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

(19) Organisation Mondiale de la Propriété Intellectuelle
Bureau international

(43) Date de la publication internationale
12 juillet 2012 (12.07.2012)

WIPO | PCT

(10) Numéro de publication internationale
WO 2012/093216 A1

(51) Classification internationale des brevets :
G06F 21/24 (2006.01) *G06F 17/30* (2006.01)

(21) Numéro de la demande internationale :
PCT/FR2011/053034

(22) Date de dépôt international :
16 décembre 2011 (16.12.2011)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
1150135 7 janvier 2011 (07.01.2011) FR

(71) Déposant *(pour tous les États désignés sauf US)* : **THOMSON LICENSING** [FR/FR]; 1-5 rue Jeanne d'Arc, F-92130 Issy Les Moulineaux (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants *(pour US seulement)* : **MONTALVO, Luis** [EC/FR]; TECHNICOLOR R&D FRANCE, 1 avenue de Belle Fontaine, F-35576 Cesson Sevigne (FR). **LE SCOUARNEC, Nicolas** [FR/FR]; TECHNICOLOR R&D FRANCE, 1 avenue de Belle Fontaine, F-35576 Cesson Sevigne (FR). **DEFRANCE, Serge** [FR/FR]; TECHNICOLOR R&D FRANCE, 1 Avenue de Belle Fontaine, F-35576 Cesson-Sévigné (FR). **LEFEBVRE, Frédéric** [FR/FR]; TECHNICOLOR R&D FRANCE, 1 avenue de Belle Fontaine, F-35576 Cesson Sevigne (FR). **PEREZ, Patrick** [FR/FR]; TECHNICOLOR R&D FRANCE, 1 avenue de Belle Fontaine, F-35576 Cesson Sevigne (FR).

(74) Mandataire : **HUCHET, Anne**; TECHNICOLOR, 1-5 rue Jeanne d'Arc, F-92130 Issy-Les-Moulineaux (FR).

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

(84) États désignés *(sauf indication contraire, pour tout titre de protection régionale disponible)* : ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ,

[Suite sur la page suivante]

(54) Title : DEVICE AND METHOD FOR ONLINE STORAGE, TRANSMISSION DEVICE AND METHOD, AND RECEIVING DEVICE AND METHOD

(54) Titre : DISPOSITIF ET PROCÈDE DE STOCKAGE EN LIGNE, DISPOSITIF ET PROCÈDE D'ÉMISSION, DISPOSITIF ET PROCÈDE DE RÉCEPTION

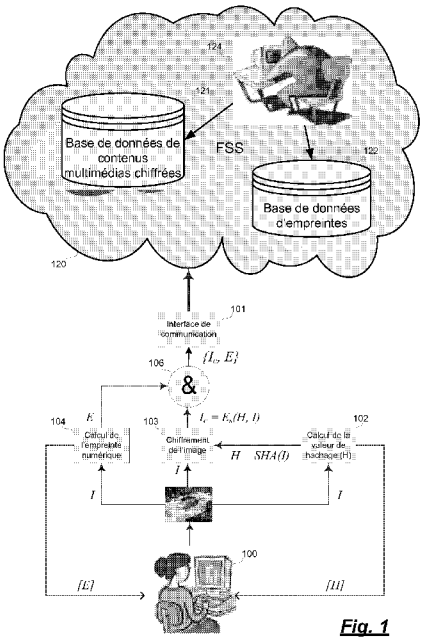

Fig. 1

101 Communication interface
102 Calculate the hashing value (H)
103 Encrypt the image
104 Calculate the digital mark
121 Encrypted-multimedia-content database
122 Mark database

(57) Abstract : The invention relates to a device and method for online storage, to a device and method for searching for similar content, to a transmission device and method, and to a receiving device and method. Encrypted data is recorded at an online service provider. With the encrypted data, encrypted hashing data is recorded with a public key, and the content to be recorded with the encrypted hashing is encrypted, thus making it possible to advantageously prevent data duplication at the online service provider while maintaining the privacy of the users of the service. In order to search for similar content having multimedia reference data, marks are also recorded at the service provider. In order to limit the number of false positives returned, the mark can also contain a search mark and/or an encrypted selection mark.

(57) Abrégé : L'invention concerne un dispositif et une méthode de stockage en ligne, dispositif et méthode de recherche de contenu similaire, un dispositif et

[Suite sur la page suivante]

WO 2012/093216 A1

UG, ZM, ZW), eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), européen (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Déclarations en vertu de la règle 4.17 :

— *relative au droit du déposant de demander et d'obtenir un brevet (règle 4.17.ii))*

— *relative à la qualité d'inventeur (règle 4.17.iv))*

Publiée :

— *avec rapport de recherche internationale (Art. 21(3))*

— *avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont reçues (règle 48.2.h))*

une méthode d'émission et un dispositif et une méthode de réception. On enregistre des données chiffrées chez un fournisseur de services en ligne. Avec les données chiffrées, on enregistre des données de hachage chiffrées avec une clé publique et l'on chiffre le contenu à enregistrer avec le hachage. Cela permet avantageusement d'éviter la duplication de données chez le fournisseur de services en ligne tout en préservant la vie privée des utilisateurs du service. Afin de rechercher des contenus similaires à des données multimédias de référence, on enregistre également des empreintes chez le fournisseur de services. Afin de limiter le nombre de faux positifs retournés, l'empreinte peut contenir outre une empreinte de recherche, une empreinte de sélection chiffrée.

DISPOSITIF ET PROCEDE DE STOCKAGE EN LIGNE, DISPOSITIF ET PROCEDE D'EMISSION, DISPOSITIF ET PROCEDE DE RECEPTION

1. Domaine de l'invention.

L'invention concerne un dispositif et une méthode de stockage partagé.

L'invention concerne encore un dispositif et une méthode d'émission de documents chiffrés permettant la déduplication et la recherche similaire.

L'invention concerne enfin un dispositif et une méthode de réception permettant de déchiffrer les documents partagés et de procéder à une recherche similaire.

Dans les systèmes de stockage en ligne de photos, les intérêts de performance du fournisseur du service de stockage, et de protection de la vie privée des utilisateurs du service, peuvent entrer en conflit. En effet, si les utilisateurs confient leurs collections de photos en clair au fournisseur du service de stockage, celui-ci peut identifier les images identiques et les stocker dans l'espace correspondant à une seule image, et ceci quels que soient les propriétaires de photos identiques ; en revanche, la confidentialité des utilisateurs est compromise. Par contre, si les utilisateurs du service de stockage chiffrent leurs images avant de les envoyer au fournisseur du service, celui-ci ne peut plus identifier les images identiques si celles-ci ont été chiffrées avec des clefs différentes.

Dans le but d'optimiser l'espace de stockage et les délais de téléchargement de fichiers, les Fournisseurs du Service de Stockage (FSS) peuvent appliquer des techniques de déduplication de fichiers aux données des utilisateurs. La déduplication signifie le fait de ne pas dupliquer un même contenu.

Quelques FSS appliquent ces techniques non seulement aux données appartenant au même compte utilisateur (déduplication intra compte) mais

également aux données appartenant à différents comptes utilisateurs (déduplication inter compte). Cette façon de gérer les données en clair des utilisateurs peut être perçue comme une atteinte à leur vie privée par les utilisateurs.

Une contre-mesure à la manipulation des données en clair est le chiffrement des données. Les utilisateurs pourraient chiffrer les images avant de les envoyer au fournisseur du service de stockage en ligne. Malheureusement, un même contenu en clair chiffré avec deux clefs différentes, produit deux contenus chiffrés très différents. La vie privée des utilisateurs est préservée mais le FSS ne peut pas détecter que les deux messages chiffrés correspondent à un même message en clair et il ne peut plus optimiser l'espace de stockage.

2. Résumé de l'invention.

L'invention propose de pallier à au moins un des inconvénients de l'art antérieur.

L'invention concerne un système de stockage en ligne qui réconcilie ces deux intérêts en apparence conflictuels. Le fournisseur du service de stockage en ligne a la capacité à identifier non seulement les données multimédias identiques mais aussi les données multimédias similaires, même si elles sont chiffrées avec des clefs différentes, sans compromettre la vie privée des utilisateurs.

À cet effet, l'invention concerne un dispositif de stockage en ligne apte à stocker des données multimédias. Le dispositif comprend des moyens d'enregistrer les données multimédias sous forme chiffrée et pour chacune des données multimédias chiffrées, une empreinte associée. Les données multimédias chiffrées sont chiffrées en utilisant une valeur de hachage. La valeur de hachage est obtenue par le hachage des données multimédias non chiffrées. L'empreinte comprend un vecteur de recherche non chiffré.

Avantageusement, l'empreinte comprend un vecteur de vérification chiffré.

Préférentiellement, le vecteur de vérification chiffré est obtenu par le chiffrage du vecteur de vérification avec la valeur de hachage.

Avantageusement, le dispositif est apte à stocker au moins un identifiant d'utilisateur et un hachage chiffré par utilisateur, le hachage chiffré étant chiffré en utilisant une clé publique de chiffrement de l'utilisateur.

Selon une variante, le dispositif est apte à stocker au moins un identifiant d'utilisateur et pour chaque utilisateur un hachage chiffré, le hachage chiffré et le vecteur de vérification chiffré sont chiffrés en utilisant une clé publique de chiffrement de l'utilisateur.

Selon un mode de réalisation particulier, le dispositif comprend des moyens de comparer, lors de chaque enregistrement de données multimédias, les données multimédias chiffrées enregistrées avec les données multimédias chiffrées à enregistrer de manière à ne pas dupliquer les données multimédias enregistrées.

Avantageusement, le dispositif comprend des moyens :

- de recevoir une première requête d'un utilisateur émetteur. Cette requête pour un document multimédia cible recherché, comprend au moins l'empreinte associée au document chiffré,
- de comparer le vecteur de recherche reçu à au moins un vecteur de recherche enregistré en mesurant une distance entre le vecteur de recherche reçu et l'au moins un vecteur de recherche enregistré,
- de transmettre à l'émetteur de la première requête, au moins un vecteur de recherche dont la distance avec le vecteur de recherche reçu est inférieure à un seuil prédéterminé, dit vecteur sélectionné,

Avantageusement, le dispositif comprend des moyens :

- de recevoir une seconde requête de la part de l'émetteur de la première. La requête, pour au moins un document multimédia recherché, comprend au moins un vecteur de recherche sélectionné parmi l'au moins un vecteur de recherche transmis.
- de transmettre à l'émetteur de la seconde requête, au moins un vecteur de vérification chiffré correspondant à au moins un vecteur de recherche sélectionné.

Selon un mode de réalisation particulier, le dispositif comprend des moyens :

- de recevoir une première requête d'un utilisateur émetteur. La requête, pour un document multimédia cible recherché, comprend au moins l'empreinte associée au document chiffré,
- de comparer le vecteur de recherche reçu à au moins un vecteur de recherche enregistré en mesurant une distance entre le vecteur de recherche reçu et l'au moins un vecteur de recherche enregistré,
- de transmettre à l'émetteur de la première requête, au moins un vecteur de vérification chiffré correspondant au vecteur de recherche sélectionné.

Avantageusement, suite à la transmission d'au moins un vecteur de vérification, le dispositif comprend des moyens :

- de recevoir une requête de l'émetteur pour au moins un document multimédia recherché. La requête comprend au moins un vecteur de vérification sélectionné parmi au moins un vecteur de vérification transmis.
- de transmettre à l'émetteur de la requête, au moins une paire correspondant à au moins un vecteur de vérification reçu, comprenant au moins une donnée chiffrée et au moins un hachage chiffré correspondant.

Avantageusement, suite à la transmission d'au moins un vecteur de vérification, le dispositif comprend des moyens :

- de recevoir une requête émise suite au déchiffrage des vecteurs de vérification par l'émetteur de la première requête et à l'élimination de faux positifs. La seconde requête comprend un identificateur des données sélectionnées suite à l'élimination des faux positifs ;
- des moyens de transmettre à l'émetteur de la requête, les paires, comprenant les données chiffrées et le hachage chiffré correspondant associé ;

L'invention concerne également un procédé de stockage en ligne apte à stocker des données multimédias. Ce procédé comprend les étapes de :

- réception de données comprenant un contenu multimédia sous forme chiffrée selon un chiffrement convergent, une empreinte associée.
- comparaison du contenu multimédia chiffré avec ceux préalablement stockés.
- enregistrement des données, si le contenu n'est pas déjà stocké.

Le procédé comprend également les étapes de :

- réception de données. Les données comprennent un identifiant d'utilisateur destinataire, une valeur de hachage chiffrée à l'attention de l'utilisateur, une empreinte associée. L'empreinte comprend un vecteur de recherche non chiffré ;
- recherche du contenu multimédia chiffré, correspondant au vecteur de recherche non chiffré ;
- transmission du résultat de la recherche, si le contenu multimédia chiffré a été trouvé. Les données transmises comprennent le contenu multimédia chiffré, la valeur de hachage chiffrée ;

Au cours de ce procédé lors de l'étape de transmission, les données transmises comprennent également l'empreinte associée au contenu multimédia chiffré.

Procédé de stockage en ligne de données multimédias comprenant les étapes de :

- hachage des données multimédias à enregistrer afin d'obtenir une valeur de hachage,
- chiffrement des données multimédias à enregistrer par la valeur de hachage,
- calcul d'une empreinte associée aux données multimédias à enregistrer,
- enregistrement des données multimédia chiffrées et de l'empreinte associée.

L'invention propose également selon un second aspect, un procédé de transmission de données multimédias. Ce procédé est utilisé par l'utilisateur souhaitant stocker ses données sur le FSS. Ce procédé comprend les étapes de :

- chiffrement des données multimédias à l'aide d'une méthode de chiffrement convergent ;
- calcul d'une empreinte correspondant aux données multimédias, à partir des données multimédias non chiffrées. L'empreinte comprend un vecteur de recherche non chiffré ;
- transmission d'un ensemble de données comprenant au moins les données multimédias chiffrées et l'empreinte associée.

Avantageusement, lors de l'étape de transmission, l'ensemble de données envoyé comprend au moins un couple utilisateur et une clé de hachage chiffrée associée à l'utilisateur. L'utilisateur est destinataire des données. La clé de hachage chiffrée est obtenue à partir de la valeur de hachage des données multimédias non chiffrées, chiffrée avec la clé publique de l'utilisateur.

Avantageusement, l'empreinte obtenue lors de l'étape de calcul, comprend un vecteur de recherche et un vecteur de vérification chiffré.

Préférentiellement, le vecteur de vérification chiffré est obtenu par le chiffrage de ce vecteur de vérification avec la valeur de hachage.

Avantageusement, le vecteur de vérification chiffré est obtenu par le chiffrage de ce vecteur de vérification avec la clé publique de chiffrement de l'utilisateur.

L'invention concerne également un dispositif de transmission de données multimédias. Ce dispositif comprend :

- des moyens de chiffrement des données multimédias à l'aide d'une méthode de chiffrement convergent ;
- des moyens de calcul d'une empreinte des données multimédias à partir des données multimédias non chiffrées ;
- des moyens de transmission d'un ensemble de données comprenant au moins les données multimédias chiffrées et l'empreinte associée.

Avantageusement, l'ensemble de données envoyé par ce dispositif comprend au moins un couple utilisateur et une clé de hachage chiffrée associée à cet utilisateur. Le hachage chiffré est destiné à permettre à l'utilisateur de déchiffrer les données multimédias. Cette clé de hachage chiffrée est obtenue à partir de la valeur de hachage des données multimédias non chiffrées, chiffrée avec la clé publique de l'utilisateur.

Selon un troisième aspect, l'invention propose un procédé de réception de données multimédias. Ce procédé est destiné à un utilisateur destinataire des données multimédias et possédant une clé publique et une clé privée associée. Ce procédé comprend les étapes de :

- réception d'un ensemble de données comprenant au moins des données multimédias sous forme chiffrée par une méthode de chiffrement convergent et une valeur de hachage chiffrée associée aux données multimédias, destinée à l'utilisateur ;
- déchiffrage de la valeur de hachage chiffré avec la clé privée pour obtenir une valeur de hachage non chiffrée ;

- déchiffrage des données multimédias avec la valeur de hachage pour obtenir les données multimédias non chiffrées.

Avantageusement, l'ensemble de données reçu à l'étape de réception comprend également un vecteur de recherche.

Avantageusement, le procédé de réception comprend les étapes de :

- transmission d'une requête de recherche de données multimédias similaires, la requête comprenant un vecteur de recherche ;
- réception d'empreintes similaires à l'empreinte transmise. Les empreintes similaires sont composées de doublets comprenant un vecteur de recherche, un vecteur de vérification chiffrée destinée à l'utilisateur ;
- déchiffrage des vecteurs de vérification avec la clé privée pour obtenir les vecteurs de vérifications non chiffrées ;

Avantageusement, le procédé de réception comprend les étapes de :

- transmission d'une requête de données multimédias similaires sélectionnées.
- réception de données multimédias sous forme chiffrée et une valeur de hachage chiffrée associée aux données multimédias. Les données multimédias sont chiffrées par une méthode de chiffrement convergent . La valeur de hachage chiffrée est destinée à l'utilisateur ;
- déchiffrage de la valeur de hachage chiffré avec la clé privée pour obtenir une valeur de hachage non chiffrée ;
- déchiffrage des données multimédias avec la valeur de hachage pour obtenir les données multimédias non chiffrées .

L'invention concerne également un dispositif de réception de données multimédias, destinées à un utilisateur possédant une clé publique et une clé privée associée, comprend des moyens :

- de réception d'un ensemble de données comprenant au moins des données multimédias sous forme chiffrée par une méthode convergente et une valeur de hachage chiffrée associée aux données multimédias, destinée à l'utilisateur ;
- de déchiffrage de la valeur de hachage chiffré avec la clé privée pour obtenir une valeur de hachage non chiffrée ;
- de déchiffrage des données multimédias avec la valeur de hachage pour obtenir les données multimédias non chiffrées.

Avantageusement, l'ensemble de données comprend également un vecteur de recherche.

Avantageusement, le dispositif de réception comprend des moyens :

- de transmission, d'une requête de recherche de données multimédias similaires, la requête comprenant le vecteur de recherche ;
- de réception de doublets comprenant un vecteur de recherche, un vecteur de vérification chiffré destinée à l'utilisateur.
- de déchiffrage des vecteurs de vérification avec la clé pour obtenir les vecteurs de vérification non chiffrés.

Avantageusement, le dispositif de réception comprend des moyens :

- de transmission d'une requête de données multimédias similaires sélectionnées.
- de réception de données multimédias sous forme chiffrée par une méthode convergente et d'une valeur de hachage chiffrée associée aux données multimédias, destinée à l'utilisateur ;
- de déchiffrage de la valeur de hachage chiffré avec la clé privée pour obtenir une valeur de hachage non chiffrée ;
- de déchiffrage des données multimédias avec la valeur de hachage pour obtenir les données multimédias non chiffrées.

3. Liste des figures.

L'invention sera mieux comprise et illustrée au moyen d'exemples de modes de réalisation et de mise en œuvre avantageux, nullement limitatifs, en référence aux figures annexées sur lesquelles :

- la **figure 1** représente un dispositif de stockage selon un mode de réalisation préféré de l'invention,
- la **figure 2** représente un système mettant en œuvre un mode de réalisation préféré relatif à la consultation de données multimédias,
- la **figure 3** représente un système mettant en œuvre un second mode de réalisation relatif à la consultation de données multimédias,
- la **figure 4** représente un organigramme de fonctionnement d'un mode de réalisation préféré d'un aspect de l'invention relatif au chiffrement des données,
- la **figure 5** représente un organigramme de fonctionnement d'un mode de réalisation préféré de la déduplication,
- la **figure 6** représente un organigramme de fonctionnement de l'invention selon un premier mode de réalisation lié au stockage,
- la **figure 7** représente un organigramme de fonctionnement de l'invention selon un second mode de réalisation lié au stockage.
- la **figure 8** représente un organigramme de fonctionnement du chiffrement de la clé de hachage à l'intention d'un utilisateur destinataire.
- la **figure 9** représente un organigramme de fonctionnement d'un mode de réalisation préféré de l'invention lié au déchiffrage,
- la **figure 10** représente un organigramme de fonctionnement de recherche de contenus multimédias similaires selon un mode de réalisation préféré.

4. Description détaillée de l'invention.

L'invention sera décrite en référence à un mode particulier de réalisation destiné au stockage de photos ou d'images. L'obtention d'une empreinte est spécifique au type du document. À ce titre, les techniques de calcul d'empreinte citées s'appliquent aux images et photos.

Dans la suite de la description les termes, données multimédias, documents et contenus seront indifféremment utilisés pour désigner la même chose. Ces termes désigneront d'une part des images ou des photos, mais également tous les contenus multimédias avec lesquels se pose ce problème comme notamment des documents textes, audio et vidéo.

De même, les termes de chiffrage et codage de données sont indifféremment utilisés pour désigner le chiffrement de données dans le but de les protéger à la consultation par quiconque. Il faut noter que pour l'ensemble de la description, les données ou contenus multimédias sont stockés systématiquement chiffrés par le FSS.

Le terme de hachage et celui de hachage cryptographique sont utilisés comme synonymes.

Enfin, le terme de document (multimédia) de référence est utilisé dans le contexte de la recherche de contenu similaire. Il désigne le contenu multimédia dont l'empreinte est utilisée pour effectuer des comparaisons avec les empreintes des contenus multimédias stockés par le FSS, afin d'identifier les contenus similaires disponibles chez le FSS.

Par convention, sur les figures 1 à 3, les données représentées entre crochets sont optionnelles. Elles peuvent être transmises, stockées, utilisées selon les variantes de réalisation.

La **figure 1** représente un système mettant en œuvre un mode de réalisation préféré de l'invention.

Un premier utilisateur souhaite transmettre des données I à un correspondant destinataire. Un fournisseur de stockage en ligne FSS (120) propose des services de mutualisation, d'archivage de contenus multimédias entre plusieurs utilisateurs.

Les utilisateurs peuvent utiliser le service pour archiver leurs collections de contenus multimédias mais ils peuvent également l'employer pour partager soit l'ensemble, soit une partie de leurs collections avec d'autres utilisateurs autorisés.

Afin de satisfaire le besoin de préserver la vie privée des utilisateurs et la capacité du FSS à détecter les copies strictement identiques des documents multimédias dans les collections des utilisateurs, le système de stockage en ligne a les caractéristiques suivantes :

- Le FSS a accès seulement aux données multimédias chiffrées I_C des utilisateurs et le FSS ne doit pas pouvoir les déchiffrer. Le FSS reçoit et stocke donc, uniquement des données chiffrées.

- Le FSS peut détecter que deux contenus multimédias chiffrés correspondent à deux documents multimédias strictement identiques.

- Seuls les utilisateurs autorisés peuvent déchiffrer l'ensemble ou une partie des contenus multimédias chiffrés qui se trouvent archivés dans le compte d'un utilisateur.

Le dispositif (100) du premier utilisateur comprend des moyens de chiffrage (103) qui chiffrent le contenu multimédia à enregistrer. Le chiffrage E_s utilisé par les moyens de chiffrage (103) est un chiffrage du type convergent, il est décrit à la *figure 4*. Le dispositif (100) comprend des moyens de calcul (102) d'une valeur de hachage H. Le chiffrage de type convergent E_s va permettre au FSS d'appliquer les méthodes de déduplication de fichiers même si les fichiers sont chiffrés avec des clés différentes. Le chiffrage peut également être mis en œuvre sur un autre dispositif qu'un ordinateur et par des moyens hardware plutôt que par un programme d'ordinateur.

L'invention concerne également la possibilité de demander une recherche de contenus multimédias similaires au contenu multimédia consulté. Dans ce cadre, pour définir le contenu multimédia consulté, le terme de contenu multimédia de référence sera utilisé dans la suite de la description.

Afin de satisfaire le besoin de recherche de contenus multimédias similaires, le contenu multimédia chiffré I_c transmis est accompagné d'une empreinte numérique E.

L'empreinte E permet une recherche efficace, elle permet l'élimination des faux positifs et préserve la confidentialité du contenu de référence.

Le dispositif (100) du premier utilisateur comprend des moyens de calcul (104) pour calculer l'empreinte E à partir du contenu multimédia non chiffré I, tel que décrit à la *figure 6* et la *figure 7*. Lors du calcul, l'empreinte E tout comme la valeur de hachage non chiffrée H peuvent être stockées localement pour une utilisation ultérieure.

La **figure 2** représente un aspect de l'invention relatif à la consultation de données.

Un premier utilisateur souhaite autoriser l'accès à des données I, déjà stockées chez un FSS (120) à un utilisateur destinataire U. Ledit FSS propose des fonctionnalités associées à ses services de stockage tel que présenté lors de la description de la *figure 1*.

Pour pouvoir transmettre ses données multimédias, le premier utilisateur a besoin que l'utilisateur destinataire lui communique sa clé de chiffrage publique K_p, par tout moyen de communication connu de l'homme du métier, par exemple, un courriel. À l'aide de celle-ci le moyen de chiffrement (105) chiffre à l'attention de l'utilisateur destinataire, la valeur de hachage desdites données et lui transmet par les moyens de l'interface de communication (101) via le FSS (220) la valeur de hachage chiffrée H_c, tel que décrit à la *figure 8*. Des moyens (102) recalculent la valeur de hachage non chiffrée H à partir des données multimédias encore stockées par le dispositif

(100) du premier utilisateur. Selon une variante, la valeur de hachage H a été stockée non chiffrée sur le dispositif (100), lors de l'étape décrite à la *figure 1* et est réutilisée. Pour permettre au FSS d'établir la relation entre ladite valeur de hachage chiffrée H_c, les données multimédias correspondantes et l'utilisateur destinataire, le dispositif du premier utilisateur transmet également un identifiant de l'utilisateur U et une empreinte E des données multimédias. Tout comme la valeur de hachage non chiffrée, l'empreinte est recalculée telle que décrit à la *figure 1*. Selon une variante, l'empreinte E a été stockée telle que décrit à la *figure 1* par le moyen de calcul (104) et elle est réutilisée par le dispositif tel que le décrit la figure 2.

Le moyen (206) prépare des données comprenant un triplet <E, U, Hc> correspondant à un identifiant U de l'utilisateur destinataire, la valeur de hachage H_c du contenu multimédia et l'empreinte E du contenu multimédia. Les méthodes d'identification d'un utilisateur sont multiples et connues de l'homme du métier. Cet identifiant pourra par exemple être une adresse mail de l'utilisateur destinataire des contenus multimédias. Enfin, les données sont transmises au FSS (220) par l'interface de communication (101).

Comme nous le verrons à la *figure 8*, le FSS stocke dans une base de données (123) le couple < U, H_c >, identifiant de l'utilisateur, valeur de hachage et il établit un lien entre ce couple et la donnée multimédia chiffrée correspondante présente dans la base de données des contenus multimédias chiffrés (121). Pour établir ce lien, le FSS (220) s'appuie sur la base de données des empreintes (122) et l'empreinte E reçue dans ledit triplet.

L'interface de communication (241) mise en œuvre par le dispositif (140) de l'utilisateur destinataire reçoit des données qui comprennent le contenu chiffré I_c, la valeur de hachage chiffrée H_c associée. Les données sont fournies à un moyen de déchiffrement (142), par un moyen d'extraction (244). À l'aide de la clé privée de l'utilisateur destinataire K_s, le moyen de déchiffrement (142) déchiffre la valeur de hachage H_c. pour obtenir H. Des moyens (143) de déchiffrement du contenu utilise la valeur de hachage H pour déchiffrer le contenu multimédia selon la description de la *figure 9*. Les

données reçues par le dispositif de l'utilisateur destinataire peuvent également inclure une empreinte E sur les données multimédias pour permettre une recherche de contenus similaires. La transmission par le FSS (220) de l'empreinte E est facultative. Elle n'est pas nécessaire pour déchiffrer le contenu multimédia. Le côté optionnel de la transmission est symbolisé par une représentation entre crochets sur la figure, entre le FSS (220), l'interface de communication (241) et le moyen d'extraction (244).

La **figure 3** représente une variante du second aspect de l'invention relatif à la consultation de données décrit à la *figure 2.*

Comme dans le scénario précédemment cité, le premier utilisateur souhaite transmettre des données multimédias I à l'utilisateur destinataire. Dans son ensemble, le scénario est similaire. Il se distingue cependant, par la méthode de transmission appliquée par l'interface de communication (301) de la valeur de hachage chiffrée H_c et de l'empreinte E. Et d'autre part par une action supplémentaire effectuée par l'interface de communication (341) mise en œuvre par ledit programme contenu sur l'ordinateur de l'utilisateur destinataire, sous la forme d'une requête émise auprès du FSS (120), pour obtenir le contenu multimédia.

En effet, l'interface de communication (301) du premier utilisateur, après avoir reçu la paire < H_c, E > constituée par les moyens (306), transmet la paire directement au dispositif (140) de l'utilisateur destinataire U. La valeur de hachage chiffrée H_c est obtenue selon la méthode (105) décrite à la *figure 2*. L'empreinte E, tout comme celle de la figure 2 comprend une empreinte de recherche V (ou vecteur de recherche). Elle peut également comprendre une empreinte de vérification S_c (ou vecteur de vérification), pour permettre à l'utilisateur destinataire U de demander une recherche de contenu similaire.

Après réception de la paire < H_c, E >, le dispositif de réception (140), à l'aide des moyens de l'interface de communications (341) adresse une requête au FSS pour obtenir le contenu multimédia chiffré I_c. La requête de l'utilisateur destinataire comporte l'empreinte E de recherche. Préférentiellement, cette

requête inclue la valeur de hachage chiffrée H_c associée à l'identifiant U de l'utilisateur destinataire, pour permettre un stockage par le FSS.

À réception de ladite requête, le FSS (120) utilise l'empreinte de recherche pour identifier le contenu multimédia chiffré I_c à fournir. Avantageusement, le FSS peut également transmettre l'empreinte associée E contenue dans la base de données d'empreinte (122), comprenant un vecteur de vérification, pour permettre une recherche de contenus similaires ultérieure.

Dans le cas où le FSS reçoit également la paire < U, H_c > valeur de hachage chiffrée, identifiant de l'utilisateur U, le FSS stocke la paire dans la base de données des valeurs de hachage. Il crée également le lien entre ladite paire et le contenu multimédia chiffré I_c, tel que décrit à la *figure 8*.

L'interface de communication (341) de l'utilisateur destinataire, reçoit alors le contenu multimédia attendu et procède selon la méthode déjà décrite à la *figure 2* pour déchiffrer et exploiter le contenu.

La **figure 4** donne un organigramme de chiffrage convergent encore appelé chiffrement convergent mise en œuvre par les moyens de chiffrement (102,103) dans le dispositif (100).

Le chiffrage convergent décrit ci-dessous, est bien connu de l'homme du métier et repose sur une méthode de chiffrement symétrique.

En référence à la *figure 4*, lors d'une étape C1, le premier utilisateur sélectionne le contenu I à transmettre au FSS. Lors d'une étape C2, une valeur de hachage cryptographique H est calculée par l'ordinateur du premier utilisateur, cette valeur de hachage cryptographique H peut-être du type SHA-256. Il est possible de choisir une autre longueur de hachage par exemple SHA-512, mais également toute autre méthode de hachage comme MD5. Cette valeur de hachage cryptographique H est ensuite utilisée en tant que clé de chiffrage, pour chiffrer avec un algorithme symétrique E_s, le contenu I que le premier utilisateur souhaite transmettre lors d'une étape C3, et obtenir le

contenu chiffré I_c. C'est le contenu chiffré obtenu qui sera transmis au FSS (120,220) dans les *figure 6* et *figure 7*.

Ainsi, le système proposé permet avantageusement de garder la confidentialité des données enregistrées par le FSS tout en permettant à celui-ci de ne pas dupliquer inutilement les données enregistrées.

Puisque le contenu du fichier est chiffré avec sa propre valeur de hachage cryptographique comme clef, le contenu chiffré est indépendant des clefs de l'utilisateur destinataire, utilisées. Il n'est dépendant que du contenu en clair. Par conséquent, le fournisseur du service de stockage (FSS), sans connaissance des clefs privées des utilisateurs, peut détecter que deux fichiers sont strictement identiques et les stocker dans l'espace correspondant à un seul fichier. Ainsi, le FSS minimise la place de stockage nécessaire pour stocker toutes les données à stocker.

Le fonctionnement de la déduplication est illustré dans l'organigramme détaillé en **figure 5**.

Lorsque le FFS reçoit un contenu à enregistrer, étape D1, il reçoit selon l'invention, un contenu chiffré I_c ainsi qu'une empreinte E associée.

Étant donné que le contenu chiffré I_c est chiffré avec sa propre valeur de hachage H, deux contenus identiques chiffrés avec leurs propres valeurs de hachage sont également identiques après chiffrage. Ainsi, le FSS peut facilement comparer deux contenus chiffrés, étape D2. Lors d'une étape D3, le FSS effectue une comparaison du contenu chiffré I_c reçu avec les contenus du FSS. Si cette comparaison est fructueuse le contenu I_c étant déjà stocké, il n'est pas une nouvelle fois enregistré. Par contre, si cette comparaison se révèle infructueuse, le nouveau contenu chiffré est enregistré avec l'empreinte associée.

La **figure 6** représente un mode de réalisation mettant en œuvre la recherche de contenus similaires dans le FSS.

Comme nous l'avons vu précédemment, la recherche de duplicatas a comme objectif l'optimisation du stockage chez le FSS et elle fait appel à des techniques de hachage cryptographique, et à titre illustratif un hachage du type SHA-256. Le résultat de la fonction de hachage change radicalement si un seul bit d'entrée change. Prenons l'exemple d'une même image sauvegardée avec deux formats de compression différents (e.g., BMP et JPEG), les valeurs de hachage de ces deux fichiers sont complètement différentes alors que les images sont visuellement similaires. Ce problème existe avec tous les contenus multimédias, dont les documents audio (e.g. Mp3, Flac), ou vidéo. (MPEG, Ogg, QuickTime). En conséquence, les techniques de hachage cryptographique sont utiles pour identifier des copies strictement identiques (bit à bit) d'une image mais elles sont inutiles pour la recherche d'images visuellement similaires à une image de référence.

Selon cet aspect de l'invention relatif à la recherche de contenu similaire, et sur demande d'un utilisateur autorisé, le FSS peut effectuer des requêtes de recherche de contenus multimédias similaires, par la méthode dite du plus proche voisin, dans les collections des données multimédias des utilisateurs sans avoir accès aux documents multimédias en clair. Le résultat de telles requêtes est équivalent au résultat que l'utilisateur aurait obtenu s'il avait exécuté les telles requêtes sur une collection de documents multimédias non chiffrées.

Pour résoudre le problème de la recherche de contenu similaire, on fait appel à des fonctions d'empreinte, ou d'ADN multimédia dédiées au type de contenus multimédias concernés. C'est par exemple le cas pour les empreintes d'images, appelées aussi descripteurs d'images. Ces descripteurs ont la particularité d'être tolérants aux distorsions des images. Il existe deux grandes classes de descripteurs :

- L'approche globale, telle que l'histogramme des niveaux de gris, décrit le contenu de l'image dans son ensemble. Cet algorithme est rapide mais son descripteur d'image est peu résistant aux distorsions de l'image.

- L'approche locale, telle que les points d'intérêt, décrit le contenu de l'image comme une collection d'empreintes de morceaux d'images appartenant à la même image. Cet algorithme est complexe et lent mais son descripteur d'image est résistant à de nombreuses distorsions.

La similitude entre deux images A et B se détermine simplement par recherche exhaustive du plus proche voisin de chaque descripteur de l'image A dans l'ensemble des descripteurs de l'image B.

Le passage à l'échelle, c'est-à-dire la recherche de similitude entre une image A et l'ensemble d'images d'une bibliothèque d'images, est beaucoup plus complexe. Ce passage à l'échelle nécessite la mise en place d'un système efficace pour résoudre le problème, dit du plus proche voisin, défini comme suit : Soit une collection de points de données et un point de requête dans un espace métrique de dimension < n >, trouver le point de données qui est le plus proche du point de requête. La manière habituelle de mettre en application un tel système est la suivante.

Un ensemble de descripteurs, dit collection de points de données, est calculé sur une bibliothèque de photos données. Ensuite, quand une requête de similitude est lancée, l'empreinte de l'image de requête est calculée afin d'obtenir le point de requête, et ensuite le point de donnée le plus proche au point de requête est déterminé.

L'efficacité d'une recherche du plus proche voisin est évaluée en fonction des mesures dites de précision et de rappel de la requête. Ces mesures dépendent essentiellement de l'algorithme d'empreinte d'image et de l'algorithme de recherche du plus proche voisin.

Les algorithmes d'empreinte d'image existants sont divers et variés et à titre illustratif, nous pouvons en citer deux : BoF (Bag of Features) ; et VLAD (acronyme anglais de Vector of Locally Aggreated Descriptors), basés sur une représentation de l'image à base d'un vecteur des descripteurs SIFT agrégés localement. Comme algorithmes d'indexation/recherche de descripteurs nous

pouvons en citer aussi deux : LSH (acronyme anglais de Locality-Sensitive Hashing) et Hamming Embedding.

Dans la suite de la description, nous définissons l'empreinte d'image comme un vecteur de taille fixe Zn appartenant à un espace métrique. Pour rappel, la norme d'un vecteur Zn fournit une mesure de distance ; de sorte que Zn avec une norme de Zn définissent un espace métrique. Une des normes de vecteur les plus populaires est la distance Euclidienne (norme L2) mais d'autres normes de vecteur existent et peuvent être employées.

Il est important de mentionner l'influence de la dimension < n > de l'empreinte d'image sur l'efficacité d'indexation des bibliothèques de photos numériques à large échelle, et sur la précision et le rappel de la requête de la base de données. Les empreintes d'images à grande dimension fournissent habituellement une meilleure précision et un meilleur rappel que les empreintes d'images à petite dimension, mais il est plus difficile à indexer efficacement des empreintes d'images à grande dimension. La capacité de discrimination d'une empreinte d'image à petite dimension est inférieure à celle d'une empreinte d'image à grande dimension et pourrait ne pas être satisfaisante.

En référence à la *figure 1*, selon ce mode de réalisation, l'ordinateur du premier utilisateur calcule et transmet outre le contenu chiffré I_c, une empreinte E relative au contenu, constituant ainsi un couple d'information $<I_c, E>$.

L'organigramme de la **figure 6** illustre ce procédé. L'étape I1 est décrite à la *figure 4* et n'est pas détaillée ici.

A l'étape I2 l'ordinateur du premier utilisateur calcule une empreinte du contenu à transmettre, selon une des méthodes connues données précédemment. . Cette empreinte est une empreinte de recherche.

Lors d'une étape I3, la paire, contenu chiffré I_c et empreinte E est transmise au FSS pour archivage s'il n'y a pas duplication.

Dans un mode de réalisation préféré, la paire < I_c ; E > envoyée par le dispositif du premier utilisateur au FSS (220) est triée et stockée en deux bases de données différentes, c'est-à-dire, une base de données pour chacun des composants du couple. Il est important de souligner que l'espace de mémoire nécessaire pour stocker l'empreinte < E > est négligeable par rapport à l'espace nécessaire pour stocker l'image chiffrée I_c. Selon une variante, le FSS utilise une unique base de données pour stocker le couple < I_c ; E >.

Le procédé de réception par l'utilisateur destinataire du contenu I transmis par le premier utilisateur est le même que celui décrit en référence à la figure 9, l'utilisateur destinataire recevant en outre, l'empreinte avec le hachage chiffré H_c et le contenu chiffré I_c.

Une caractéristique très importante d'une bibliothèque de photos est la possibilité, pour les utilisateurs autorisés, d'interroger la base de données d'images en fonction de son contenu. Par exemple, les utilisateurs, qui peuvent être le premier utilisateur ou l'utilisateur destinataire, doivent pouvoir rechercher dans la bibliothèque de photos des images presque identiques ou des images semblables à une image présentée au système comme exemple.

Comme expliqué ci-dessus, la manière habituelle de répondre à une telle exigence est d'associer une empreinte E (un vecteur Zn appartenant à un espace métrique) à chacune des images de la bibliothèque de photos. Afin de déterminer si deux images sont presque identiques ou semblables, l'utilisateur calcule la distance Euclidienne (norme L2) entre les empreintes correspondantes aux deux images et il compare ce résultat à un seuil donné. Puisque le FSS a accès aux empreintes en clair des images, le FSS peut, sur demande des utilisateurs, lancer des requêtes sur la base de données des images. Nous devons mentionner que nous supposons que le FSS ne peut obtenir aucune information, concernant l'image en clair, par sa connaissance de l'empreinte en clair de l'image. Cela implique que l'empreinte de l'image ne permet pas de reconstituer l'image à partir de la connaissance de l'empreinte de l'image. Ainsi, on favorise les empreintes de petite dimension pour limiter la fuite d'information vers le FSS.

Ainsi, lorsque l'utilisateur destinataire transmet un triplet au FSS pour une recherche d'images similaires, le FSS peut effectuer une mesure de similarité sur les images qu'il stocke, en utilisant l'une des méthodes décrites précédemment et fournir à l'utilisateur destinataire une à plusieurs image similaire, associée(s) avec son hachage cryptographique et son empreinte.

Étant donné que la mesure de similarité est effectuée par comparaison d'empreintes de petite dimension, la précision obtenue peut être insuffisante et l'utilisateur destinataire peut recevoir un ou plusieurs faux positifs. Ainsi, le mode de réalisation proposé ci-après améliore la robustesse en diminuant le nombre de faux positifs tout en garantissant une confidentialité des données stockées sur le FSS.

La **figure 7** illustre ce mode de réalisation. L'étape I'1 est décrite à la *figure 4* et ne l'est pas à nouveau ici.

Lors de l'étape I'2, une empreinte de recherche est calculée, par exemple selon la méthode des VLADs pour produire une empreinte V.

Comme suite à l'étape I'2, on passe à une étape I'3 dans laquelle une empreinte de sélection S est calculée. S est un vecteur de sélection, appelé sac de descripteurs. Lors d'une étape I'4, on chiffre S. Préférentiellement, le chiffrement sera obtenu avantageusement avec la valeur de hachage cryptographique H. Dans une mise en œuvre alternative le chiffrement sera obtenu en utilisant la clé publique que l'utilisateur destinataire a transmis au premier utilisateur, utilisée également pour chiffrer la valeur de hachage cryptographique H.

Lors d'une étape I'5, on transmet le triplet comprenant le contenu chiffré I_c, l'empreinte de recherche V et l'empreinte de sélection chiffrée S_c, soit $<I_c, V, S_c>$.

La **figure 8** illustre la création d'une valeur de hachage chiffrée à l'attention d'un utilisateur destinataire. Ce processus suppose que l'utilisateur destinataire a au préalable fourni sa clé de chiffrement publique K_p au premier

utilisateur et que celui-ci à déjà transmis les données multimédias au FSS selon le principe de la *figure 1*.

Lors d'une étape H1 la valeur de hachage non chiffrée est fournie au dispositif, soit par sélection d'une clé stockée sur l'ordinateur du premier utilisateur, soit par un nouveau calcul (102) à partir des données multimédias choisies.

Lors de l'étape suivante H2, la valeur de hachage cryptographique H est chiffrée en utilisant un hachage du type asymétrique E_a, avec ladite clé publique K_p.

Enfin, lors d'une étape H3, un triplet < E, U, Hc >, comprenant, l'empreinte E du contenu multimédia, l'identifiant de l'utilisateur destinataire U et le hachage cryptographique chiffré H_c associé est transmis au FSS (120,220). Le FSS les transmet alors, à l'utilisateur destinataire. Le couple utilisateur U, hachage chiffré H_c, est enregistré et un lien est créé par le FSS entre le contenu enregistré et la pluralité de hachages associés enregistrés, car chacun des hachages est chiffré avec une clé publique différente et il est donc nécessaire de les enregistrer pour chaque utilisateur. Le couple utilisateur, hachage chiffré étant de petite taille, ceci ne pose pas de problème de place de stockage sur le FSS et reste négligeable par rapport à la taille des contenus multimédias stockés.

Dans une mise en œuvre alternative, le dispositif (100) utilisé par le premier utilisateur réalise l'ensemble des étapes décrites à la *figure 6* ou à la *figure 7* et celles de la *figure 8* pour procéder à un envoi groupé de l'ensemble des données. C'est le cas par exemple lorsque les données multimédias n'ont pas préalablement été transmises au FSS.

L'utilisateur destinataire du contenu multimédia récupère alors les données multimédias pour les utiliser.

Pour cela, en référence à la **figure 9**, l'utilisateur destinataire demande au FSS le contenu que le premier utilisateur a fait enregistrer sur le FSS (120,220) à son attention, à l'étape V1.

Le dispositif de l'utilisateur destinataire reçoit le contenu chiffré I_c et la valeur de hachage cryptographique chiffrée H_c associée. Grâce à la clé privée K_s de l'utilisateur destinataire, le dispositif déchiffre le hachage cryptographique H_c lors d'une étape V2. Ensuite, lors d'une étape V3, le contenu chiffré I_c est déchiffré à l'aide du hachage cryptographique déchiffré H, utilisé en tant que clé de déchiffrage. L'utilisateur destinataire peut alors lire le contenu I.

Lors de la réception des données multimédias, lesdites données peuvent également inclure une empreinte E associée. L'empreinte reçue permet à l'utilisateur destinataire de faire de la recherche de contenu similaire.

La **figure 10** illustre la recherche par l'utilisateur destinataire de l'ensemble des données similaires à un contenu de référence dans le FSS.

Lors d'une étape R1, l'utilisateur destinataire transmet au FSS une requête de recherche de données similaires contenant son identifiant U et l'empreinte V des données de référence pour lesquelles il souhaite retrouver des données similaires.

Lors d'une étape R2, le FSS effectue une recherche de données similaires en utilisant les empreintes de recherche V stockées avec celle transmise par l'utilisateur.

Dans un mode de réalisation préféré, lors d'une étape R3, le FSS transmet à l'utilisateur destinataire l'ensemble des triplets $<H_c, V, S_c>$ correspondant à l'ensemble des empreintes proches voisins de V qu'il trouve. En raison de la faible précision de la recherche effectuée par le FSS, l'utilisateur destinataire reçoit un certain nombre d'empreintes de données multimédias qui sont des faux positifs.

Lors d'une étape R4 l'ordinateur de l'utilisateur destinataire déchiffre, avec la clé privée de l'utilisateur destinataire, les valeurs de hachage cryptographiques (H_c), pour obtenir la valeur de hachage non chiffrée (H). Lesdites valeurs de hachage obtenues (H) sont alors utilisées pour déchiffrer les empreintes d'images de sélection S_c des images reçues afin d'obtenir les empreintes d'images en clair S qu'il utilise, lors d'une étape R5, pour éliminer les faux positifs de l'ensemble des triplets $<H_c, S_c, V>$ qu'il a reçu en provenance du FSS. Cela est rendu possible car l'utilisateur destinataire possède sa clé privée et les empreintes S sont de dimension suffisamment grande pour détecter les faux positifs.

Lors d'une étape R6, l'utilisateur destinataire envoie une seconde requête au FSS pour demander les doublets $<I_c, H_c>$ des données multimédias sélectionnées. L'empreinte associée à chacun desdits doublets peut également être fournie par le FSS.

Dans un mode de réalisation alternatif, lors de lors d'une étape R3, le FSS transmet à l'utilisateur destinataire l'ensemble des doublets $<S_c, V>$ correspondant à l'ensemble des empreintes proches voisins de V qu'il trouve, avec le même problème de précision évoqué précédemment.

Dans ce mode alternatif, lors d'une étape R4 l'ordinateur de l'utilisateur destinataire déchiffre, avec la clé privée K_s de l'utilisateur destinataire, les empreintes d'images de sélection S_c des images reçues afin d'obtenir les empreintes d'images en clair S qu'il utilise, lors d'une étape R5, pour éliminer les faux positifs de l'ensemble des doublets $<S_c, V>$ qu'il a reçu en provenance du FSS. Cela est rendu possible car l'utilisateur destinataire possède sa clé privée et les empreintes S sont de dimension suffisamment grande pour détecter les faux positifs.

L'étape R6 du mode alternatif est identique au mode de réalisation préféré.

Nous pouvons noter qu'un avantage supplémentaire de l'invention est aussi la protection du droit d'auteur. En effet, si le FSS comprend une photo de

référence en clair, il peut déterminer si les utilisateurs ont des copies identiques à cette photo de référence. Par exemple, si le FSS a une photo de la tour Eiffel, réalisée par un photographe de renom, il peut déterminer si un ou plusieurs de ses clients a une copie identique de cette photo, stockée dans leurs collections de photos.

Si le FSS n'a pas de copie en clair des photos stockées dans ses serveurs, il lui est impossible d'exploiter la base de données d'images chiffrées stockées chez lui. D'autre part, si le FSS a légalement une copie en clair d'une photo dont il connaît le propriétaire, il peut déterminer si les utilisateurs de son service de stockage en ligne ont une copie illégale de cette photo stockée dans ses serveurs.

Bien que la description porte principalement sur un contenu de type image, l'invention ne se limite pas aux modes de réalisation décrits précédemment. Comme cela a été dit, ce système de stockage en ligne FSS est également applicable pour d'autres types de documents et notamment des vidéos.

Pour rappel, l'invention fait appel à deux techniques connues, le chiffrement convergent, permettant l'application de la déduplication, et l'empreinte de document pour pouvoir faire de la recherche de documents similaires.

À propos de la déduplication, la technique de chiffrement appliquée reste indépendante du type de document multimédia à traiter, elle garantit la possibilité d'appliquer la déduplication lors du stockage des documents multimédias. C'est sur l'obtention de l'empreinte, afin de permettre la recherche de documents similaires que des différences apparaissent. L'homme du métier sait que les méthodes de calcul d'empreintes sont spécifiques au type de contenu. Cependant, la comparaison pour identifier les contenus multimédias similaires reste basée sur la technique du plus proche voisin, déjà décrite à la *figure 6*.

Par exemple, l'invention, selon une première variante, peut s'appliquer à des documents du type audio. En effet, pour des problématiques de recherche dans les bases de données, de falsification / authentification, de filigrane des documents audio, de nombreuses méthodes de création d'empreintes ont été mises au point. Les algorithmes de calcul d'empreintes de documents audio sont multiples. Cela est décrit notamment dans la publication « Robust audio hashing for audio identification » de Hamza Özer, Bülent Sankur et Nasir Memon, faite en 2001 (Proc. Content-Based Multimedia Indexing).

De même, selon une seconde variante, il est possible de traiter des documents du type vidéo. Tout comme pour les autres types de documents, il existe des méthodes de calcul d'empreinte. Par exemple, une description de calcul d'empreinte est faite dans les publications numéro 2297-23000, "A video fingerprint based on visual digest and local fingerprints" par Massoudi, A., Lefebvre, F., Demarty, C.-H., Oisel, L. and Chupeau, B (Proc. IEEE Int. Conf. on Image Processing 2006) et numéro 3411-3414 "Global motion estimation for MPEG-encoded streams", de Coudray R. et Besserer B., (Proc. IEEE Int. Conf. on Image Processing 2004)

Selon une autre variante, il est également possible de traiter des documents du type texte. Les méthodes d'empreinte texte ont été développées notamment pour de la détection de plagiat dans les documents sous forme électronique. La publication numéro 342-353 « New Algorithms for Text Fingerprinting » par Roman Kolpakov, Mathieu Raffinot (au Combinatorial Pattern Matching, 2006) décrit des méthodes d'obtention d'empreinte texte.

Selon une variante, les données multimédias chiffrées Ic et les empreintes E associées sont stockées chez deux FSS différents (FSS1 et FSS2), l'un stockant les données multimédias chiffrées (FSS1), l'autre les empreintes (FSS2). Cela présente l'avantage de rassurer le premier utilisateur à propos du niveau de confidentialité de ses données. En effet, le dépositaire des données multimédias ne possédant pas d'empreintes et en particulier le vecteur de recherche qui n'est pas chiffré, il ne peut pas à l'aide de ce vecteur

de recherche tenter de reconstituer les données multimédias du premier utilisateur, dans le cas par exemple où le vecteur de recherche est de grande dimension. Il ne peut pas non plus faire de recherche de données multimédias similaires. Pour que cela fonctionne, il faut également que les valeurs de hachages chiffrées associées aux données multimédias soient stockées sur au moins l'un des FSS, toujours avec l'identifiant de l'utilisateur pour qui la valeur a été chiffrée.

Cependant, la méthode nécessite une coordination entre les FSS (ici FSS1 et FSS2) pour maintenir la relation qui existe entre le triplet, données multimédias, valeurs de hachage chiffrées pour un utilisateur et empreinte, mais elle garantit la possibilité d'appliquer la déduplication. Pour effectuer une recherche de données multimédias similaires, le dispositif de réception de l'utilisateur destinataire envoie une requête à celui des FSS stockant les empreintes E (ici FSS2). À réception de la réponse, le dispositif de réception de l'utilisateur destinataire demande au second FSS les données multimédias sélectionnées (ici FSS1). Il est possible de passer par une étape de recherche (à l'aide du vecteur de recherche), suivie d'une étape d'élimination des faux positifs à l'aide de vecteurs de sélection pour filtrer les données multimédias à demander au dépositaire de celle-ci.

Ce principe de répartition entre deux FSS peut s'étendre avec un nuage comprenant une multitude de FSS répartis dans de multiples pays tout autour de la planète.

REVENDICATIONS

1. Dispositif de stockage (120,220) en ligne apte à stocker des données multimédias (I) caractérisé en ce qu'il comprend des moyens d'enregistrer (121,122,223) lesdites données multimédias sous forme chiffrée (I_c) et pour chacune desdites données multimédias chiffrées (I_c), une empreinte (E) associée, lesdites données multimédias chiffrées (I_c) étant chiffrées en utilisant une valeur de hachage (H), ladite valeur de hachage (H) étant obtenue par le hachage de lesdites données multimédias non chiffrées (I), ladite empreinte (E) comprenant un vecteur de recherche non chiffré (V).

2. Dispositif selon la revendication 1 caractérisé en ce que ladite empreinte comprend un vecteur de vérification chiffré (S_c).

3. Dispositif selon la revendication 2, caractérisé en ce que ledit vecteur de vérification chiffré (S_c), est obtenu par le chiffrage dudit vecteur de vérification (S_c) avec ladite valeur de hachage (H).

4. Dispositif selon la revendication 3, apte à stocker au moins un identifiant d'utilisateur (U) caractérisé en ce que pour chaque utilisateur (U) le dispositif est apte à stocker un hachage chiffré (H_c) par utilisateur (U), ledit hachage chiffré (S_c) étant chiffré en utilisant une clé publique de chiffrement (K_p) dudit utilisateur (U).

5. Dispositif selon la revendication 2, caractérisé en ce qu'il est apte à stocker au moins un identifiant d'utilisateur (U) et pour chaque utilisateur un hachage chiffré (H_c), ledit hachage chiffré (H_c) et ledit vecteur de vérification (S_c) chiffré étant chacun chiffrés en utilisant une clé publique de chiffrement (K_p) dudit utilisateur (U).

6. Dispositif selon la revendication 1 caractérisé en ce qu'il comprend des moyens de comparer (124), lors de chaque enregistrement de données multimédias (I), lesdites données multimédias chiffrées enregistrées (I_c) avec les données multimédias chiffrées à enregistrer (I_c) de manière à ne pas dupliquer les données multimédias enregistrées (I_c).

7. Dispositif selon la revendication 4 ou 5 caractérisé en ce qu'il comprend des moyens :
   - de recevoir une première requête d'un utilisateur émetteur, comprenant pour un document multimédia cible recherché, au moins ladite empreinte associée audit document chiffré,
   - de comparer ledit vecteur de recherche (V) reçu à au moins un vecteur de recherche (V) enregistré en mesurant une distance entre ledit vecteur de recherche (V) reçu et le au moins un vecteur de recherche (V) enregistré,
   - de transmettre à l'émetteur de ladite première requête, au moins un vecteur de recherche (V) dont ladite distance avec ledit vecteur de recherche (V) reçu est inférieure à un seuil prédéterminé, dit vecteur sélectionné,

8. Dispositif selon la revendication 7 caractérisé en ce qu'il comprend des moyens :
   - de recevoir une seconde requête dudit émetteur, pour au moins un document multimédia recherché, comprenant au moins un vecteur de recherche (V) sélectionné parmi le au moins un vecteur de recherche (V) transmis.
   - de transmettre audit émetteur de ladite seconde requête, au moins un vecteur de vérification (S_c) chiffré correspondant audit au moins un vecteur de recherche (V) sélectionné.

9. Dispositif selon la revendication 4 ou 5 caractérisé en ce qu'il comprend des moyens :

- de recevoir une première requête d'un utilisateur émetteur, comprenant pour un document multimédia cible recherché, au moins ladite empreinte associée audit document chiffré,
- de comparer ledit vecteur de recherche (V) reçu à au moins un vecteur de recherche (V) enregistré en mesurant une distance entre ledit vecteur de recherche (V) reçu et le au moins un vecteur de recherche (V) enregistré,
- de transmettre à l'émetteur de ladite première requête, au moins un vecteur de vérification (S_c) chiffré correspondant audit vecteur de recherche (V) sélectionné.

10. Dispositif selon la revendication 8 ou 9 caractérisé en ce qu'il comprend, suite à la transmission dudit au moins un vecteur de vérification des moyens :
    - de recevoir une requête dudit émetteur pour au moins un document multimédia recherché, comprenant au moins un vecteur de vérification (S_c) sélectionné parmi le au moins un vecteur de vérification (S_c) transmis.
    - de transmettre à l'émetteur de ladite requête, au moins une paire correspondant à au moins un dit vecteur de vérification (S_c) reçu, comprenant au moins une donnée chiffrée (I_c) et au moins un hachage chiffré (H_c) correspondant.

11. Procédé de stockage en ligne apte à stocker des données multimédias (I_c) caractérisé en ce qu'il comprend les étapes de :
    - réception (D1) de données, comprenant un contenu multimédia sous forme chiffré (I_c) selon un chiffrement convergent, une empreinte associée (E).
    - comparaison (D2) dudit contenu multimédia chiffré (I_c) avec ceux préalablement stockés.
    - enregistrement (D4) desdites données, si ledit contenu n'est pas déjà stocké.

12. Procédé de transmission de données multimédias caractérisé en ce qu'il comprend les étapes de :
    - chiffrement desdites données multimédias (I) à l'aide d'une méthode de chiffrement convergent ;
    - calcul d'une empreinte (E), desdites données multimédias (I) à partir desdites données multimédias non chiffrées (I), ladite empreinte comprenant un vecteur de recherche non chiffré (V) ;
    - transmission d'un ensemble de données, comprenant au moins lesdites données multimédias chiffrées (I_c) et ladite empreinte (E) associée.
13. Dispositif de transmission de données multimédias caractérisé en ce que pour la transmission de données multimédias, il comprend des moyens :
    - de chiffrement (105) desdites données multimédias (I) à l'aide d'une méthode de chiffrement convergent ;
    - de calcul (104) d'une empreinte (E), desdites données multimédias (I) à partir desdites données multimédias non chiffrées (I) ;
    - de transmission (101) d'un ensemble de données, comprenant au moins lesdites données multimédias chiffrées (I_c) et ladite empreinte (E) associée.

14. Procédé de réception de données multimédias, destinées à un utilisateur (U) possédant une clé publique (K_p) et une clé privée associée (K_s) caractérisé en ce qu'il comprend les étapes de :
    - réception (V1) d'un ensemble de données, comprenant au moins des données multimédias (I_c) sous forme chiffrée par une méthode de chiffrement convergent et une valeur de hachage chiffrée (H_c) associée auxdites données multimédias (I_c), destinée audit utilisateur ;

- déchiffrage (V2) de ladite valeur de hachage chiffré (H_c) avec ladite clé privée (K_s) pour obtenir une valeur de hachage non chiffrée (H) ;
- déchiffrage (V3) desdites données multimédias (I_c) avec ladite valeur de hachage (H) pour obtenir les données multimédias non chiffrées (I).

15. Dispositif de réception de données multimédias, destinées à un utilisateur (U) possédant une clé publique (K_p) et une clé privée associée (K_s) caractérisé en ce qu'il comprend des moyens :
    - de réception (141,341) d'un ensemble de données, comprenant au moins des données multimédias (I_c) sous forme chiffrée par une méthode convergente et une valeur de hachage chiffrée (H_c) associée auxdites données multimédias (I_c), destinée audit utilisateur U ;
    - de déchiffrage (142) de ladite valeur de hachage chiffré (H_c) avec ladite clé privée (K_s) pour obtenir une valeur de hachage non chiffrée (H) ;
    - de déchiffrage (143) desdites données multimédias (I_c) avec ladite valeur de hachage (H) pour obtenir les données multimédias non chiffrées (I).

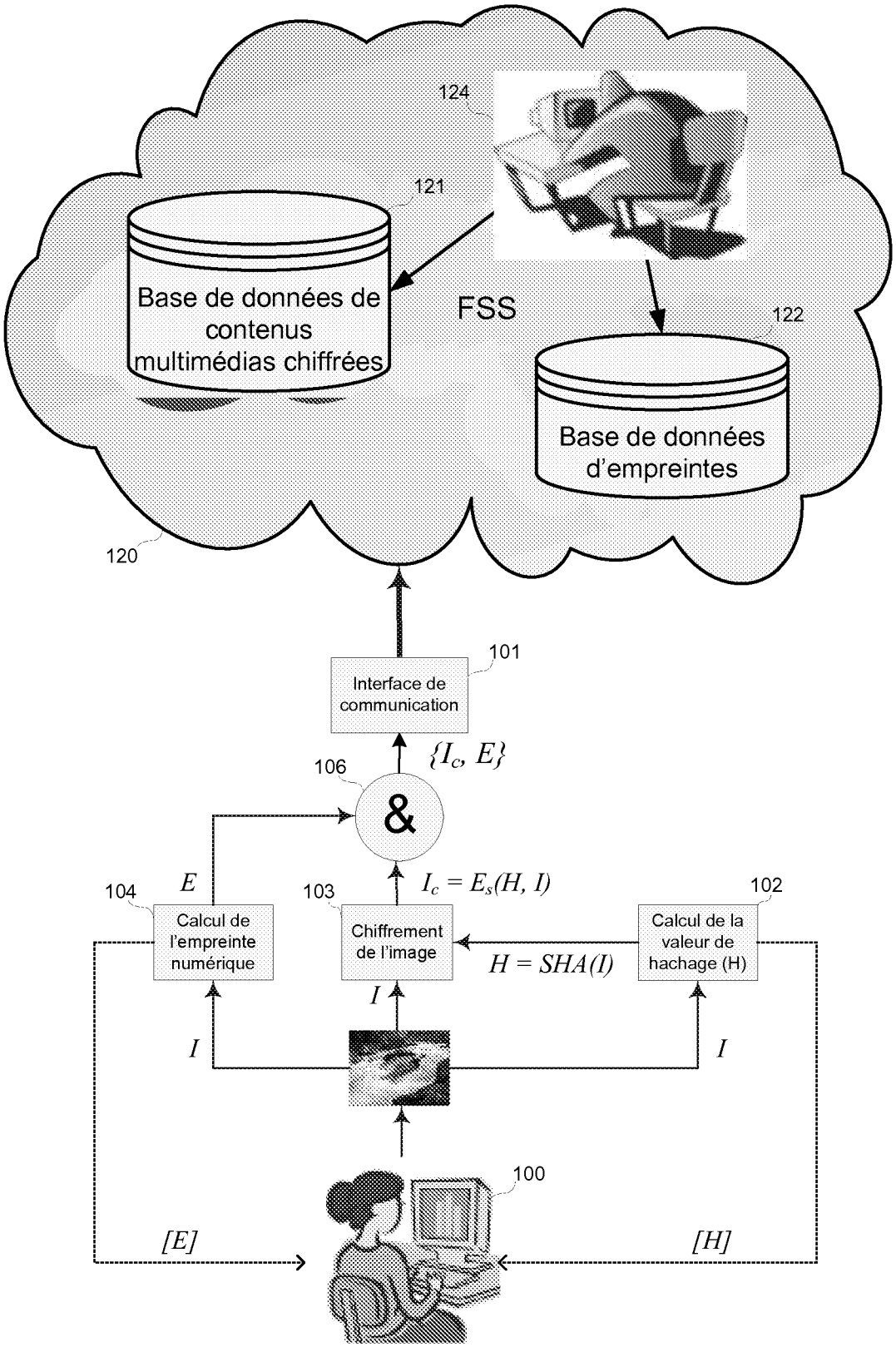
124
121
Base de données de contenus multimédias chiffrées
FSS
122
Base de données d'empreintes
120
101
Interface de communication
{Ic, E}
106
&
104
E
Calcul de l'empreinte numérique
103
Ic = Es(H, I)
Chiffrement de l'image
102
Calcul de la valeur de hachage (H)
H = SHA(I)
I
I
I
100
[E]
[H]

Fig. 1

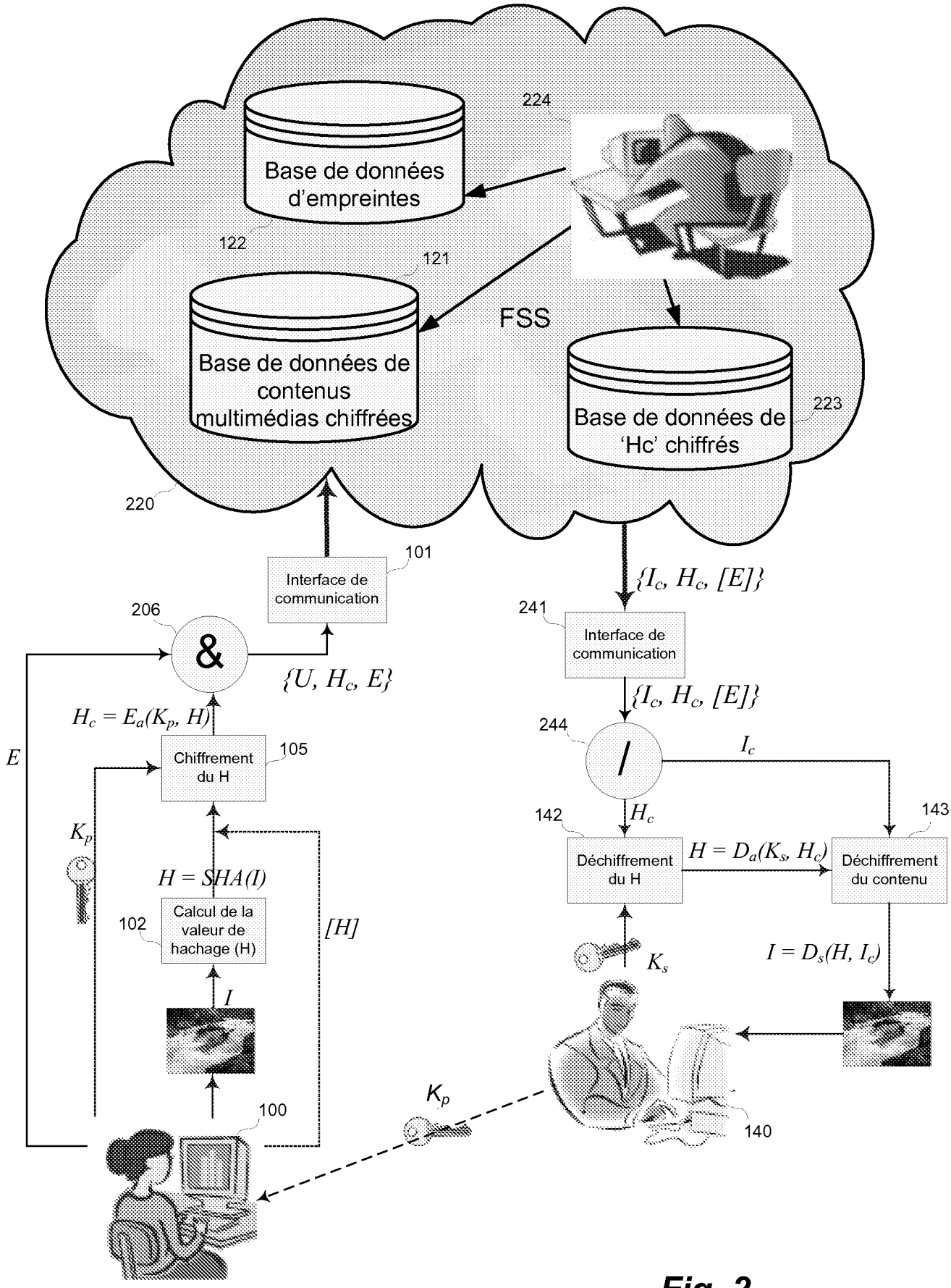
Base de données d'empreintes
122
224
121
Base de données de contenus multimédias chiffrées
FSS
Base de données de 'Hc' chiffrés
223
220
101
Interface de communication
206
&
{U, Hc, E}
Hc = Ea(Kp, H)
E
Chiffrement du H
105
Kp
H = SHA(I)
102
Calcul de la valeur de hachage (H)
[H]
I
100
Kp
{Ic, Hc, [E]}
241
Interface de communication
{Ic, Hc, [E]}
244
/
Ic
142
Hc
143
Déchiffrement du H
H = Da(Ks, Hc)
Déchiffrement du contenu
Ks
I = Ds(H, Ic)
140

Fig. 2

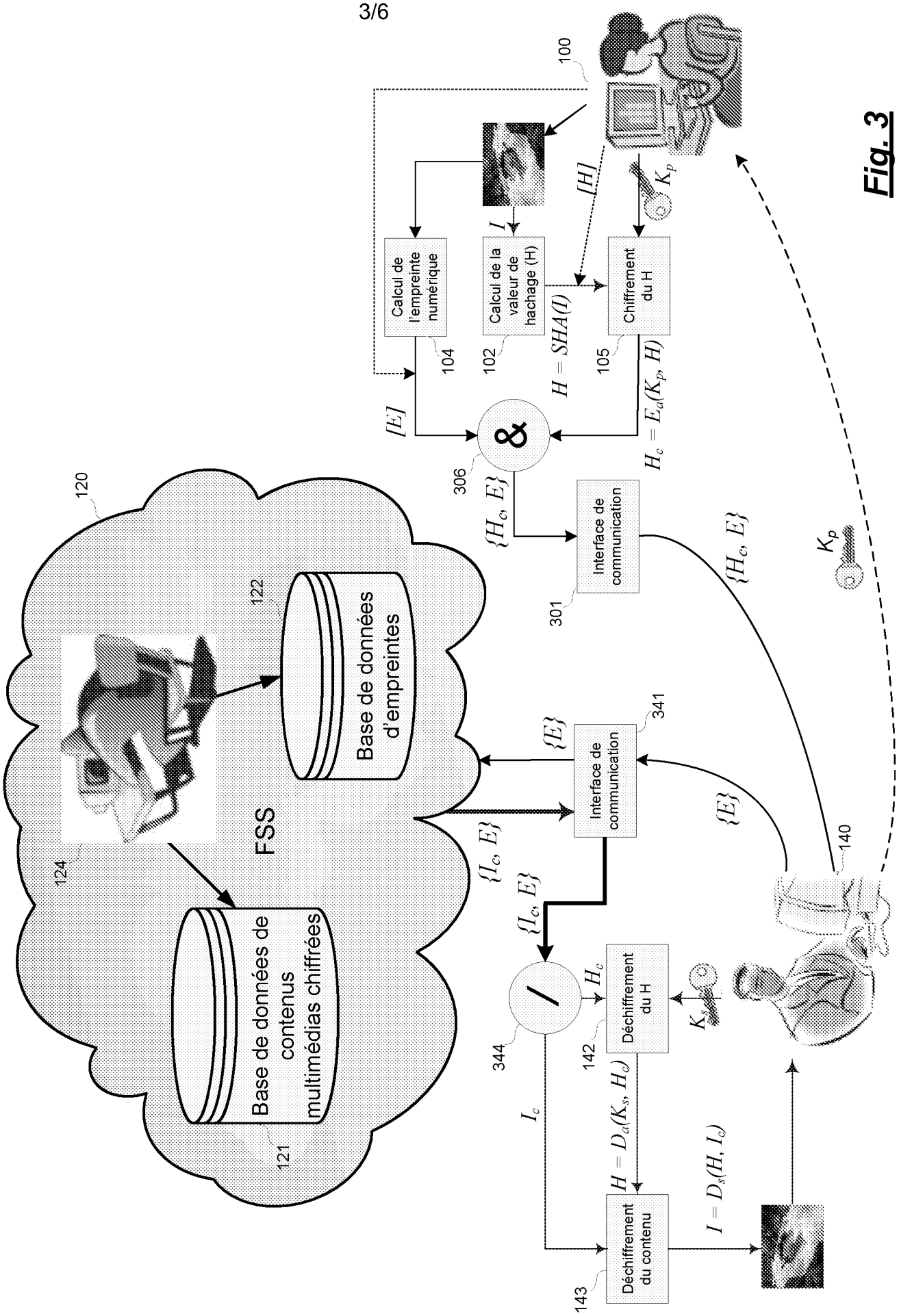
100
Calcul de l'empreinte numérique
104
Calcul de la valeur de hachage (H)
102
Chiffrement du H
105
[E]
I
[H]
K_p
$H = SHA(I)$
$H_c = E_a(K_p, H)$
&
306
$\{H_c, E\}$
Interface de communication
301
120
122
Base de données d'empreintes
FSS
124
Base de données de contenus multimédias chiffrées
121
$\{E\}$
341
Interface de communication
$\{I_c, E\}$
$\{I_c, E\}$
/
344
H_c
Déchiffrement du H
142
K_s
I_c
$H = D_a(K_s, H_c)$
Déchiffrement du contenu
143
$I = D_s(H, I_c)$
$\{H_c, E\}$
K_p
140

Fig. 3

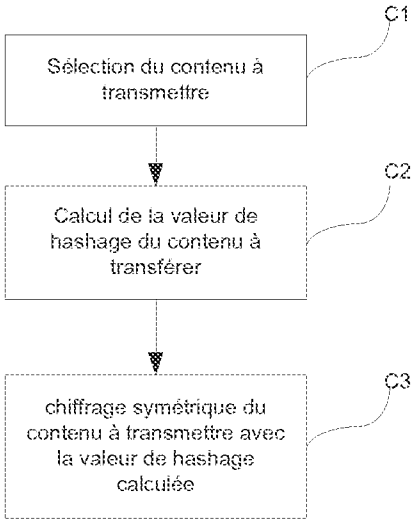
C1
Sélection du contenu à transmettre
C2
Calcul de la valeur de hashage du contenu à transférer
C3
chiffrage symétrique du contenu à transmettre avec la valeur de hashage calculée

Fig. 4

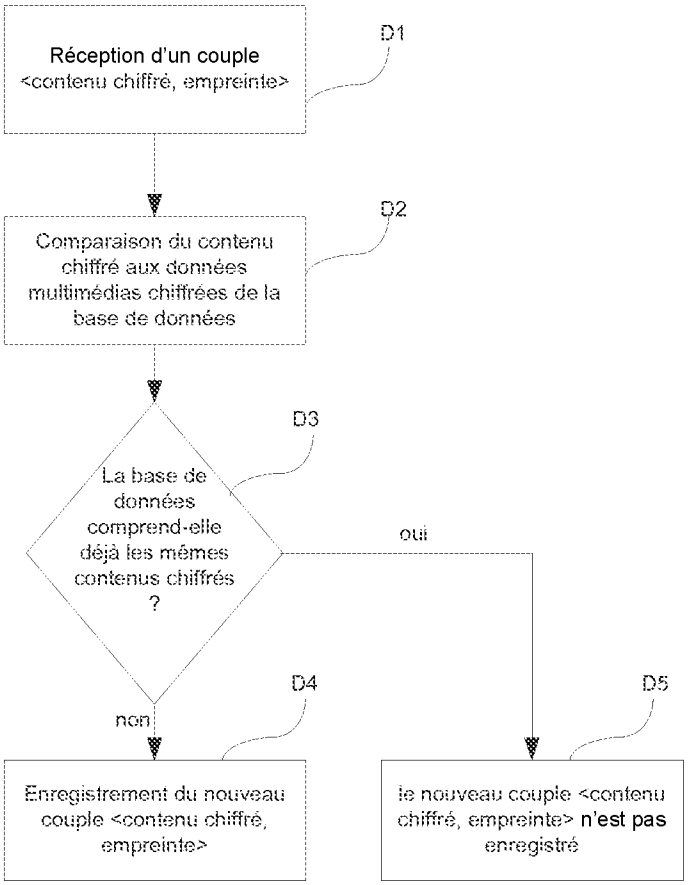
D1
Réception d'un couple <contenu chiffré, empreinte>
D2
Comparaison du contenu chiffré aux données multimédias chiffrées de la base de données
D3
La base de données comprend-elle déjà les mêmes contenus chiffrés ?
oui
non
D4
Enregistrement du nouveau couple <contenu chiffré, empreinte>
D5
le nouveau couple <contenu chiffré, empreinte> n'est pas enregistré

Fig. 5

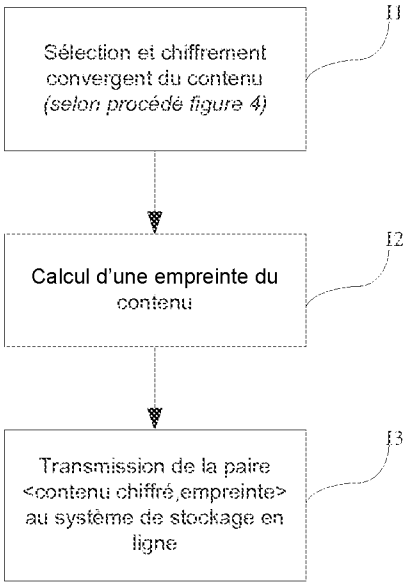
I1
Sélection et chiffrement convergent du contenu (selon procédé figure 4)
I2
Calcul d'une empreinte du contenu
I3
Transmission de la paire <contenu chiffré,empreinte> au système de stockage en ligne

Fig. 6

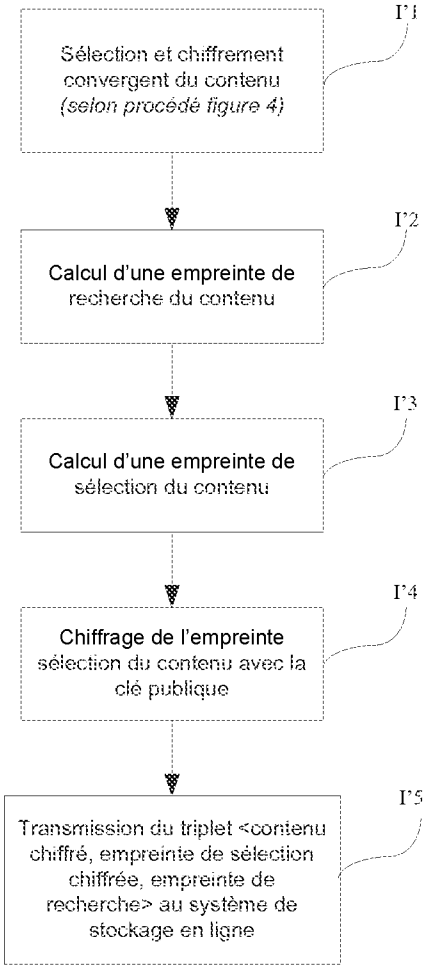
I'1
Sélection et chiffrement convergent du contenu (selon procédé figure 4)
I'2
Calcul d'une empreinte de recherche du contenu
I'3
Calcul d'une empreinte de sélection du contenu
I'4
Chiffrage de l'empreinte sélection du contenu avec la clé publique
I'5
Transmission du triplet <contenu chiffré, empreinte de sélection chiffrée, empreinte de recherche> au système de stockage en ligne

Fig. 7

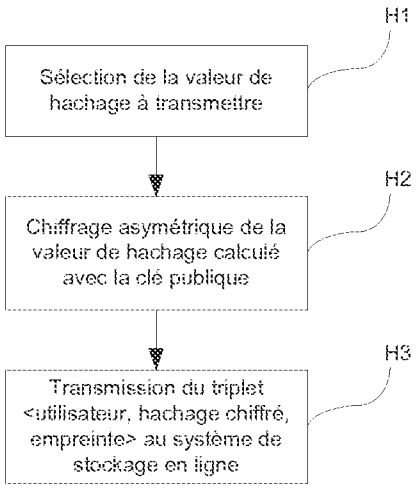

Fig. 8

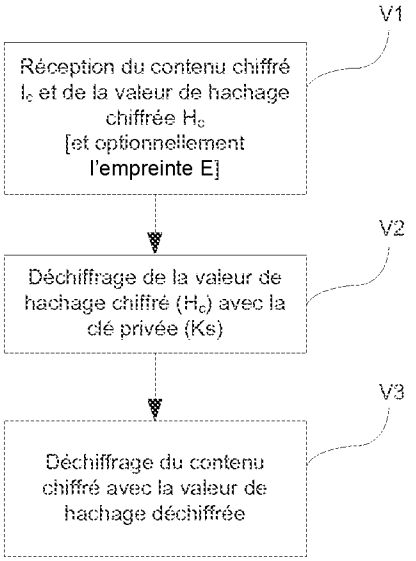

Fig. 9

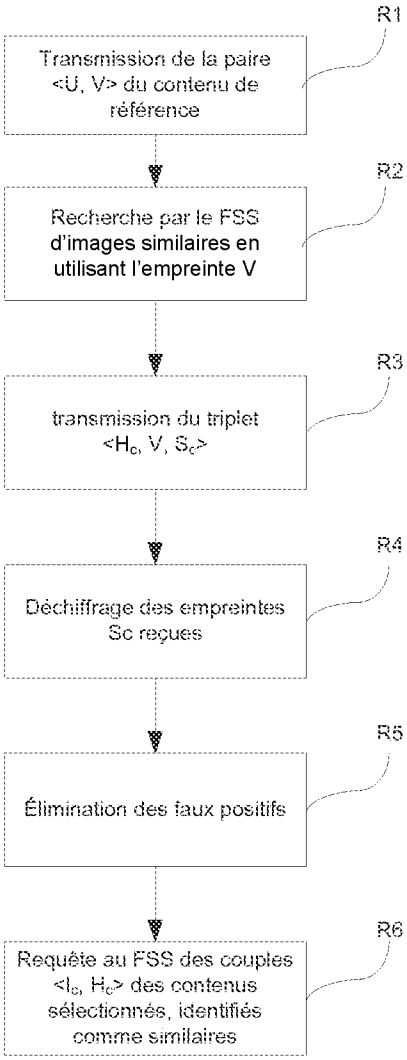

Fig. 10

INTERNATIONAL SEARCH REPORT

International application No
PCT/FR2011/053034

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F21/24 G06F17/30
ADD.

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

B. FIELDS SEARCHED

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

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, INSPEC, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	US 2004/162808 A1 (MARGOLUS NORMAN H [US] ET AL) 19 August 2004 (2004-08-19) paragraphs [0010], [0011], [0020], [0021], [0048], [0059], [0060], [0081] - [0083] figures 1, 5	1-6, 11-15 7-10

X A	WO 03/019412 A2 (DATACT TECHNOLOGIES N V [BE]; DE SPIEGELEER KRISTOF [BE]) 6 March 2003 (2003-03-06) page 1, line 4 - page 1, line 7 page 3, line 4 - page 3, line 22 page 5, line 12 - page 6, line 27 figures 1, 2	1,11-15 2-6

	-/--	

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
11 June 2012	21/06/2012

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Volpato, Gian Luca

INTERNATIONAL SEARCH REPORT

International application No
PCT/FR2011/053034

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2010/313040 A1 (LUMB CHRISTOPHER R [US]) 9 December 2010 (2010-12-09) paragraphs [0021], [0032], [0039], [0042] - [0047] figures 4, 5	1,6, 11-15
Y	US 2010/185615 A1 (MONGA VISHAL [US]) 22 July 2010 (2010-07-22) paragraphs [0001], [0005], [0006], [0028], [0033] - [0038] figures 2A, 2B	7-10
A	US 6 084 595 A (BACH JEFFREY R [US] ET AL) 4 July 2000 (2000-07-04) column 1, line 9 - column 1, line 12 column 2, line 13 - column 3, line 2 column 4, line 24 - column 4, line 31 column 5, line 59 - column 6, line 13 column 8, line 54 - column 9, line 9; figure 3	7-10
A	US 6 463 432 B1 (MURAKAWA AKIRA [JP]) 8 October 2002 (2002-10-08) column 1, line 11 - column 1, line 15 column 2, line 33 - column 2, line 46 column 6, line 15 - column 6, line 22 column 7, line 20 - column 9, line 18 figure 7	7-10

INTERNATIONAL SEARCH REPORT

International application No.

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See additional sheet

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/FR2011/053034

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2004162808 A1	19-08-2004	AU 3852401 A	27-08-2001
		US 2002038296 A1	28-03-2002
		US 2004139098 A1	15-07-2004
		US 2004139303 A1	15-07-2004
		US 2004143578 A1	22-07-2004
		US 2004143743 A1	22-07-2004
		US 2004143744 A1	22-07-2004
		US 2004143745 A1	22-07-2004
		US 2004162808 A1	19-08-2004
		US 2004255140 A1	16-12-2004
		US 2005131903 A1	16-06-2005
		US 2005131904 A1	16-06-2005
		US 2005131905 A1	16-06-2005
		US 2005131961 A1	16-06-2005
		US 2010185855 A1	22-07-2010
		WO 0161438 A2	23-08-2001
WO 03019412 A2	06-03-2003	AU 2002304842 A1	10-03-2003
		CN 1543617 A	03-11-2004
		EP 1419457 A2	19-05-2004
		HK 1069651 A1	04-05-2007
		JP 4446738 B2	07-04-2010
		JP 2005501342 A	13-01-2005
		US 2004236803 A1	25-11-2004
		US 2008034021 A1	07-02-2008
		WO 03019412 A2	06-03-2003
US 2010313040 A1	09-12-2010	NONE	
US 2010185615 A1	22-07-2010	NONE	
US 6084595 A	04-07-2000	NONE	
US 6463432 B1	08-10-2002	NONE	

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°
PCT/FR2011/053034

A. CLASSEMENT DE L'OBJET DE LA DEMANDE
INV. G06F21/24 G06F17/30
ADD.

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)
G06F

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés)
EPO-Internal, INSPEC, WPI Data

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X Y	US 2004/162808 A1 (MARGOLUS NORMAN H [US] ET AL) 19 août 2004 (2004-08-19) alinéas [0010], [0011], [0020], [0021], [0048], [0059], [0060], [0081] - [0083] figures 1, 5	1-6, 11-15 7-10

X A	WO 03/019412 A2 (DATACT TECHNOLOGIES N V [BE]; DE SPIEGELEER KRISTOF [BE]) 6 mars 2003 (2003-03-06) page 1, ligne 4 - page 1, ligne 7 page 3, ligne 4 - page 3, ligne 22 page 5, ligne 12 - page 6, ligne 27 figures 1, 2	1,11-15 2-6

	-/--	

[X] Voir la suite du cadre C pour la fin de la liste des documents

[X] Les documents de familles de brevets sont indiqués en annexe

* Catégories spéciales de documents cités:

"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent

"E" document antérieur, mais publié à la date de dépôt international ou après cette date

"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)

"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens

"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée	Date d'expédition du présent rapport de recherche internationale
11 juin 2012	21/06/2012

Nom et adresse postale de l'administration chargée de la recherche internationale	Fonctionnaire autorisé
Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Volpato, Gian Luca

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°
PCT/FR2011/053034

C(suite). DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	US 2010/313040 A1 (LUMB CHRISTOPHER R [US]) 9 décembre 2010 (2010-12-09) alinéas [0021], [0032], [0039], [0042] - [0047] figures 4, 5	1,6, 11-15
Y	US 2010/185615 A1 (MONGA VISHAL [US]) 22 juillet 2010 (2010-07-22) alinéas [0001], [0005], [0006], [0028], [0033] - [0038] figures 2A, 2B	7-10
A	US 6 084 595 A (BACH JEFFREY R [US] ET AL) 4 juillet 2000 (2000-07-04) colonne 1, ligne 9 - colonne 1, ligne 12 colonne 2, ligne 13 - colonne 3, ligne 2 colonne 4, ligne 24 - colonne 4, ligne 31 colonne 5, ligne 59 - colonne 6, ligne 13 colonne 8, ligne 54 - colonne 9, ligne 9; figure 3	7-10
A	US 6 463 432 B1 (MURAKAWA AKIRA [JP]) 8 octobre 2002 (2002-10-08) colonne 1, ligne 11 - colonne 1, ligne 15 colonne 2, ligne 33 - colonne 2, ligne 46 colonne 6, ligne 15 - colonne 6, ligne 22 colonne 7, ligne 20 - colonne 9, ligne 18 figure 7	7-10

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°
PCT/FR2011/053034

Cadre n°. II Observations - lorsqu'il a été estimé que certaines revendications ne pouvaient pas faire l'objet d'une recherche (suite du point 2 de la première feuille)

Le rapport de recherche internationale n'a pas été établi en ce qui concerne certaines revendications conformément à l'article 17.2)a) pour les raisons suivantes :

1. ☐ Les revendications n^{os}
se rapportent à un objet à l'égard duquel l'administration chargée de la recherche internationale n'est pas tenue de procéder à la recherche, à savoir :

2. ☐ Les revendications n^{os}
parce qu'elles se rapportent à des parties de la demande internationale qui ne remplissent pas suffisamment les conditions prescrites pour qu'une recherche significative puisse être effectueé, en particulier :

3. ☐ Les revendications n^{os}
parce qu'elles sont des revendications dépendantes et ne sont pasrédigées conformément aux dispositions de la deuxième et de la troisième phrases de la règle 6.4.a).

Cadre n°. III Observations - lorsqu'il y a absence d'unité de l'invention (suite du point 3 de la première feuille)

L'administration chargée de la recherche internationale a trouvé plusieurs inventions dans la demande internationale, à savoir:

voir feuille supplémentaire

1. ☒ Comme toutes les taxes additionnelles exigées ont été payées dans les délais par le déposant, le présent rapport de recherche internationale porte sur toutes les revendications pouvant faire l'objet d'une recherche.

2. ☐ Comme toutes les revendications qui se prêtent à la recherche ont pu faire l'objet de cette recherche sans effort particulier justifiant des taxes additionnelles, l'administration chargée de la recherche internationale n'a sollicité le paiement d'aucunes taxes de cette nature.

3. ☐ Comme une partie seulement des taxes additionnelles demandées a été payée dans les délais par le déposant, le présent rapport de recherche internationale ne porte que sur les revendications pour lesquelles les taxes ont été payées, à savoir les revendications n^{os}:

4. ☐ Aucunes taxes additionnelles demandées n'ont été payées dans les délais par le déposant. En conséquence, le présent rapport de recherche internationale ne porte que sur l'invention mentionnée en premier lieu dans les revendications; elle est couverte par les revendications n^{os}:

Remarque quant à la réserve

☐ Les taxes additionnelles étaient accompagnées d'une réserve de la part du déposant et, le cas échéant, du paiement de la taxe de réserve.

☐ Les taxes additionnelles étaient accompagnées d'une réserve de la part du déposant mais la taxe de réserve n'a pas été payée dans le délai prescrit dans l'invitation.

☒ Le paiement des taxes additionnelles n'était assorti d'aucune réserve.

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale n°
PCT/FR2011/053034

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 2004162808 A1	19-08-2004	AU 3852401 A	27-08-2001
		US 2002038296 A1	28-03-2002
		US 2004139098 A1	15-07-2004
		US 2004139303 A1	15-07-2004
		US 2004143578 A1	22-07-2004
		US 2004143743 A1	22-07-2004
		US 2004143744 A1	22-07-2004
		US 2004143745 A1	22-07-2004
		US 2004162808 A1	19-08-2004
		US 2004255140 A1	16-12-2004
		US 2005131903 A1	16-06-2005
		US 2005131904 A1	16-06-2005
		US 2005131905 A1	16-06-2005
		US 2005131961 A1	16-06-2005
		US 2010185855 A1	22-07-2010
		WO 0161438 A2	23-08-2001
WO 03019412 A2	06-03-2003	AU 2002304842 A1	10-03-2003
		CN 1543617 A	03-11-2004
		EP 1419457 A2	19-05-2004
		HK 1069651 A1	04-05-2007
		JP 4446738 B2	07-04-2010
		JP 2005501342 A	13-01-2005
		US 2004236803 A1	25-11-2004
		US 2008034021 A1	07-02-2008
		WO 03019412 A2	06-03-2003
US 2010313040 A1	09-12-2010	AUCUN	
US 2010185615 A1	22-07-2010	AUCUN	
US 6084595 A	04-07-2000	AUCUN	
US 6463432 B1	08-10-2002	AUCUN	

SUITE DES RENSEIGNEMENTS INDIQUES SUR PCT/ISA/ 210

L'administration chargée de la recherche internationale a trouvé plusieurs (groupes d') inventions dans la demande internationale, à savoir:

1. revendications: 1-6, 11-15

   Dispositif et procédé de stockage en ligne

   ---

2. revendications: 7-10

   Dispositif et procédé pour rechercher des documents similaires

   ---