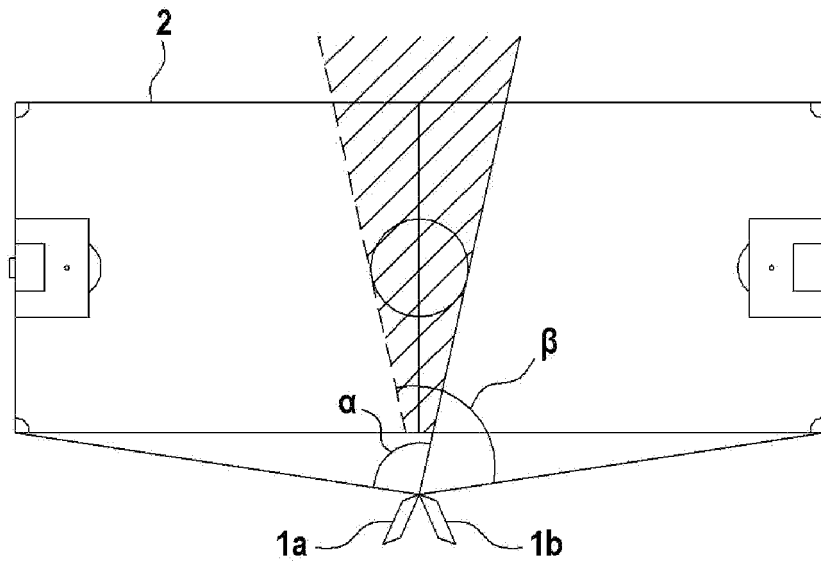
[Fig. 7]



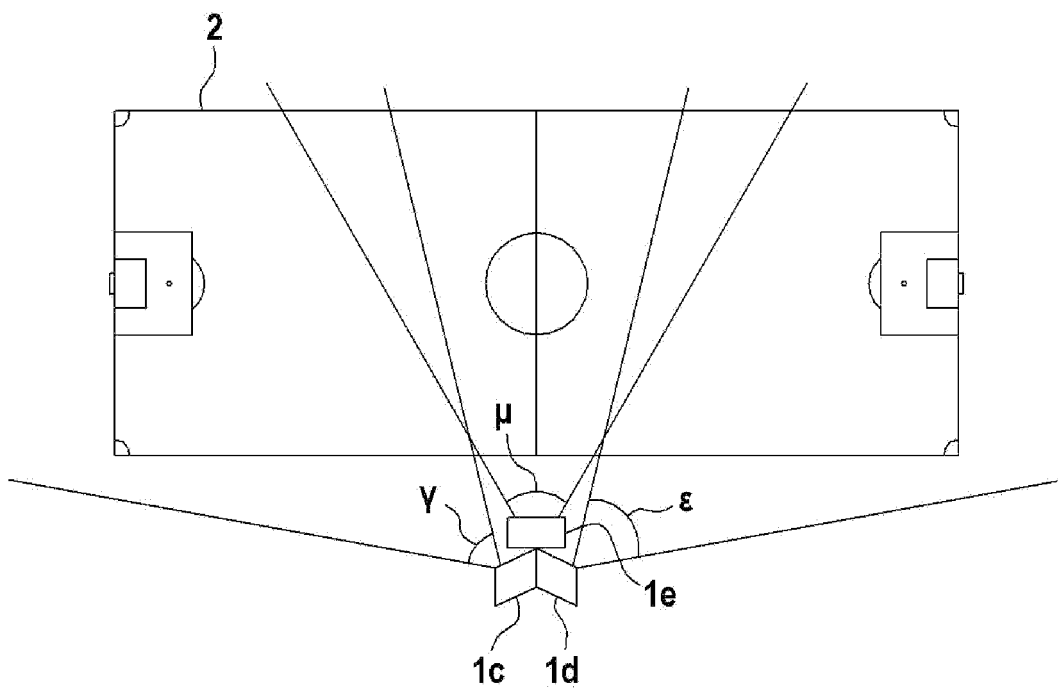
[Fig. 8]



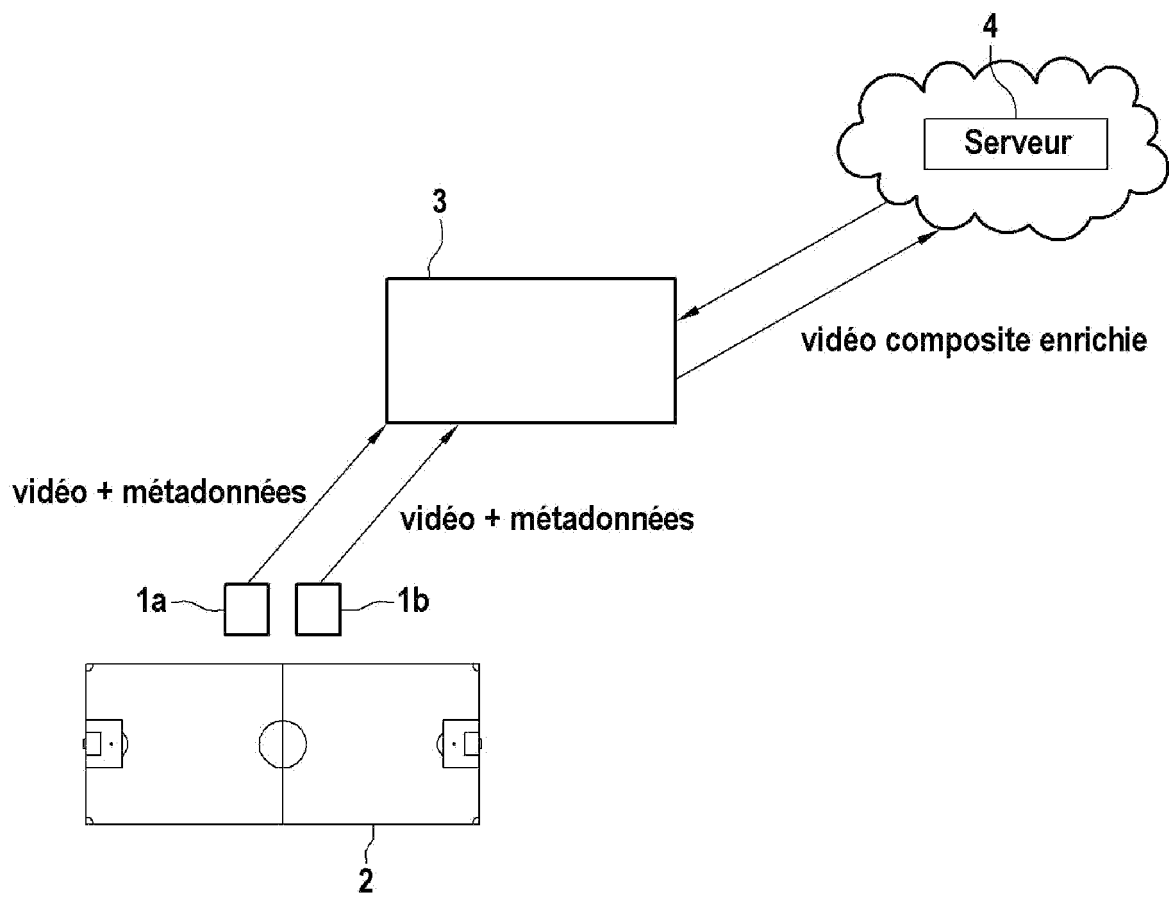
[Fig. 9]



[Fig. 10]



[Fig. 11]



RAPPORT DE RECHERCHE

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

OBJET DU RAPPORT DE RECHERCHE

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

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

CONDITIONS D'ETABLISSEMENT DU PRESENT RAPPORT DE RECHERCHE

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

☐ Le demandeur a maintenu les revendications.

☒ Le demandeur a modifié les revendications.

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

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

☐ Un rapport de recherche préliminaire complémentaire a été établi.

DOCUMENTS CITES DANS LE PRESENT RAPPORT DE RECHERCHE

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

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

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

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

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

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

US 2019/191146 A1 (KOYAMA TATSUYA [JP] ET
AL) 20 juin 2019 (2019-06-20)

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

US 11 043 008 B2 (PANASONIC IP MAN CO LTD
[JP]) 22 juin 2021 (2021-06-22)

US 11 188 776 B2 (GENETEC INC [CA])
30 novembre 2021 (2021-11-30)

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

NEANT