

19 RÉPUBLIQUE FRANÇAISE
INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE
COURBEVOIE

11 N° de publication : **3 143 248**
(à n'utiliser que pour les
commandes de reproduction)
21 N° d'enregistrement national : **22 13086**

51 Int Cl⁸ : **H 04 N 19/50 (2023.01), G 06 T 9/00, G 06 N 3/08**

12

DEMANDE DE BREVET D'INVENTION

A1

22 Date de dépôt : 09.12.22.

30 Priorité :

43 Date de mise à la disposition du public de la demande : 14.06.24 Bulletin 24/24.

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

60 Références à d'autres documents nationaux apparentés :

○ Demande(s) d'extension :

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

72 Inventeur(s) : **HENRY Felix, LADUNE Theo, PHILIPPE Pierrick et CLARE Gordon.**

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

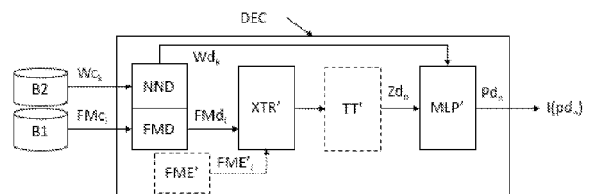
54 **Matériaux** : **Procédé et dispositif de codage et décodage d'images.**

57 Procédé et dispositif de codage et décodage d'images

L'invention concerne un procédé et dispositif de codage et décodage d'un signal comprenant une pluralité d'échantillons (Pdn). Le procédé de décodage comporte les étapes suivantes :

- décodage (E31, FMD) d'un premier groupe de cartes de caractéristiques (FMdi) ;
- décodage (E35, NND) d'un ensemble de paramètres (Wdk) représentatifs d'un réseau de neurones (MLP') ;
- pour au moins un échantillon, dit échantillon courant (Pdn), du signal à décoder, associé à une position (xn, yn) dans le signal à décoder :
 - construction (E33, E34, XTR', TT') d'un vecteur caractéristique (Zdn) à partir des cartes de caractéristiques dudit premier groupe (FMdi), en fonction de ladite position (xn, yn) dudit échantillon courant, et :
 - traitement (E36, MLP') dudit vecteur caractéristique (Zdn) par un réseau de neurones artificiels (MLP') défini par les paramètres décodés (Wdk) pour fournir un vecteur représentatif d'une valeur décodée de l'échantillon courant (Pdn).

Figure pour l'abrégié : Fig. 2.



FR 3 143 248 - A1



Description

Titre de l'invention : Procédé et dispositif de codage et décodage d'images.

Technique antérieure

- [0001] L'invention se rapporte au domaine général du codage de signaux mono ou multidimensionnels. Elle concerne plus particulièrement la compression des images et vidéo numériques.
- [0002] Les images numériques font en général l'objet d'un codage source visant à les compresser afin de limiter les ressources nécessaires à leur transmission et/ou à leur stockage. Il existe de nombreux standards de codage, tels que, pour les images fixes, les standards de la famille JPEG, et pour les images animées, ou vidéo, les standards des organismes ITU/MPEG (H.264/AVC, H.265/HEVC, H.266/VVC, etc.) ainsi que leurs extensions (MVC, SVC, 3D-HEVC, etc.)
- [0003] L'encodage d'une image est généralement effectué en divisant l'image en plusieurs blocs rectangulaires, et en encodant ces blocs de pixels selon une séquence de traitement donnée. Dans les techniques existantes de compression vidéo, le traitement d'un bloc comprend typiquement une prédiction des pixels du bloc effectuée à l'aide de pixels précédemment codés puis décodés présents dans l'image en cours d'encodage, auquel cas on parle de « prédiction Intra », ou d'images précédemment codées, auquel cas on parle de « prédiction Inter ». Cette exploitation des redondances spatiales et/ou temporelles permet d'éviter de transmettre ou de stocker la valeur des pixels de chaque bloc de pixels, en représentant certains au moins des blocs par un résiduel représentant une différence entre les valeurs de prédiction des pixels du bloc et les valeurs réelles des pixels du bloc prédit.
- [0004] Les formats vidéo ne cessant d'évoluer pour compresser toujours davantage et s'adapter à la variété de formats attendus et des réseaux de communication, les possibilités de prédiction deviennent de plus en plus grandes et les algorithmes de codage et décodage classiques très complexes.
- [0005] En plus de ces approches classiques proposées par les normes de compression (JPEG, MPEG, ITU), les approches basées sur l'intelligence artificielle, et notamment neuronales, tendent à se développer.
- [0006] Certaines de ces approches neuronales peuvent être vues comme une simple extension de la notion de compétition des techniques de compression susmentionnées, telles que la compétition de mode de prédiction et transformation en codage vidéo.
- [0007] D'autres approches utilisent le concept « d'auto-encodeur ». Les auto-encodeurs sont des algorithmes d'apprentissage à base de réseaux de neurones artificiels, qui

permettent de construire une nouvelle représentation d'un jeu de données.

L'architecture d'un auto-encodeur est constituée de deux parties : l'encodeur et le décodeur. L'encodeur est constitué par un ensemble de couches de neurones, qui traitent les données afin de construire de nouvelles représentations dites "encodées", appelées aussi « représentations latentes ». À leur tour, les couches de neurones du décodeur reçoivent ces représentations et les filtrent afin d'essayer de reconstruire les données de départ. Les différences entre les données reconstruites et les données initiales permettent de mesurer l'erreur commise par l'auto-encodeur. L'entraînement consiste à modifier les paramètres de l'auto-encodeur afin de réduire l'erreur de reconstruction mesurée sur les différents échantillons du jeu de données. Les performances de tels systèmes à base d'auto-encodeur se font au prix d'une augmentation considérable de l'empreinte mémoire et de la complexité comparativement aux approches conventionnelles telles que proposées par les normes de compression. Ils peuvent avoir des millions de paramètres et peuvent nécessiter jusqu'à un million de MAC (multiplication-accumulation) pour décoder un pixel. Cela rend de tels décodeurs bien plus complexes que les décodeurs conventionnels, ce qui pourrait entraver l'adoption de la compression basée sur l'apprentissage.

[0008] Plus récemment, une technique simple d'encodage fondé sur un réseau de neurones a été décrite dans l'article « COmpression with Implicit Neural representations » d'Emilien Dupont et al. (arXiv:2103.03123). La technique de codage proposée consiste à ajuster un réseau de neurones à une image, quantifier les poids du réseau et les transmettre. Au moment du décodage, le réseau de neurones est évalué à chaque position de pixel pour reconstruire l'image. Une telle technique reste cependant inefficace en termes de compression.

[0009] Il existe donc un besoin pour une solution permettant de coder/compresser une image ou une séquence d'images de manière simple et efficace.

Exposé de l'invention

[0010] L'invention vise un procédé de codage d'un signal comprenant une pluralité d'échantillons à coder comportant les étapes suivantes :

- une étape de construction, comportant les sous-étapes de :
 - construction d'un premier groupe de cartes de caractéristiques ;
 - pour au moins un échantillon, dit échantillon courant, du signal à coder, associé à une position dans le signal à coder :
 - construction d'un vecteur caractéristique à partir desdites cartes de caractéristiques dudit premier groupe, en fonction de ladite position dudit échantillon courant; - traitement dudit vecteur caractéristique par un réseau de neurones artificiels défini par un ensemble de paramètres, pour fournir un vecteur représentatif d'une valeur décodée

de l'échantillon courant ;

- mise à jour d'au moins une valeur d'une desdites cartes de caractéristiques dudit premier groupe et/ou d'au moins un paramètre dudit réseau, en fonction d'une mesure de performance de codage,

- une étape de codage dudit premier groupe de cartes de caractéristiques et dudit ensemble de paramètres.

[0011] L'invention vise aussi un procédé de décodage d'un signal comprenant une pluralité d'échantillons à décoder comportant les étapes suivantes :

- décodage d'un premier groupe de cartes de caractéristiques ;

- décodage d'un ensemble de paramètres représentatifs d'un réseau de neurones ;

- pour au moins un échantillon, dit échantillon courant, du signal à décoder, associé à une position dans le signal à décoder :

- construction d'un vecteur caractéristique à partir des cartes de caractéristiques dudit premier groupe, en fonction de ladite position dudit échantillon courant, et :

- traitement dudit vecteur caractéristique par un réseau de neurones artificiels défini par les paramètres décodés pour fournir un vecteur représentatif d'une valeur décodée de l'échantillon courant.

[0012] Au sens de l'invention, on entend par encodage, ou « codage », l'opération qui consiste à représenter un ensemble d'échantillons sous une forme compacte portée par exemple par un train binaire numérique. On entend par décodage l'opération qui consiste à traiter un train binaire numérique pour restituer des échantillons décodés.

[0013] Par « échantillon » du signal, on entend une valeur prélevée dans le signal.

L'échantillonnage du signal produit une suite de valeurs discrètes nommées échantillons. Dans le cas d'un signal d'image, l'échantillon prend le nom de pixel qui peut être par exemple un pixel couleur représenté traditionnellement par un triplet de valeurs, par exemple (R,G,B) ou (Y,U,V). De manière générale, un échantillon est représenté par un vecteur d'une ou plusieurs composantes, ou données.

[0014] Par « signal comprenant une pluralité d'échantillons », on entend un signal à une (audio, son), deux (image) ou plus de deux (image stéréoscopique, multiscopique, image associée à une carte de profondeur, vidéo, etc.) dimensions. En fonction de cette dimensionnalité, l'échantillon possède une, deux ou plusieurs coordonnées dans le signal. Dans le cas d'un signal d'image, la position de l'échantillon est repérée par ses coordonnées en abscisse (x) et ordonnée (y).

[0015] Par « cartes de caractéristiques », on entend une représentation abstraite du signal comportant une pluralité de données variables, discrètes ou non, que l'on appelle aussi valeurs, par exemple des nombres réels ou entiers. De manière connue, ces cartes sont aussi désignées sous le terme de « représentation latente » du signal d'entrée.

[0016] Par « vecteur caractéristique de données construit à partir des cartes de caracté-

ristiques en fonction d'une position » on entend un vecteur constitué d'un ou plusieurs éléments, ou données, de préférence discrètes, les données étant construite à partir des cartes de caractéristiques à une position déterminée par la position de l'échantillon en cours de traitement dans le signal. Ce vecteur caractéristique est celui qui est appliqué à l'entrée du réseau de neurones. Par exemple, dans le cas d'un signal audio monodimensionnel, un tel vecteur peut être constitué à partir d'une pluralité de valeurs prélevées dans chacune des cartes de caractéristiques à la même coordonnée que l'échantillon à coder. Dans le cas d'une image, un tel vecteur peut être constitué à partir d'une pluralité de valeurs prélevées dans chacune des cartes de caractéristiques aux mêmes coordonnées en abscisse et en ordonnée que l'échantillon à coder. Il peut aussi être constitué à partir d'une pluralité de valeurs prélevées dans chacune des cartes de caractéristiques à une position qui dépend non seulement de la position de l'échantillon mais aussi de la taille de la carte de caractéristiques. Une fois que ces valeurs sont prélevées dans les cartes de caractéristiques, elles peuvent être traitées pour constituer le vecteur caractéristique, avant l'entrée dans le réseau de neurones, par exemple par une quantification, un filtrage, une interpolation, etc.

- [0017] Par « réseau de neurones », on entend un réseau neuronal tels un réseau neuronal convolutif, un perceptron multicouche, un LSTM (pour « Long Short Term Memory » en anglais), etc. Le réseau de neurones est défini par exemple par une pluralité de couches de neurones artificiels qui comportent un ensemble de fonctions d'activation, de pondération et d'addition (par exemple, une couche peut calculer $y = f(Ax+b)$, où y et b sont des vecteurs de dimension N , x est un vecteur de dimension M , A est une matrice de dimension $M \times N$, et f est la fonction d'activation). Par la suite on parle de « réseau de neurones » ou « réseau de neurones de synthèse ».
- [0018] Par « paramètre du réseau de neurones », on entend une des valeurs qui caractérise le réseau de neurones, par un exemple un poids associé à l'un des neurones (coefficient de filtre, pondération, biais, valeur affectant le fonctionnement de la non-linéarité), etc.
- [0019] Par « traitement par un réseau de neurones », on entend l'application d'une fonction exprimée par un réseau de neurones au vecteur caractéristique d'entrée pour produire un vecteur de sortie représentatif de l'échantillon à coder (resp. à décoder). Ce vecteur de sortie peut comporter une ou plusieurs données représentatives de l'échantillon.
- [0020] Par « mesure de performance », on entend une mesure entre au moins une valeur d'un échantillon à coder et une valeur décodée dudit échantillon. La mesure peut évaluer par exemple une distorsion, ou une erreur perceptive. Elle peut être effectuée sur un échantillon ou une pluralité d'échantillons (par exemple, l'image complète). La mesure peut comporter aussi une mesure du débit, notamment associé au codage du réseau de neurones et/ou au codage des cartes de caractéristiques. La mesure peut être une mesure conjointe entre le débit et la distorsion au travers de leur pondération.

Comme il est bien connu de l'état de l'art, il est généralement procédé à une minimisation de la valeur de cette mesure jusqu'à atteindre une valeur cible.

- [0021] Par « étape de construction » on entend une étape qui vise à construire les paramètres représentatifs de l'image, avant leur codage effectif. Les sous-étapes de construction peuvent être répétées autant que nécessaire pour obtenir une mesure de performance acceptable.
- [0022] De manière générale, on considère que les étapes d'un procédé de codage ou décodage ne doivent pas être interprétées comme étant liées à une notion de succession temporelle. Autrement dit, les étapes peuvent être effectuées dans un ordre différent de celui indiqué dans la revendication indépendante de codage ou de décodage, voire en parallèle.
- [0023] Le procédé de codage selon l'invention réalise une construction des paramètres de codage, à partir du signal d'entrée, par exemple une image, en entraînant un réseau de neurones sur des vecteurs caractéristiques associés à une position d'un échantillon à coder. Ces vecteurs caractéristiques sont construits à partir de cartes de caractéristiques. Au cours de l'entraînement, ou construction, les paramètres du réseau de neurones et les valeurs des cartes de caractéristiques sont mis à jour en fonction d'une mesure de performance, par exemple de type débit-distorsion. Lorsque l'entraînement est terminé, c'est-à-dire que la mesure de performance obtenue est satisfaisante, le codage effectif des paramètres du réseau de neurones et/ou des valeurs des cartes de caractéristiques peut être effectué et mémorisé ou transmis à destination du décodeur.
- [0024] Avantagusement, le processus d'entraînement permet de raffiner les paramètres du réseau de neurones et/ou les valeurs des cartes de caractéristiques jusqu'à obtenir une représentation adéquate en termes de performance, par exemple un équilibre souhaité entre le débit généré et la distorsion subie par le signal d'entrée. L'entraînement des valeurs des cartes de caractéristiques et des paramètres du réseau de neurones peut être conjoint. Avantagusement, le procédé de codage selon l'invention permet de compresser efficacement le signal.
- [0025] Avantagusement, le procédé de décodage est simple puisqu'il suffit de décoder les cartes de caractéristiques et le réseau de neurones pour reconstituer une version décodée du signal, par exemple une image.
- [0026] Un tel réseau de neurones peut avantagusement être de structure très simple avec peu de paramètres.
- [0027] De surcroît, le décodage peut être effectué échantillon par échantillon, de manière progressive.
- [0028] Selon des modes de réalisation du procédé de codage et/ou de décodage :
- [0029] - Le procédé comporte une étape de construction d'un second groupe de cartes de caractéristiques et le vecteur caractéristique est en outre construit à partir desdites cartes

de caractéristiques. Avantageusement, ces cartes supplémentaires du second groupe, construites de manière identique au codeur et au décodeur, ne sont ni stockées ni transmises au niveau du codeur, ni décodées au niveau du décodeur. Elles permettent ainsi de bénéficier de données supplémentaires pour améliorer la compression sans dégrader le débit. Elles peuvent par exemple comprendre des coordonnées, des données disponibles dans les cartes du premier groupe, des données concernant d'autres images déjà traitées par le codeur ou décodeur, etc.

- [0030] – Une au moins des cartes de caractéristiques est de résolution inférieure à celle du signal à coder (resp. décoder). Avantageusement selon ce mode, la compression des cartes de caractéristiques est plus efficace puisque l'une au moins des cartes de caractéristiques comporte moins de valeurs que si elle était à la résolution du signal. Par exemple, dans le cas d'une image numérique, l'une des cartes de caractéristiques peut être à la résolution 1/2, c'est-à-dire qu'elle comporte deux fois moins de valeurs en abscisse et en ordonnée que le signal d'entrée comporte d'échantillons, soit au total 4 fois moins de valeurs qu'une carte de caractéristiques à la résolution du signal.
- [0031] – Une au moins des dites cartes de caractéristiques est de résolution identique à celle du signal à coder (resp. décoder). Avantageusement selon ce mode, l'une au moins des cartes de caractéristiques, à la même résolution que le signal d'entrée (resp. à décoder), permet une fidélité importante et le respect des détails de la résolution initiale du signal.
- [0032] – La construction dudit vecteur caractéristique comporte une sous-étape d'extraction d'une valeur de ladite au moins une carte de caractéristiques à une position identique à celle de l'échantillon courant dans le signal.
- [0033] Avantageusement, il est ainsi possible d'extraire une valeur d'une carte de caractéristique du premier ou du second groupe, à la même position que l'échantillon dans le signal (signal d'entrée pour le codage, signal à décoder pour le décodage), pour constituer un élément du vecteur caractéristique. Ce procédé est simple à mettre en œuvre. Par exemple, si l'on dispose de J cartes de caractéristiques en entrée, de même résolution que le signal, une simple extraction des valeurs des cartes aux coordonnées de l'échantillon courant (à la même abscisse et la même coordonnée dans la carte de caractéristiques) permet de construire directement le vecteur caractéristique de J éléments.
- [0034] – La construction dudit vecteur caractéristique comporte les sous-étapes suivantes :
- extraction d'une pluralité de valeurs des cartes de caractéristiques en fonction de ladite position dudit échantillon courant ;
 - traitement des dites valeurs extraites pour obtenir le vecteur caractéristique.
- [0035] Avantageusement, il est possible de traiter les valeurs extraites des cartes de caractéristiques pour construire le vecteur caractéristique. Un tel traitement peut par exemple

correspondre à une quantification des données extraites, ou à une mise à l'échelle, ou à un filtrage, etc. Au codage, la quantification est indispensable au bon fonctionnement du système si les cartes de caractéristiques comportent par exemple des valeurs flottantes, ou réelles. Il est nécessaire de les quantifier avant de les coder et/ou de les fournir en entrée du réseau de neurones. Au décodage en revanche, la quantification inverse n'est pas nécessaire, selon les modes de réalisation.

[0036] - Le codage (resp. décodage) dudit premier groupe de cartes de caractéristiques comporte une sous-étape de codage (resp. décodage) entropique. Avantageusement, le codage entropique permet d'exploiter les redondances du signal qui peuvent persister dans les cartes de caractéristiques, le signal étant ainsi compressé plus efficacement.

[0037] Corrélativement, l'invention vise aussi un dispositif de codage et un dispositif de décodage.

[0038] Les caractéristiques et avantages du procédé de codage ou décodage s'appliquent de la même façon au dispositif de codage ou décodage selon l'invention et vice versa.

[0039] L'invention vise également un programme d'ordinateur sur un support d'enregistrement, ce programme étant susceptible d'être mis en œuvre dans un ordinateur ou un dispositif de contrôle conforme à l'invention. Ce programme comporte des instructions adaptées à la mise en œuvre du procédé correspondant. Ce programme peut utiliser n'importe quel langage de programmation, et être sous la forme de code source, code objet, ou de code intermédiaire entre code source et code objet, tel que dans une forme partiellement compilée, ou dans n'importe quelle autre forme souhaitable.

[0040] L'invention vise aussi un support d'information ou un support d'enregistrement lisible par un ordinateur, et comportant des instructions de programme d'ordinateur mentionné ci-dessus. Les supports d'information ou d'enregistrement peuvent être n'importe quel entité ou dispositif capable de stocker les programmes. Par exemple, les supports peuvent comporter un moyen de stockage, tel qu'une ROM, par exemple un CD-ROM ou une ROM de circuit microélectronique, ou encore un moyen d'enregistrement magnétique, par exemple une disquette ou un disque dur, une séquence d'ADN, ou une mémoire flash. D'autre part, les supports d'information ou d'enregistrement peuvent être des supports transmissibles tels qu'un signal électrique ou optique, qui peut être acheminé via un câble électrique ou optique, par lien radio, par lien optique sans fil ou par d'autres moyens.

[0041] Le programme selon l'invention peut être en particulier téléchargé sur un réseau de type Internet.

[0042] Alternativement, chaque support d'informations ou d'enregistrement peut être un circuit intégré dans lequel le programme est incorporé, le circuit étant adapté pour exécuter ou pour être utilisé dans l'exécution d'un procédé conforme à l'invention.

Breve description des dessins

- [0043] D'autres caractéristiques et avantages de la présente invention ressortiront de la description faite ci-dessous, en référence aux dessins annexés qui en illustrent des exemples de réalisation dépourvus de tout caractère limitatif.
- [0044] [Fig.1] La [Fig.1] représente schématiquement un dispositif de codage utilisé dans le cadre de l'invention ;
- [0045] [Fig.2] la [Fig.2] représente schématiquement un dispositif de décodage utilisé dans le cadre de l'invention ;
- [0046] [Fig.3] la [Fig.3] illustre un exemple de réseau de neurones artificiels de synthèse utilisé dans le cadre de l'invention au codage et au décodage ;
- [0047] [Fig.4] la [Fig.4] est un logigramme représentant un exemple de procédé de codage qui peut être mis en œuvre par le dispositif de codage de la [Fig.1] ;
- [0048] [Fig.5] la [Fig.5] représente de manière illustrée un procédé de codage utilisé dans un mode de réalisation de l'invention ;
- [0049] [Fig.6] la [Fig.6] représente de manière illustrée un procédé de codage utilisé dans un autre mode de réalisation de l'invention ;
- [0050] [Fig.7] la [Fig.7] est un logigramme représentant un exemple de procédé de décodage qui peut être mis en œuvre par le dispositif de décodage de la [Fig.2] ;
- [0051] [Fig.8] la [Fig.8] représente de manière illustrée un procédé de décodage utilisé dans un mode de réalisation de l'invention.

Description des modes de réalisation

- [0052] La [Fig.1] représente schématiquement un dispositif de codage ENC.
- [0053] Ce dispositif de codage ENC comprend un module GEN de génération de cartes de caractéristiques, un module TT de traitement et quantification, un module XTR d'extraction de données, un module MLP correspondant à un réseau de neurones artificiels, un module NNC de codage de réseau de neurones, un module FMC de codage des cartes de caractéristiques, un module EVAL d'évaluation d'une fonction de performance, un module MAJ de mise à jour, un module FME optionnel de génération de cartes de caractéristiques supplémentaires.
- [0054] Le dispositif de codage ENC peut être mis en œuvre au moyen d'un dispositif électronique comprenant un processeur et une mémoire, non représentés ; chacun des modules mentionnés ci-dessus peut alors être réalisé par la coopération du processeur et d'instructions de programme d'ordinateur mémorisées dans la mémoire susmentionnée et conçues pour effectuer les fonctionnalités du module concerné, notamment comme décrit ci-dessous, lorsque ces instructions sont exécutées par le processeur.
- [0055] Le dispositif de codage ENC de la [Fig.1] reçoit en entrée une succession

d'échantillons à coder, notés P_n , par exemple une succession temporelle d'échantillons sonores, ou un ensemble de données d'image noté $I(P_n)$. Dans ce second cas, le signal d'image $I(P_n)$ peut représenter une image bidimensionnelle, ou une pluralité d'images bidimensionnelles (vidéo, composantes de couleur, composantes stéréoscopiques, multiscopiques, etc.). P_n désigne un échantillon n du signal d'entrée comportant N échantillons. Dans un mode de réalisation, le signal est un signal d'image couleur représentée au moyen d'au moins une représentation bidimensionnelle, telle qu'une matrice de pixels, de largeur W et de hauteur H (avec $N = W \times H$) chaque pixel comportant une composante rouge (R), verte (G), bleue (B), ou, en variante, une composante de luminance (Y) et au moins une composante de chrominance. L'emplacement de chaque pixel est défini par ses coordonnées en abscisse et ordonnée (x et y) dans l'image. Dans un mode de réalisation, l'image est en niveaux de gris représentée au moyen d'une représentation bidimensionnelle, telle qu'une matrice de pixels, chaque pixel comportant une composante de niveaux de gris, ou luminance. Dans ce cas le vecteur représentatif du pixel est réduit à une seule composante, ou donnée.

[0056] Comme cela sera décrit plus en détail plus loin en référence aux figures 3 à 8 :

[0057] Le module GEN de génération de cartes de caractéristiques est configuré pour générer, et notamment initialiser, une pluralité de M cartes de caractéristiques notées FM_i . Le module FME facultatif peut générer une ou plusieurs cartes supplémentaires, au nombre de L , qui ne seront ni codées ni transmises, notées FME_i .

[0058] Le module XTR réalise une extraction de valeurs dans les cartes de caractéristiques FM_i et optionnellement FME_i , pour un échantillon P_n à coder, en fonction de ses coordonnées dans le signal d'entrée.

[0059] Le module TT réalise un traitement des valeurs extraites, pour générer le vecteur Z_n . Z_n est un J -uplet, c'est-à-dire qu'il comporte J éléments, ou données z_i . Le vecteur Z_n d'indice n réfère au vecteur caractéristique du pixel P'_n .

[0060] Le module TT réalise une quantification des données extraites de cet ensemble de M cartes FM_i , ou du vecteur Z_n constitué à partir de ces données. Pour préparer le codage ultérieur (par exemple de type entropique), le système de codage quantifie la représentation latente des données d'entrée en utilisant un quantificateur Q pour générer une collection ordonnée de valeurs quantifiées. On rappelle que la quantification d'une valeur fait référence à la mise en correspondance de cette valeur avec un membre d'un ensemble discret de symboles de code possibles. Par exemple, l'ensemble de symboles de code possibles peut être constitué de valeurs entières, et le système de quantification réalise un simple arrondi d'une valeur réelle à une valeur entière. Selon un autre exemple, la quantification consiste en une multiplication par une valeur donnée puis un arrondi.

- [0061] Le module MLP est un réseau de neurones défini par K paramètres W_k , apte à traiter le vecteur Z_n , ou J-Uplet, en entrée, pour générer en sortie un second vecteur représentatif de l'échantillon P_n à coder. Le réseau de neurones est selon un mode de réalisation un MLP, ou Multi Layer Perceptron, constitué d'une couche d'entrée adaptée au format d'entrée (le J-uplet), optionnellement une ou plusieurs couche(s) cachée(s), et une couche de sortie adaptée au format de sortie du vecteur de sortie, de manière générale un vecteur comportant A éléments. Selon un mode de réalisation, A est égal à 3 et le vecteur de sortie est le triplet (R,G,B) du pixel P'_n codé puis décodé.
- [0062] Le module NNC réalise le codage du réseau de neurones, notamment de ses paramètres W_k . Durant le processus d'entraînement, ou de construction, du codage, c'est-à-dire tant que l'étape d'évaluation d'une performance n'est pas satisfaisante, le module NNC réalise une simulation de codage, suivi d'un décodage, à destination du module d'évaluation. Il met à jour les valeurs des paramètres en fonction des résultats d'une mesure de performance réalisée par le module EVAL. Par la suite, il effectue le codage effectif des paramètres du réseau de neurones W_k . Les paramètres codés sont notés W_{c_k} . De manière connue, la simulation de codage peut être identique au codage effectif, ou en réaliser une approximation.
- [0063] Le module FMC réalise le codage des cartes FM_i , c'est-à-dire des valeurs quantifiées de chaque carte du premier groupe (à l'exclusion des cartes FME_1). Durant le processus d'entraînement, ou de construction, du codage, c'est-à-dire tant que l'étape d'évaluation d'une performance n'est pas satisfaisante, le module FMC réalise une simulation de codage, suivi d'un décodage, à destination du module d'évaluation. Par la suite, il effectue le codage effectif des valeurs des cartes FM_i . Les cartes codées sont notées FM_{c_i} . De manière connue, la simulation de codage peut être identique au codage effectif, ou en réaliser une approximation. Le module de codage compresse les données quantifiées en utilisant par exemple un codage de type entropique.
- [0064] Le module EVAL réalise une évaluation et minimisation d'une performance de codage. La fonction d'évaluation est par exemple de type débit-distorsion. La minimisation peut être effectuée par une descente de gradient, ou tout autre procédé à la portée de l'homme du métier.
- [0065] Le module MAJ réalise une mise à jour des valeurs des cartes FM_i à encoder en fonction des résultats de la fonction de performance. Optionnellement il peut effectuer une mise à jour des valeurs des cartes supplémentaires FME_1 .
- [0066] La [Fig.2] représente schématiquement un dispositif de décodage DEC.
- [0067] Le dispositif de codage DEC de la [Fig.2] reçoit en entrée un premier groupe de données encodées organisées en M cartes de caractéristiques FM_{c_i} (dites aussi couches FM) et les paramètres encodés W_{c_k} du réseau de neurones
- [0068] Ce dispositif de décodage DEC comprend un module NND de décodage de réseau de

neurones, un module FMD de décodage des cartes de caractéristiques, un module XTR' d'extraction de données, un module TT' de traitement inverse, un module MLP' correspondant à un réseau de neurones, un module FME' optionnel de génération de cartes de caractéristiques supplémentaires.

- [0069] Les cartes décodées par le module FMD, au nombre de M , sont notées FMd_i . Les paramètres décodés par le module NND sont notés Wd_k .
- [0070] Le module FME' du décodeur peut aussi générer une ou plusieurs cartes supplémentaires, notées FME'_i , au nombre de L , identiques aux cartes supplémentaires FME_i générées par le codeur.
- [0071] Le module XTR' est identique au module XTR de la [Fig.1]. Il réalise une extraction de valeurs des M cartes de caractéristiques FMd_i décodées, et éventuellement des L cartes supplémentaires FME'_i , pour construire le vecteur Z_n prenant la forme d'un J -uplet pour un échantillon P_n à décoder, en fonction de ses coordonnées dans le signal à décoder. Dans un mode de réalisation, $J=M$. Dans un mode de réalisation, $J=M+L$.
- [0072] Le module TT' effectue optionnellement une quantification inverse correspondant à la quantification effectuée au codeur. La quantification inverse n'est pas nécessaire si le quantificateur Q du codeur a réalisé un simple arrondi des valeurs réelles qui lui sont soumises. La quantification inverse n'est pas nécessaire si le réseau de neurones est apte à prendre en compte une quantification de ses données d'entrée. Sinon, le décodeur réalise l'opération inverse du quantificateur Q . Par ailleurs, comme le module TT du codeur, le module TT' réalise éventuellement un traitement des valeurs extraites des cartes de caractéristiques, pour générer le vecteur décodé Zd_n .
- [0073] Le module MLP' est un réseau de neurones défini par K paramètres Wd_k , apte à traiter le vecteur Zd_n , ou J -Uplet, en entrée, pour générer en sortie un second vecteur représentatif de l'échantillon P_n à décoder, de manière générale un vecteur comportant A éléments. Selon un mode de réalisation, $A=3$ et le vecteur de sortie est le triplet (R,G,B) du pixel Pd_n décodé. Le module MLP' est de structure identique au module MLP, et ses paramètres sont soit identiques si le codage de ses paramètres W_k est sans perte, soit différents si le codage est réalisé avec pertes.
- [0074] Lorsque tous les échantillons P_n du signal ont été décodés, on dispose d'un signal reconstruit $I(Pd_n)$, selon un exemple une image I comportant N échantillons décodés sous la forme de N vecteurs Pd_n .
- [0075] Le dispositif de décodage DEC peut être mis en œuvre au moyen d'un dispositif électronique comprenant un processeur et une mémoire, non représentés ; chacun des modules mentionnés ci-dessus peut alors être réalisé par la coopération du processeur et d'instructions de programme d'ordinateur mémorisées dans la mémoire susmentionnée et conçues pour effectuer les fonctionnalités du module concerné, notamment comme décrit ci-dessous, lorsque ces instructions sont exécutées par le

processeur.

- [0076] la [Fig.3] illustre un exemple de réseau de neurones artificiels de synthèse utilisé au codage et au décodage dans le cadre de modes de réalisation de l'invention.
- [0077] Le réseau de neurones artificiels de synthèse MLP au codage et le réseau de neurones artificiels de synthèse MLP' au décodage sont définis par une structure identique, comprenant par exemple une pluralité de couches de neurones artificiels, et par un ensemble de poids et fonctions d'activation associés respectivement aux neurones artificiels du réseau concerné.
- [0078] Une représentation vectorielle d'un échantillon courant (un vecteur Z_n ou Zd_n issu des cartes de caractéristiques FM_i et FME_i ou FMd_i et FME'_i) est appliquée en entrée (c'est-à-dire sur une couche d'entrée) du réseau de neurones artificiels de synthèse MLP ou MLP'. Le réseau de neurones artificiels produit en sortie un vecteur P'_n ou Pd_n représentatif de l'échantillon décodé, selon un mode de réalisation les composantes de couleur (R, G, B) constitutives d'un pixel couleur d'une image.
- [0079] La concaténation de tous ces pixels reconstruits dans une image (2D, 3D) constitue l'image décodée, ou reconstruite.
- [0080] Au codeur, le réseau de neurones artificiels de synthèse MLP est entraîné sur l'image, de sorte à minimiser les différences entre la représentation en entrée de l'image courante $I(P_n)$ et sa représentation $I(P'_n)$ en sortie, tout en minimisant également la quantité de données à encoder. Le module EVAL effectue une mesure de performance en ce sens.
- [0081] Une fois que l'entraînement est terminé, les paramètres du réseau sont encodés, soit sans pertes, auquel cas le réseau de neurones MLP' est identique à MLP, soit avec pertes, auquel cas le réseau MLP' peut être légèrement différent de MLP.
- [0082] La [Fig.4] est un logigramme représentant un exemple de procédé de codage qui peut être mis en œuvre par le dispositif de codage de la [Fig.1].
- [0083] Selon ce mode de réalisation, le signal est une image bidimensionnelle, chaque échantillon à coder est donc un pixel P_n de coordonnées (x_n, y_n) .
- [0084] L'encodage se déroule en deux phases principales :
- [0085] Dans une première phase, dite phase de construction, un apprentissage est réalisé, afin de déterminer, pour un signal d'entrée $I(P_n)$, les valeurs des cartes FM_i et des paramètres W_k pour optimiser une fonction de coût globale. L'apprentissage est par exemple réalisé par une descente de gradient, suivie d'une mise à jour des paramètres du réseau de neurones MLP et des valeurs des cartes de caractéristiques FM_i . Comme il est connu dans l'état de l'art, la fonction de coût peut être de type débit-distorsion, ou débit, ou distorsion, ou perceptuelle. Pour mesurer le débit R, il est nécessaire de simuler le codage des cartes FM_i , puis de mesurer le débit de codage associé (la taille du flux B1). Selon un mode de réalisation, on ne simule pas le codage des paramètres

W_k car leur influence est moins importante que celle des cartes de caractéristiques. Selon un mode de réalisation, on simule aussi le codage des paramètres W_k et on mesure le débit associé (la taille du flux B2). Pour mesurer la distorsion D, il est nécessaire de simuler le codage puis le décodage d'une partie au moins de l'image I, pour obtenir au moins un pixel P'_n résultant d'une simulation de codage puis de décodage, puis de mesurer l'écart entre cette partie de l'image $I(P_n)$ en entrée et une partie correspondante de l'image $I(P'_n)$ codée puis décodée.

- [0086] Puis lors d'une deuxième phase, dite phase de codage, les cartes FM_i et les paramètres W_k sont encodés pour produire les valeurs codées FMc_i et Wc_k avant transmission ou stockage. Ils constituent la représentation compressée du signal d'entrée $I(P_n)$.
- [0087] On va maintenant décrire les étapes d'un procédé selon un mode de réalisation de l'invention.
- [0088] Lors d'une étape E20, un signal $I(P_n)$ à coder, comportant une pluralité de N échantillons P_n , est fourni en entrée du procédé.
- [0089] Lors d'une étape E21, les M cartes FM_i du premier groupe sont initialisées par le module GEN. Par la suite, les paramètres W_k du réseau de neurones MLP et les valeurs des cartes FM_i doivent être optimisés durant la phase de construction.
- [0090] Selon un mode de réalisation, les cartes FM_i sont de même résolution que le signal d'entrée $I(P_n)$ et comportent donc chacune le même nombre de valeurs N qu'il y a d'échantillons P_n à coder.
- [0091] Selon un mode de réalisation, les cartes FM_i sont de résolution inférieure ou égale à celle du signal d'entrée $I(P_n)$ et comportent donc, pour au moins l'une d'entre elles, un nombre N' de valeurs à coder inférieur à N.
- [0092] Selon un mode de réalisation, la première carte FM_i est à la résolution du signal et chaque carte suivante est à une résolution moitié de la précédente.
- [0093] Selon un mode de réalisation, plusieurs cartes FM_i sont de même résolution, inférieure à la résolution du signal.
- [0094] Selon un mode de réalisation, les cartes FM_i sont initialisées par des valeurs constantes prédéfinies.
- [0095] Selon un autre mode de réalisation, les cartes de caractéristiques sont initialisées par un ensemble de valeurs réelles aléatoires.
- [0096] Selon un mode de réalisation, une ou plusieurs cartes FME_i , formant un second groupe de L cartes de caractéristiques supplémentaires, sont générées et ajoutées au premier groupe. Elles servent à la construction du vecteur caractéristique mais ne seront ni stockées ni transmises.
- [0097] Ces cartes de caractéristiques FM_i du premier groupe sont par la suite mises à jour, ou raffinées, lors d'une étape E22 par le module de mise à jour MAJ du codeur au

cours de son apprentissage lors de la phase de construction.

- [0098] Lors d'une étape E23, les cartes FM_i du premier groupe sont codées par le module FMC du codeur. Durant la phase de construction, cette opération est une simulation de codage. Durant la phase de codage, cette opération est un codage effectif et les valeurs codées constituent le flux B1. La simulation peut être identique au codage effectif mais elle peut aussi être différente (par exemple, simplifiée). Pour ce codage, on peut utiliser toute technique connue visant à compresser les valeurs des cartes.
- [0099] Dans un mode de réalisation, les cartes FM_i sont codées dans l'ordre (FM_1, FM_2, \dots, FM_4), et les variables de chaque carte dans un ordre prédéfini, par exemple lexicographique. Chaque carte subit un codage entropique. Le codage entropique produit un flux compressé B1 dont le débit est mesuré ultérieurement au cours d'une étape E28.
- [0100] Lors d'une étape E24, des valeurs sont extraites par le module XTR des cartes FM_i et éventuellement FME_i . Cette extraction est réalisée en fonction des coordonnées (x_n, y_n) de l'échantillon P_n du signal d'entrée.
- [0101] Selon un mode de réalisation, les valeurs de coordonnées (x_n, y_n) de chaque carte sont extraites.
- [0102] Selon un mode de réalisation, les valeurs de coordonnées sont mises à l'échelle pour procéder à l'extraction des couches de résolutions inférieures.
- [0103] Selon un mode de réalisation, des valeurs voisines de celles précédemment citées sont également extraites pour le traitement ultérieur de l'étape E25.
- [0104] Lors d'une étape E25, un vecteur Z_n est construit par le module TT à partir des valeurs extraites des cartes FM_i et éventuellement FME_i pour chaque échantillon P_n de coordonnées (x_n, y_n) du signal d'entrée. Le traitement comporte une quantification des valeurs extraites des cartes FM_i , si nécessaire. Le traitement peut comporter d'autres opérations, comme un filtrage, une mise à l'échelle, l'application d'une fonction quelconque, de préférence monotone, etc.
- [0105] Dans un mode de réalisation, Z_n comporte autant de valeurs que de cartes ou cartes FM_i (et optionnellement FME_i) en entrée. Dans ce cas on a $J=M(+L)$.
- [0106] Dans un mode de réalisation, Z_n est un J-uplet (z_1, z_2, \dots, z_J) , constitué des valeurs des cartes FM_i (et optionnellement FME_i) situées aux coordonnées (x_n, y_n) du pixel courant P_n , comme il sera illustré à l'appui de la [Fig.5]. La valeur située aux coordonnées (x_n, y_n) de la carte FM_i donne la valeur finale (z_i) de l'élément i du vecteur Z_n relative à cette carte FM_i ou FME_i . Dans ce cas le module TT n'applique aucun traitement autre que la quantification.
- [0107] Dans un mode de réalisation, Z_n est un J-uplet construit à partir de valeurs prélevées dans les cartes FM_i (et optionnellement FME_i) à des coordonnées qui peuvent être différentes selon les cartes. Par exemple, si les cartes FM_i (et/ou FME_i) sont à des résolutions différentes parce qu'elles ont été sous-échantillonnées, les coordonnées sont

adaptées (par une mise à l'échelle) pour correspondre à la résolution de chaque carte, comme il sera illustré à l'appui de la [Fig.6].

- [0108] Dans un mode de réalisation, Z_n est un J-uplet construit à partir de valeurs prélevées dans les cartes FM_i (et FME_i) en appliquant le traitement à une ou plusieurs valeurs des cartes, par exemple un filtrage des valeurs voisines de la valeur ciblée dans une carte. Par exemple, dans une carte FM_i qui se trouve à la même résolution que le signal d'entrée, on peut extraire les valeurs situées aux coordonnées (x_n, y_n) , (x_n-1, y_n) , $(x_n, y_n - 1)$ et (x_n-1, y_n-1) et appliquer un traitement à ces valeurs (filtrages, moyennage, interpolation, etc.) pour obtenir la valeur finale (z_i) de l'élément i du vecteur Z_n relative à cette carte FM_i ou FME_i . Selon un autre exemple, dans une carte FM_i qui se trouve à une résolution moitié du signal d'entrée, on peut considérer les valeurs situées aux coordonnées $(x_n/2, y_n/2)$, $(x_n/2-1, y_n/2)$, $(x_n/2, y_n/2-1)$ et $(x_n/2-1, y_n/2-1)$ et appliquer un traitement à ces valeurs (filtrages, moyennage, interpolation, etc.) pour obtenir la valeur finale (z_i) de l'élément i du vecteur Z_n relative à cette carte FM_i ou FME_i .
- [0109] Les échantillons à coder sont par exemple traités par ordre séquentiel, de $n=1$ à $n=N$.
- [0110] Lors d'une étape E26, le vecteur Z_n est traité par le réseau de neurones MLP pour générer en sortie un vecteur représentatif de l'échantillon P_n à coder, selon un mode de réalisation le triplet (R, G, B) de l'échantillon P'_n (l'échantillon P_n codé puis décodé).
- [0111] La structure et les paramètres W_k du réseau de neurones sont initialisés par exemple lors de la première itération de cette étape. Ces paramètres sont par la suite mis à jour, ou raffinés, au cours de la phase de construction, lors des itérations ultérieures du procédé.
- [0112] Selon un mode de réalisation, les paramètres du réseau de neurones sont initialisés par des valeurs prédéfinies connues pour donner un résultat satisfaisant (par exemple, suite à un entraînement sur un corpus d'images).
- [0113] Selon un autre mode de réalisation, les paramètres W_k du réseau de neurones sont initialisés par un ensemble de valeurs aléatoires.
- [0114] Lors d'une étape E27, les paramètres W_k du réseau de neurones MLP sont quantifiés et codés. Durant la phase de construction, cette opération est une simulation de codage. Durant la phase de codage, cette opération est un codage effectif et les valeurs codées constituent le flux B2. La simulation peut être identique au codage effectif mais elle peut aussi être différente (par exemple, simplifiée). On peut utiliser à cette fin toute technique connue, par exemple la norme de codage de réseaux de neurones proposée par le standard MPEG-7 partie 17, aussi appelée NNR (Neural Network Representation). On notera que dans ce cas, il faut choisir la quantité de dégradation que le codage apporte aux paramètres W_k .
- [0115] Lors d'une étape E28, une mesure de performance est évaluée.
- [0116] A cet effet, les débits de simulation de codage associés aux cartes de caractéristiques

(simulation du flux B1 par codage et décodage des cartes FM_i) et optionnellement aux paramètres du réseau de neurones (simulation du flux B2 par codage et décodage des paramètres W_k) sont mesurés.

- [0117] Selon un mode de réalisation, la fonction de coût est de type débit-distorsion, notée $(D+L*R)$, où D est par exemple l'erreur quadratique mesurée entre le signal d'entrée et le signal décodé (ou l'erreur mesurée sur un sous-ensemble d'échantillon du signal). Selon un autre exemple, D est calculé à partir d'une fonction perceptive telle que le SSIM (pour *Structural SIMilarity*), ou MSSSIM (pour *Multi-scale Structural SIMilarity*). Selon un mode de réalisation, R est le débit simulé du flux B1 ; selon un autre mode de réalisation, R est le débit total utilisé pour coder cette image, c'est-à-dire la somme des débits simulés de B1 et B2. L est un paramètre qui règle le compromis débit-distorsion. D'autres fonctions de coût sont possibles.
- [0118] Tant que la fonction de coût n'a pas atteint son minimum, la mesure de performance n'est pas satisfaisante, et le procédé est réitéré à partir de l'étape E22. Cette minimisation peut être effectuée par un mécanisme connu comme une descente de gradient avec mise à jour des paramètres au cours de l'étape E22 pour les valeurs des cartes de caractéristiques et E26 pour les paramètres du réseau.
- [0119] Lors d'une étape E29, si la fonction de coût a atteint son minimum, l'entraînement s'arrête. Si une version codée correspondant à la dernière simulation des paramètres du réseau de neurones (W_k) et des cartes de caractéristiques (FM_i) est disponible, les flux B1 et B2 peuvent en être constitués. Selon un autre mode de réalisation, le codage effectif des paramètres mis à jour du réseau de neurones (W_k) et des valeurs des cartes de caractéristiques (FM_i) est effectué à cette étape pour produire les paramètres encodés Wc_k et FMc_i qui constituent des flux B1 et B2.
- [0120] Les flux B1 et B2 peuvent être concaténés pour produire un flux final BS. Selon un mode de réalisation, le flux B2 des paramètres codés du réseau de neurones est stocké ou transmis avant le flux B1, afin de pouvoir être décodé avant le flux B1.
- [0121] La [Fig.5] représente de manière illustrée un procédé de codage utilisé dans un mode de réalisation de l'invention pour coder une image $I(P_n)$ (non représentée).
- [0122] Dans ce mode de réalisation, les cartes FM_i sont au nombre de 4. Dans un mode préféré elles sont au nombre de 7. Les cartes FM_i sont de même résolution que l'image $I(P_n)$, et comportent donc chacune $W \times H$ valeurs, où W représente la largeur de l'image en pixels et H sa hauteur ($N=W \times H$).
- [0123] Selon une variante présentée en pointillés, les cartes de caractéristiques sont au nombre de 5 : une carte supplémentaire FME_0 a été introduite, qui ne sera ni stockée ni transmise. Cette carte supplémentaire comporte typiquement des données qui peuvent assister le réseau MLP dans la tâche de reconstruction du signal. Ainsi, les cartes ajoutées peuvent être une ou plusieurs parmi la liste suivante, non limitative :

- [0124] – Une carte comportant en chaque point l'abscisse de ce point.
- [0125] – Une carte comportant en chaque point l'ordonnée de ce point.
- [0126] – Une carte comportant en chaque point un codage positionnel (comme décrit par exemple à l'adresse internet <https://skosmos.loterre.fr/P66/fr/page/-K0D65X2X-X>).
- [0127] – Une carte représentant une image distincte de l'image en cours de traitement, susceptible d'apporter des informations sur l'image à coder, par exemple l'image précédemment traitée si l'image courante fait partie d'une série d'images à coder comme une vidéo, un ensemble d'images médicales, une représentation multivues, etc.
- [0128] – Une carte représentant une carte de caractéristiques d'une image distincte de l'image en cours de traitement susceptible d'apporter des informations sur l'image à coder, par exemple une carte précédemment traitée si l'image courante fait partie d'une série d'images à coder comme une vidéo, un ensemble d'images médicales, une représentation multivues, etc.
- [0129] – Une carte comportant la valeur d'un échantillon déjà décodé de la même carte, par exemple l'échantillon précédent dans l'ordre de décodage.
- [0130] Dans ce mode de réalisation, le vecteur Z_n est un 4-uplet $(z_1...z_4)$ constitué des valeurs extraites des cartes FM_i situées aux coordonnées (x_n, y_n) du pixel courant P_n . Les valeurs extraites sont quantifiées par le module TT et le vecteur Z_n constitué des valeurs quantifiées est traité par le réseau de neurones MLP pour générer en sortie un second vecteur, selon l'exemple un triplet (R, G, B) représentatif de l'échantillon P_n à coder. Le vecteur de sortie est dans ce cas le triplet (R,G,B) du pixel P'_n codé puis décodé. Le triplet est inséré dans l'image décodée $I(P'_n)$ aux positions (x_n, y_n) des composantes couleur (R', G', B') .
- [0131] Dans un mode de réalisation, les 4 cartes FM_i sont quantifiées.
- [0132] Selon la variante présentée en pointillés, le vecteur Z_n est un 5-uplet $(z_0...z_4)$, la valeur z_0 étant extraite de la carte supplémentaire FME_0 .
- [0133] La [Fig.6] représente de manière illustrée un procédé de codage utilisé dans un autre mode de réalisation de l'invention.
- [0134] Dans ce mode de réalisation, les cartes FM_i sont au nombre de 4. De même que dans le mode précédent, la première carte FM_1 a la même résolution que l'image I, et comporte donc $W \times H$ valeurs, où W représente la largeur de l'image en pixels, et H sa hauteur. La seconde carte FM_2 est de résolution moitié (dans chaque dimension) de la carte FM_1 . Chaque carte supplémentaire est de résolution moitié de la carte précédente. Cette structure permet de réduire le nombre de variables des cartes de caractéristiques ce qui facilite le codage et l'apprentissage tout en minimisant le coût de codage.
- [0135] Dans ce mode de réalisation, un 4-uplet Z_n est construit à partir des valeurs des cartes FM_i situées respectivement aux coordonnées (x_n, y_n) , $(x_n/2, y_n/2)$, $(x_n/4, y_n/4)$, $(x_n/8, y_n/8)$ du pixel courant $P_n(x_n, y_n)$. Ces valeurs extraites sont traitées et quantifiées pour

générer le 4-uplet Z_n . Selon un mode de réalisation, pour la carte FM_2 , on extrait la valeur située aux coordonnées $(x_n/2, y_n/2)$ pour obtenir la valeur finale (z_2) du vecteur Z_n relative à cette carte FM_2 . Selon un autre mode de réalisation, pour la carte FM_2 , on extrait les valeurs situées aux coordonnées $(x_n/2, y_n/2)$, $(x_n/2-1, y_n/2)$, $(x_n/2, y_n/2-1)$, $(x_n/2-1, y_n/2-1)$ et on leur applique un filtrage pour obtenir la valeur finale (z_2) du vecteur Z_n relative à cette carte FM_2 . Un suréchantillonnage est réalisé de cette manière. Le filtrage ou suréchantillonnage peut par exemple utiliser un filtrage bi-cubique ou filtrage de Lanczos, connus en soi. On peut également utiliser un traitement plus général : un réseau de neurones peut être utilisé pour produire la valeur finale (z_2) par pondération des valeurs situées aux coordonnées $(x_n/2, y_n/2)$, $(x_n/2-1, y_n/2)$, $(x_n/2, y_n/2-1)$, $(x_n/2-1, y_n/2-1)$, par exemple un réseau de type MLP.

- [0136] Comme précédemment, le vecteur Z_n est traité par le réseau de neurones MLP pour générer en sortie un second vecteur, ici un triplet (R, G,B) représentatif de l'échantillon P_n à coder. Le vecteur de sortie est dans ce cas le triplet (R,G,B) du pixel P'_n codé puis décodé. Le triplet (R,G,B) est réinjecté dans l'image décodée aux positions (x_n, y_n) des composantes couleur (R', G', B') de l'image $I(p'_n)$.
- [0137] D'autres types de structure sont possibles, par exemple on peut utiliser un taux de réduction différent de un demi entre les cartes (un quart, ou un tiers, etc.) L'étape de suréchantillonnage est alors adaptée pour produire la valeur finale du vecteur Z_n .
- [0138] Comme il a été décrit à l'appui de la [Fig.5], d'autres cartes FME_i peuvent être introduites, qui serviront à la génération du vecteur Z_n mais ne seront ni stockées ni transmises.
- [0139] La [Fig.7] est un logigramme représentant un exemple de procédé de décodage qui peut être mis en œuvre par le dispositif de décodage de la [Fig.2].
- [0140] Lors d'une étape E30, les flux B1 et B2 sont extraits du flux encodé BS. Ils contiennent respectivement les représentations codées des cartes du premier groupe FMc_i et des paramètres Wc_k .
- [0141] Lors d'une étape E31, les M cartes FMd_i sont générées par décodage des valeurs FMc_i . Pour ce décodage, on peut utiliser toute technique connue similaire à celle utilisée au codeur, de préférence un décodage entropique. Dans un mode de réalisation, les cartes FMd_i sont décodées dans l'ordre ($FMd_1, FMd_2, \dots FMd_4$), et les variables de chaque carte dans un ordre prédéfini, par exemple lexicographique.
- [0142] Selon un mode de réalisation tel que décrit pour le codeur, les cartes FMd_i sont de même résolution que le signal $I(Pd_n)$ à reconstituer, c'est-à-dire qu'elles comportent N valeurs.
- [0143] Selon un mode de réalisation tel que décrit pour le codeur, les cartes FMd_i sont à une résolution inférieure ou égale à celle du signal $I(Pd_n)$ à reconstituer.
- [0144] Selon un mode de réalisation tel que décrit pour le codeur, plusieurs cartes FMd_i sont

de même résolution, inférieure à la résolution du signal.

- [0145] Lors d'une étape E32, selon un mode de réalisation, une ou plusieurs cartes FME'_1 , formant un second groupe de L cartes de caractéristiques supplémentaires, sont générées et complètent le premier groupe. Elles ne sont pas décodées mais générées par le décodeur de manière identique à la génération du codeur. Elles comportent typiquement des données qui peuvent assister le réseau MLP' dans la tâche de reconstruction du signal. La liste non limitative de cartes de caractéristiques supplémentaires possibles décrite à l'appui de la [Fig.5] pour le codeur s'applique ici.
- [0146] Lors d'une étape E33, des valeurs sont extraites par le module XTR' des cartes FMD_i et éventuellement FME'_1 . Cette extraction est réalisée en fonction des coordonnées (x_n, y_n) du signal d'entrée de l'échantillon P_n du signal à reconstruire. Cette étape est identique à l'étape E24 qui a été décrite pour le codeur à l'appui de la [Fig.4] et les modes de réalisation décrits s'appliquent.
- [0147] Les échantillons à décoder sont par exemple traités par ordre séquentiel, de $n=1$ à $n=N$.
- [0148] Lors d'une étape E34, un vecteur Zd_n est construit par le module TT' à partir des valeurs extraites des cartes FMD_i et éventuellement FME'_1 pour un échantillon P_n à décoder de coordonnées (x_n, y_n) . Le traitement peut comporter une quantification inverse des valeurs extraites ou du vecteur Zd_n constitué, si nécessaire. Le traitement peut comporter d'autres opérations, comme un filtrage, une mise à l'échelle, etc. Cette étape est identique à l'étape E25 qui a été décrite pour le codeur à l'appui de la [Fig.4] et les modes de réalisation décrits s'appliquent. Cependant la quantification inverse n'est pas obligatoire. Dans ce cas particulier, l'étape E34 et le module TT' sont donc optionnels.
- [0149] Lors d'une étape E35, les paramètres Wd_k du réseau de neurones MLP' sont générés par décodage des valeurs Wc_k du flux B2. On peut utiliser à cette fin toute technique connue de décodage correspondant à la technique de codage qui a été utilisée par le codeur. Le réseau de neurones MLP' est semblable au réseau MLP, c'est-à-dire qu'il est de même structure et comporte les mêmes paramètres, au codage près, qui peut être réalisé avec ou sans pertes.
- [0150] Selon un mode de réalisation, le flux B2 est décodé avant le flux B1, afin de pouvoir disposer du réseau de neurones avant de commencer à décoder les échantillons.
- [0151] Lors d'une étape E36, le vecteur Zd_n est traité par le réseau de neurones MLP' pour générer en sortie un second vecteur représentatif de l'échantillon Pd_n à décoder, selon un mode de réalisation un triplet qui est injecté dans l'image décodée $I(Pd_n)$ aux positions (x_n, y_n) des composantes couleur (Rd, Gd, Bd). Cette étape est identique à l'étape E26 qui a été décrite pour le codeur à l'appui de la [Fig.4].
- [0152] Lorsque tous les échantillons du signal ont été traités, le signal décodé correspondant

par exemple à l'image $I(Pd_n)$ est disponible.

- [0153] La [Fig.8] représente de manière illustrée un procédé de décodage utilisé dans un mode de réalisation de l'invention pour décoder une image $I(Pd_n)$.
- [0154] Dans ce mode de réalisation, les cartes FMD_i sont au nombre de 4. Dans un mode préféré elles sont au nombre de 7.
- [0155] Dans ce mode de réalisation, les cartes FMD_i sont de même résolution que l'image à décoder, et comportent donc $W \times H$ variables, où W représente la largeur de l'image en pixels, et H sa hauteur.
- [0156] Dans ce mode de réalisation, le vecteur Zd_n est un 4-uplet $(z_1 \dots z_4)$ constitué des valeurs des cartes FMD_i situées aux coordonnées (x_n, y_n) du pixel courant Pd_n . Le vecteur Zd_n est optionnellement déquantifié puis traité par le réseau de neurones MLP' pour générer en sortie un second vecteur, selon l'exemple un triplet (R, G, B) représentatif de l'échantillon Pd_n à décoder. Le triplet (R, G, B) est réinjecté dans l'image décodée $I(Pd_n)$ aux coordonnées (x_n, y_n) dans les composantes couleur (Rd, Gd, Bd) .
- [0157] Selon une variante présentée en pointillés, les cartes sont au nombre de 5 : une carte supplémentaire FME'_0 a été introduite. Dans ce mode de réalisation, le vecteur Zd_n est un 5-uplet.

Revendications

- [Revendication 1] Procédé de codage d'un signal ($I(P_n)$) comprenant une pluralité d'échantillons (P_n) à coder comportant les étapes suivantes :
- une étape de construction, comportant les sous-étapes de :
 - construction (E21, E22) d'un premier groupe de cartes de caractéristiques (FM_i);
 - pour au moins un échantillon, dit échantillon courant (P_n), du signal à coder, associé à une position (x_n, y_n) dans le signal à coder :
 - construction (E24, E25) d'un vecteur caractéristique (Z_n) à partir desdites cartes de caractéristiques (FM_i) dudit premier groupe, en fonction de ladite position (x_n, y_n) dudit échantillon courant (P_n) ;
 - traitement (E25) dudit vecteur caractéristique (Z_n) par un réseau de neurones artificiels (MLP) défini par un ensemble de paramètres (W_k), pour fournir un vecteur représentatif d'une valeur décodée (P'_n) de l'échantillon courant ;
 - mise à jour (E22, E26) d'au moins une valeur d'une desdites cartes de caractéristiques dudit premier groupe et/ou d'au moins un paramètre dudit réseau, en fonction d'une mesure de performance de codage,
 - une étape de codage (E23, E27, E29) dudit premier groupe de cartes de caractéristiques (FM_i) et dudit ensemble de paramètres (W_k).
- [Revendication 2] Procédé de codage d'un signal selon la revendication 1, caractérisé en ce qu'il comporte une étape de construction (E21) d'un second groupe de cartes de caractéristiques (FME_1), et en ce que le vecteur caractéristique est en outre construit à partir desdites cartes de caractéristiques du second groupe.
- [Revendication 3] Procédé de codage d'un signal selon l'une des revendications précédentes, caractérisé en ce qu'une au moins desdites cartes de caractéristiques (FM_i, FME_1) est de résolution inférieure à celle du signal à coder.
- [Revendication 4] Procédé de codage d'un signal selon l'une des revendications précédentes, caractérisé en ce qu'une au moins desdites cartes de caractéristiques (FM_i, FME_1) est de résolution identique à celle du signal à coder.
- [Revendication 5] Procédé de codage d'un signal selon l'une des revendications précédentes, caractérisé en ce que la construction dudit vecteur (Z_n) caractéristique comporte une sous-étape d'extraction d'une valeur de ladite au moins une carte de caractéristiques (FM_i, FME_1) à une position

identique (x_n, y_n) à celle de l'échantillon courant (P_n) dans le signal d'entrée.

- [Revendication 6] Procédé de codage d'un signal selon l'une des revendications précédentes, caractérisé en ce que la construction dudit vecteur (Z_n) caractéristique comporte les sous-étapes suivantes :
- extraction d'une pluralité de valeurs desdites cartes de caractéristiques (FM_i, FME_i) en fonction de ladite position (x_n, y_n) dudit échantillon courant (P_n) ;
 - traitement desdites valeurs extraites pour obtenir le vecteur caractéristique.
- [Revendication 7] Procédé de codage d'un signal selon l'une des revendications précédentes, caractérisé en ce que le codage dudit premier groupe de cartes de caractéristiques comporte une sous-étape de codage entropique.
- [Revendication 8] Procédé de décodage d'un signal comprenant une pluralité d'échantillons (Pd_n) à décoder comportant les étapes suivantes :
- décodage (E31) d'un premier groupe de cartes de caractéristiques (FMd_i) ;
 - décodage (E35) d'un ensemble de paramètres (Wd_k) représentatifs d'un réseau de neurones (MLP') ;
 - pour au moins un échantillon, dit échantillon courant (Pd_n) , du signal à décoder, associé à une position (x_n, y_n) dans le signal à décoder :
 - construction (E33, E34) d'un vecteur caractéristique (Zd_n) à partir des cartes de caractéristiques dudit premier groupe (FMd_i) , en fonction de ladite position (x_n, y_n) dudit échantillon courant, et :
 - traitement (E36) dudit vecteur caractéristique (Zd_n) par un réseau de neurones artificiels (MLP') défini par les paramètres décodés (Wd_k) pour fournir un vecteur représentatif d'une valeur décodée de l'échantillon courant (Pd_n) .
- [Revendication 9] Procédé de décodage selon la revendication 8, caractérisé en ce en ce qu'il comporte une étape de construction (E32) d'un second groupe de cartes de caractéristiques (FME'_i) et en ce que le vecteur caractéristique est en outre construit à partir desdites cartes de caractéristiques du second groupe.
- [Revendication 10] Procédé de décodage selon la revendication 8 ou 9, caractérisé en ce qu'une au moins desdites cartes de caractéristiques (FMd_i, FME'_i) est de résolution inférieure à celle du signal à décoder.
- [Revendication 11] Procédé de décodage d'un signal selon l'une des revendications 8 à 10, caractérisé en ce qu'une au moins desdites cartes de caractéristiques

- (FM_{d_i} , FME'_{i_1}) est de résolution identique à celle du signal à décoder.
- [Revendication 12] Procédé de décodage d'un signal selon l'une des revendications 8 à 11, caractérisé en ce que la construction dudit vecteur (Z_{d_n}) caractéristique comporte une sous-étape d'extraction d'une valeur de ladite au moins une carte de caractéristiques (FM_{d_i} , FME'_{i_1}) à une position identique (x_n , y_n) à celle de l'échantillon courant (P_{d_n}) dans le signal à décoder.
- [Revendication 13] Procédé de décodage d'un signal selon l'une des revendications 8 à 12, caractérisé en ce que la construction dudit vecteur caractéristique (Z_{d_n}) comporte les sous-étapes suivantes :
- extraction d'une pluralité de valeurs des cartes de caractéristiques en fonction de ladite position (x_n , y_n) dudit échantillon courant (P_{d_n}) ;
 - traitement desdites valeurs extraites pour obtenir le vecteur caractéristique.
- [Revendication 14] Procédé de décodage d'un signal selon l'une des revendications 8 à 13, caractérisé en ce que le décodage (E31) dudit premier groupe de cartes de caractéristiques comporte une sous-étape de décodage entropique.
- [Revendication 15] Dispositif de codage d'un signal ($I(P_n)$) comprenant une pluralité d'échantillons (P_n) à coder caractérisé en ce que ledit dispositif est configuré pour mettre en œuvre :
- construction (GEN, MAJ) d'un premier groupe de cartes de caractéristiques (FM_i) ;
 - pour au moins un échantillon, dit échantillon courant (P_n), du signal à coder, associé à une position (x_n , y_n) dans le signal à coder :
 - construction (XTR, TT) d'un vecteur caractéristique (Z_n) à partir desdites cartes de caractéristiques (FM_i) dudit premier groupe, en fonction de ladite position (x_n , y_n) dudit échantillon courant (P_n), et ;
 - traitement (MLP) dudit vecteur caractéristique (Z_n) par un réseau de neurones artificiels (MLP) défini par un ensemble de paramètres (W_k), pour fournir un vecteur représentatif d'une valeur décodée (P'_n) de l'échantillon courant ;
 - mise à jour (MAJ, NNC) d'au moins une valeur d'une desdites cartes de caractéristiques dudit premier groupe et/ou d'au moins un paramètre dudit réseau, en fonction d'une mesure de performance de codage,
 - codage (FMC, NNC) dudit premier groupe de cartes de caractéristiques (FM_i) et dudit ensemble de paramètres (W_k).
- [Revendication 16] Dispositif de décodage d'un signal comprenant une pluralité d'échantillons (P_{d_n}) à décoder caractérisé en ce que ledit dispositif est configuré pour mettre en œuvre :

- décodage (FMD) d'un premier groupe de cartes de caractéristiques (FMd_i) ;
- décodage (NND) d'un ensemble de paramètres (Wd_k) représentatifs d'un réseau de neurones (MLP') ;
- pour au moins un échantillon, dit échantillon courant (Pd_n), du signal à décoder, associé à une position (x_n, y_n) dans le signal à décoder :
- construction (XTR', TT') d'un vecteur caractéristique (Zd_n) à partir des cartes de caractéristiques dudit premier groupe (FMd_i), en fonction de ladite position (x_n, y_n) dudit échantillon courant , et :
- traitement (MLP') dudit vecteur caractéristique (Zd_n) par un réseau de neurones artificiels (MLP') défini par les paramètres décodés (Wd_k) pour fournir un vecteur représentatif d'une valeur décodée de l'échantillon courant (Pd_n).

[Revendication 17]

Programme d'ordinateur comportant des instructions pour l'exécution des étapes d'un procédé de codage ou de décodage selon la revendication 1 ou 8 lorsque ledit programme est exécuté par un ordinateur.

[Fig. 4]

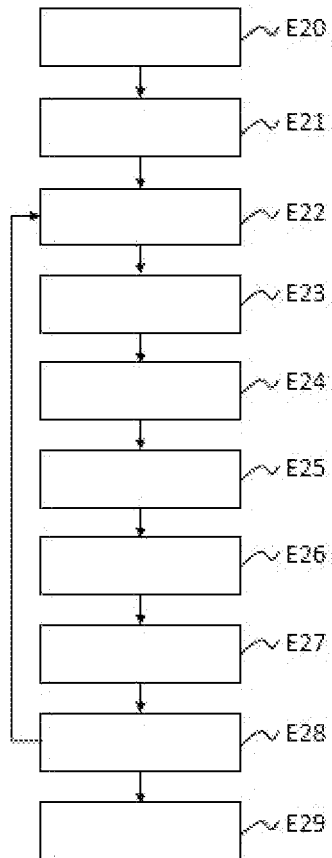


Figure 4

[Fig. 5]

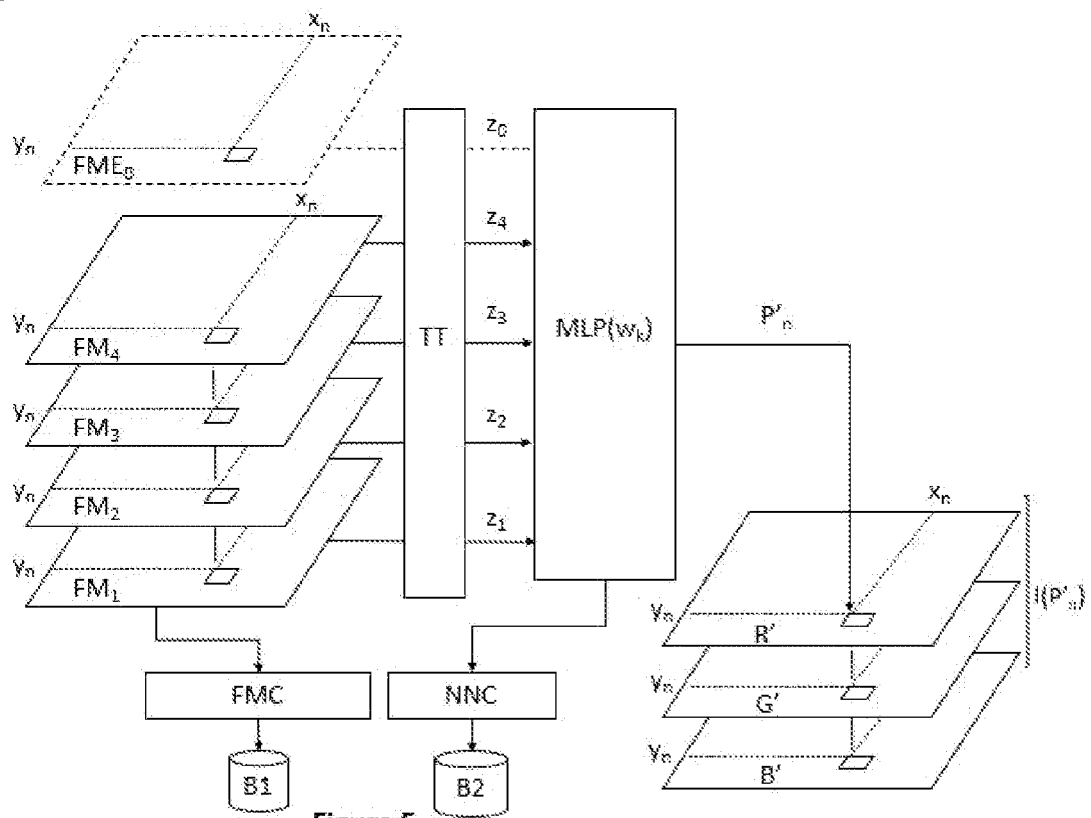
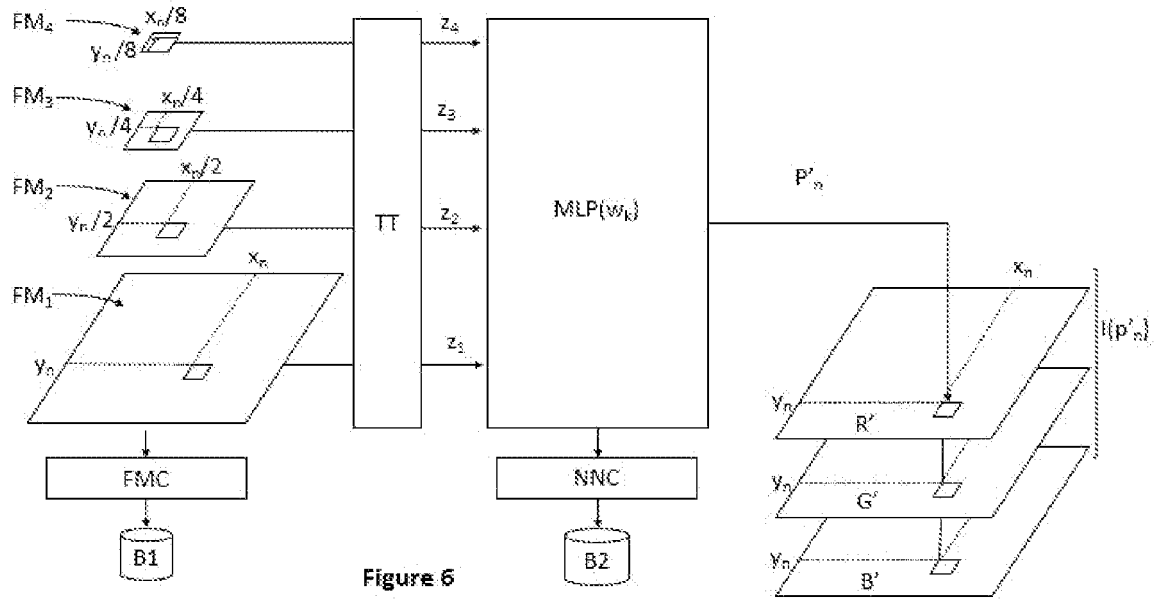
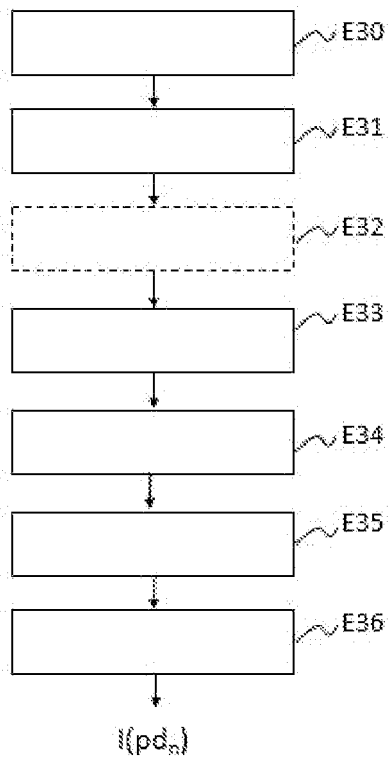


Figure 5

[Fig. 6]



[Fig. 7]



[Fig. 8]

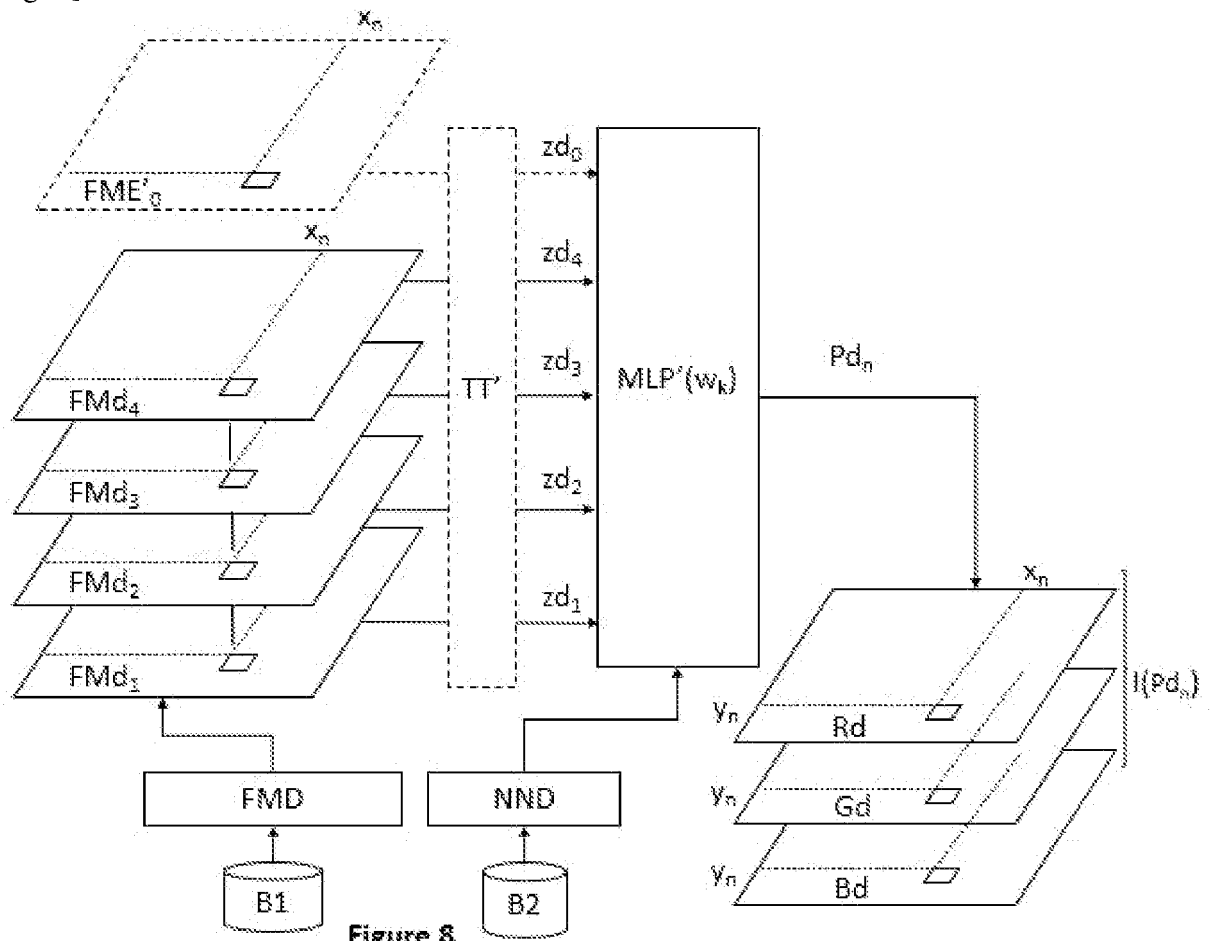


Figure 8

**RAPPORT DE RECHERCHE
PRÉLIMINAIRE**

N° d'enregistrement
national

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

FA 916194
FR 2213086

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
X	<p>SUBIN KIM ET AL: "Scalable Neural Video Representations with Learnable Positional Features", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 13 octobre 2022 (2022-10-13), XP091342999, * sections 1 et 3 *</p> <p>-----</p>	1-17	H04N19/50 G06T9/00 G06N3/08
X	<p>WENTAO SHANGGUAN ET AL: "Learning Cross-Video Neural Representations for High-Quality Frame Interpolation", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 28 février 2022 (2022-02-28), XP091166479, * section 3 *</p> <p>-----</p>	1, 8, 15-17	
A	<p>Yuan Zhongzheng: "Block-based Image Coding with Autoencoder and Border Information", / 20 mai 2020 (2020-05-20), pages 1-34, XP055938495, Extrait de l'Internet: URL:https://s18798.pcdn.co/videolab/wp-content/uploads/sites/10258/2020/10/Zhongzhen g_Yuan_Thesis_Final_Submitted.pdf [extrait le 2022-07-05] * Chapitre 4 *</p> <p>-----</p> <p style="text-align: center;">-/--</p>	1-17	<p>DOMAINES TECHNIQUES RECHERCHÉS (IPC)</p> <p>G06N H04N</p>
Date d'achèvement de la recherche		Examineur	
8 août 2023		Ferré, Pierre	
CATÉGORIE DES DOCUMENTS CITÉS		<p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant</p>	
<p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire</p>			

**RAPPORT DE RECHERCHE
PRÉLIMINAIRE**

N° d'enregistrement
national

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

FA 916194
FR 2213086

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
A	<p>YUAN ZHONGZHENG ET AL: "Block-based Learned Image Coding with Convolutional Autoencoder and Intra-Prediction Aided Entropy Coding", 2021 PICTURE CODING SYMPOSIUM (PCS), 1 juin 2021 (2021-06-01), pages 1-5, XP055837685, DOI: 10.1109/PCS50896.2021.9477503 ISBN: 978-1-6654-2545-2 * section III *; figures 1,2,3</p> <p style="text-align: center;">-----</p>	1-17	
T	<p>TH\ 'EO LADUNE ET AL: "COOL-CHIC: Coordinate-based Low Complexity Hierarchical Image Codec", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 11 décembre 2022 (2022-12-11), XP091392072, * le document en entier *</p> <p style="text-align: center;">-----</p>		
Date d'achèvement de la recherche		Examineur	
8 août 2023		Ferré, Pierre	
CATÉGORIE DES DOCUMENTS CITÉS		<p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant</p>	
<p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire</p>			