

12

DEMANDE DE BREVET D'INVENTION

A1

22 Date de dépôt : 25.07.00.

30 Priorité :

43 Date de mise à la disposition du public de la
demande : 01.02.02 Bulletin 02/05.

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 :

71 Demandeur(s) : CANON KABUSHIKI KAISHA — JP.

72 Inventeur(s) : LE FLOCH HERVE.

73 Titulaire(s) :

74 Mandataire(s) : RINUY SANTARELLI.

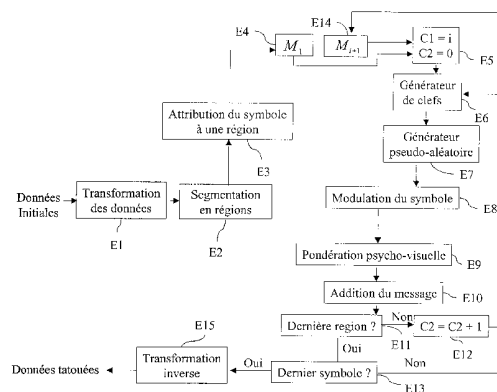
54 INSERTION ET EXTRACTION DE MESSAGE DANS DES DONNEES NUMERIQUES.

57 L'invention concerne un procédé d'insertion d'un message dans des données numériques représentatives de grandeurs physiques, le message comportant des symboles ordonnés, comportant les étapes de:

- segmentation (E2) des données en régions,
- association (E3) d'au moins une région à chaque symbole à insérer,

caractérisé en ce qu'il comporte, pour chaque région dans laquelle un symbole considéré est à insérer, les étapes de:

- détermination (E7) d'une fonction pseudo-aléatoire, à partir d'une clé qui dépend:
 - d'une clé initiale, et
 - de la longueur du message,
- modulation (E8) du symbole considéré par la fonction pseudo-aléatoire précédemment déterminée pour fournir une séquence pseudo-aléatoire,
- addition (E10) de la séquence pseudo-aléatoire à la région considérée.



5

10 La présente invention concerne un procédé d'insertion d'une information supplémentaire telle qu'une marque secrète dans un signal numérique.

Elle concerne également un procédé d'extraction d'une marque secrète insérée dans un signal numérique.

15 Corrélativement, la présente invention concerne un dispositif d'insertion d'une information supplémentaire et un dispositif d'extraction de l'information supplémentaire, adaptés respectivement à mettre en œuvre les procédés d'insertion et d'extraction conformes à l'invention.

20 Le signal numérique considéré dans la suite sera plus particulièrement un signal d'image numérique.

Les procédés d'insertion et extraction conformes à l'invention s'inscrivent dans le domaine technique du marquage (watermarking en anglais) des données numériques qui peut s'interpréter comme l'insertion d'un sceau dans les données numériques, permettant par exemple d'authentifier le contenu d'un fichier de données numériques. Ce marquage est également
25 appelé tatouage numérique.

Des méthodes de marquage sont connues, par exemple d'après le document de F. Hartung et B. Girod, Université de Erlangen-Nuremberg, « Watermarking of uncompressed and compressed video », Signal Processing
30 66 (1998), pp283-301, ou encore d'après le document US 5 915 027.

Cependant, aucun de ces documents ne mentionne la possibilité d'insérer un message dont la taille est inconnue lors de l'extraction de ce message.

La présente invention vise à remédier aux inconvénients de la technique antérieure, en fournissant un procédé et un dispositif d'insertion de message dans des données numériques qui permettent d'insérer un message dont la taille est inconnue lors de l'extraction ultérieure de ce message.

A cette fin, l'invention propose un procédé d'insertion d'un message dans des données numériques représentatives de grandeurs physiques, le message comportant des symboles ordonnés, comportant les étapes de :

- segmentation des données en régions,
- association d'au moins une région à chaque symbole à insérer, caractérisé en ce qu'il comporte, pour chaque région dans laquelle un symbole considéré est à insérer, les étapes de :
- détermination d'une fonction pseudo-aléatoire, à partir d'une clé qui dépend :
 - d'une clé initiale, et
 - de la longueur du message,

- modulation du symbole considéré par la fonction pseudo-aléatoire précédemment déterminée pour fournir une séquence pseudo-aléatoire,
- addition de la séquence pseudo-aléatoire à la région considérée.

Corrélativement, l'invention propose un dispositif d'insertion d'un message dans des données numériques représentatives de grandeurs physiques, le message comportant des symboles ordonnés, comportant :

- des moyens de segmentation des données en régions,
- des moyens d'association d'au moins une région à chaque symbole à insérer,

caractérisé en ce qu'il comporte, pour chaque région dans laquelle un symbole est à insérer :

- des moyens de détermination d'une fonction pseudo-aléatoire, pour chaque région dans laquelle un symbole considéré est à insérer, à partir d'une clé qui dépend :

- d'une clé initiale, et
- 5 - de la longueur du message,
- des moyens de modulation du symbole considéré par la fonction pseudo-aléatoire précédemment déterminée pour fournir une séquence pseudo-aléatoire,
- des moyens d'addition de la séquence pseudo-aléatoire à la région
- 10 considérée.

Le procédé et le dispositif selon l'invention permettent d'insérer une information supplémentaire, ou message, dont la longueur, exprimée en nombre de symboles, n'est pas connue lors de l'extraction ultérieure.

En outre, le procédé et le dispositif selon l'invention permettent

15 d'insérer une information supplémentaire dont la longueur est arbitraire, tout en restant inférieure au nombre de régions formées.

Selon une caractéristique préférée, la dépendance de la clé vis à vis de la longueur du message est assurée par la dépendance de la clé vis à vis :

- du nombre de fois où le symbole à insérer a déjà été
- 20 inséré dans d'autres régions, et
- du rang du symbole parmi les symboles ordonnés.

Ainsi, la mise en œuvre de l'invention demeure simple et ne nécessite pas de calculs complexes.

Selon une caractéristique préférée, le procédé comporte une étape

25 préalable de transformation des données numériques par une transformation réversible.

En effet, l'invention s'applique aussi bien à des données « originales », telles qu'une image, qu'à des données transformées.

L'invention concerne aussi un procédé d'extraction d'un message

30 dans des données numériques représentatives de grandeurs physiques, le message comportant des symboles ordonnés, comportant les étapes de :

- segmentation des données en régions,
- extraction de la longueur du message inséré,
- extraction du message inséré.

Le procédé d'extraction permet de retrouver le message qui a été
5 inséré selon l'invention.

Selon une caractéristique préférée, l'extraction de la longueur du message inséré comporte les étapes de :

- sélection d'un ensemble de valeurs de longueur, et
- calcul d'une corrélation entre le message et les données
10 numériques, pour chacune de ces valeurs,
- détermination d'un maximum local parmi les valeurs de corrélation.

La longueur du message est extraite de manière fiable.

Selon une autre caractéristique préférée, l'extraction de la longueur du message inséré est effectuée en traitant F fois moins de coefficients que
15 n'en comportent les données numériques. Les calculs sont ainsi accélérés.

Dans ce cas, le procédé comporte les étapes de :

- détermination du nombre total de coefficients à considérer,
- sélection d'un nombre maximum de coefficients correspondant à un même symbole inséré, puis, si le nombre total de coefficients à considérer n'est
20 pas atteint,
- réitération de l'étape de sélection, pour un autre symbole.

La corrélation est alors plus efficace, car elle est effectuée sur un nombre réduit de symboles. La détection est ainsi améliorée.

L'invention concerne un dispositif d'extraction mettant en œuvre les
25 caractéristiques précédentes. Le procédé et le dispositif d'extraction présentent des avantages analogues à ceux qui ont été présentés plus haut.

L'invention concerne aussi un appareil numérique incluant le dispositif d'insertion ou d'extraction, ou des moyens de mise en œuvre du procédé d'insertion ou d'extraction. Cet appareil numérique est par exemple un
30 appareil photographique numérique, un caméscope numérique, un scanner,

une imprimante, un photocopieur, un télécopieur. Les avantages du dispositif et de l'appareil numérique sont identiques à ceux précédemment exposés.

Un moyen de stockage d'information, lisible par un ordinateur ou par un microprocesseur, intégré ou non au dispositif, éventuellement amovible, 5 mémorise un programme mettant en œuvre le procédé d'insertion ou d'extraction.

Les caractéristiques et avantages de la présente invention apparaîtront plus clairement à la lecture d'un mode préféré de réalisation illustré par les dessins ci-joints, dans lesquels :

- 10 - la figure 1 représente un dispositif d'insertion d'un message dans des données numériques selon l'invention,
- la figure 2 représente un dispositif de détection de message inséré dans des données numériques selon l'invention,
- la figure 3 représente un dispositif mettant en œuvre l'invention,
- 15 - la figure 4 représente un procédé d'insertion d'un message dans des données numériques selon l'invention,
- la figure 5 représente un procédé de détection de message inséré dans des données numériques selon l'invention,
- les figures 6a, 6b, 7a et 7b représentent différents modes de 20 réalisation de procédé de détection de taille de message inséré dans des données numériques, inclus dans le procédé de la figure 5,
- la figure 8 représente un procédé d'accélération de calculs, inclus dans le procédé de la figure 7,
- la figure 9a représente un premier mode de réalisation de procédé 25 de détermination de message, inclus dans le procédé de la figure 5,
- la figure 9b représente un second mode de réalisation de procédé de détermination de message, inclus dans le procédé de la figure 5,
- les figures 10 et 11 représentent des exemples d'attribution de clé selon l'invention.

30 On va décrire tout d'abord un mode de réalisation d'un dispositif d'insertion d'une information supplémentaire dans des données numériques en

référence à la **figure 1**. Ce dispositif est intégré dans un dispositif de traitement de données TD1, tel qu'un ordinateur, un appareil photographique numérique, un scanner, par exemple.

Une source 1 de données non codées comporte par exemple un
5 moyen de mémoire, telle que mémoire vive, disque dur, disquette, disque compact, pour mémoriser des données non codées, ce moyen de mémoire étant associé à un moyen de lecture approprié pour y lire les données. Un moyen pour enregistrer les données dans le moyen de mémoire peut également être prévu. La source 1 peut également être intégrée ou non à
10 l'appareil numérique.

On considérera plus particulièrement dans la suite que les données à coder sont une suite d'échantillons numériques représentant une image numérique IM. Une image originale IM peut être représentée par une série de pixels codés par exemple sur 8 bits ou octet. L'image IM noir et blanc peut ainsi
15 être décomposée dans le domaine spatial en un ensemble de coefficients sur 256 niveaux de gris, chaque valeur de coefficient représentant un pixel de l'image I.

La source de signal 1 est reliée à un circuit 2 de transformation réversible. Cette transformation est par exemple une transformation en
20 ondelettes de l'image ou encore une transformation en cosinus discrète, dite DCT, par blocs. La transformation réalise une décomposition de l'image numérique, et fournit un ensemble de coefficients. Si cette transformation est une transformation en ondelettes, ces coefficients portent une information spatio-temporelle. Si cette transformation est une transformation DCT par
25 blocs, les coefficients sont des coefficients spectraux. Cette transformation n'est pas essentielle pour l'invention qui peut être mise en œuvre sur les coefficients spatiaux d'une image.

Le circuit 2 est relié à un circuit de segmentation 3 qui segmente l'image en blocs, ou régions, de taille prédéterminée.

Un message M à insérer est mémorisé dans une mémoire 4. Chaque symbole du message M est inséré dans au moins un bloc formé dans l'image, par un circuit d'insertion 5.

Pour insérer un symbole, un générateur de clé 6 génère une clé en fonction d'une clé initiale K_{init} et de la longueur L du message à insérer. La clé générée est transmise à un générateur pseudo-aléatoire 7 qui génère une séquence pseudo-aléatoire.

La séquence pseudo-aléatoire est transmise à un circuit de modulation 8 qui reçoit également le symbole courant à insérer. Le circuit 8 module le symbole avec la séquence pseudo-aléatoire, ce qui produit une seconde séquence pseudo-aléatoire. Cette séquence est modifiée en amplitude de manière à assurer l'invisibilité du message inséré puis est fournie au circuit d'insertion 5, qui reçoit les blocs formés par le circuit de segmentation 3.

La modification d'amplitude de chaque coefficient de la séquence pseudo-aléatoire dépend d'un modèle psycho-visuel. Le modèle psycho-visuel dépend du mode de transformation utilisé.

Par exemple, le modèle psycho-visuel est spatial et attribue une modification maximale à chaque pixel de l'image. Chaque pixel de l'image est ainsi pondéré de façon à atteindre cette amplitude maximale. La modification maximale attribuée à chaque pixel augmente en fonction du degré d'activité du voisinage du pixel considéré. Le degré d'activité est une mesure de la texture. Ainsi, un pixel situé dans une zone fortement texturée de l'image sera plus modifié qu'un pixel situé dans une zone homogène.

Le circuit 5 additionne la séquence pseudo-aléatoire fournie par le circuit 8 au bloc courant, en fonction de l'association entre chaque symbole à insérer et au moins un bloc. Le circuit 5 fournit un bloc marqué.

Des moyens 9 utilisateurs de données codées sont reliés en sortie du circuit d'insertion 5.

Les moyens utilisateurs 9 comportent par exemple des moyens de mémorisation de données codées, et/ou des moyens de transmission des données codées.

Un dispositif de détection d'information supplémentaire, correspondant au dispositif d'insertion précédent, est représenté à la **figure 2**. Ce dispositif de détection d'une information supplémentaire dans des données est intégré dans un dispositif de traitement de données TD2, tel qu'un ordinateur, un appareil photographique numérique, un télécopieur, par exemple.

Le dispositif de détection comporte une source 20 de données dans lesquelles un message a été inséré.

Une sortie de la source 20 est reliée à un circuit 21 de transformation réversible, identique au circuit 2 du dispositif TD1 précédemment décrit.

Une sortie du circuit 21 est reliée à un circuit 22 de détection de la taille du message inséré. Le fonctionnement de ce circuit sera décrit dans la suite.

Une sortie du circuit 22 est reliée à un circuit 23 d'extraction du message inséré. Le fonctionnement de ce circuit sera décrit dans la suite.

Le message détecté est ensuite fourni à un circuit d'exploitation 24, qui comporte par exemple un écran permettant de lire le message.

Le fonctionnement du dispositif de détection sera détaillé dans la suite.

Comme représenté à la **figure 3**, un dispositif mettant en œuvre l'invention est par exemple un micro-ordinateur 10 connecté à différents périphériques, par exemple une caméra numérique 107 (ou un scanner, ou tout moyen d'acquisition ou de stockage d'image) reliée à une carte graphique et fournissant des informations à traiter selon l'invention.

Le dispositif 10 comporte une interface de communication 112 reliée à un réseau 113 apte à transmettre des données numériques à traiter ou inversement à transmettre des données traitées par le dispositif. Le dispositif 10 comporte également un moyen de stockage 108 tel que par exemple un

disque dur. Il comporte aussi un lecteur 109 de disque 110. Ce disque 110 peut être une disquette, un CD-ROM, ou un DVD-ROM, par exemple. Le disque 110 comme le disque 108 peuvent contenir des données traitées selon l'invention ainsi que le ou les programmes mettant en œuvre l'invention qui, une fois lu par
5 le dispositif 10, sera stocké dans le disque dur 108. Selon une variante, le programme permettant au dispositif de mettre en œuvre l'invention, pourra être stocké en mémoire morte 102 (appelée ROM sur le dessin). En seconde variante, le programme pourra être reçu pour être stocké de façon identique à celle décrite précédemment par l'intermédiaire du réseau de communication
10 113.

Le dispositif 10 est relié à un microphone 111. Les données à traiter selon l'invention seront dans ce cas du signal audio.

Ce même dispositif possède un écran 104 permettant de visualiser les données à traiter ou de servir d'interface avec l'utilisateur qui peut ainsi
15 paramétrer certains modes de traitement, à l'aide du clavier 114 ou de tout autre moyen (souris par exemple).

L'unité centrale 100 (appelée CPU sur le dessin) exécute les instructions relatives à la mise en œuvre de l'invention, instructions stockées dans la mémoire morte 102 ou dans les autres éléments de stockage. Lors de
20 la mise sous tension, les programmes de traitement stockés dans une mémoire non volatile, par exemple la ROM 102, sont transférés dans la mémoire vive RAM 103 qui contiendra alors le code exécutable de l'invention ainsi que des registres pour mémoriser les variables nécessaires à la mise en œuvre de l'invention.

25 De manière plus générale, un moyen de stockage d'information, lisible par un ordinateur ou par un microprocesseur, intégré ou non au dispositif, éventuellement amovible, mémorise un programme mettant en œuvre le procédé de codage, de transmission et respectivement de décodage.

Le bus de communication 101 permet la communication entre les
30 différents éléments inclus dans le micro-ordinateur 10 ou reliés à lui. La représentation du bus 101 n'est pas limitative et notamment l'unité centrale 100

est susceptible de communiquer des instructions à tout élément du micro-ordinateur 10 directement ou par l'intermédiaire d'un autre élément du micro-ordinateur 10.

5 Le fonctionnement des dispositifs d'insertion de marquage et de détection de marquage selon l'invention va maintenant être décrit au moyen d'algorithmes.

L'algorithme de la **figure 4** représente le fonctionnement général du dispositif d'insertion selon l'invention et comporte des étapes E1 à E15.

10 Cet algorithme peut être mémorisé en totalité ou en partie dans tout moyen de stockage d'information capable de coopérer avec le microprocesseur. Ce moyen de stockage est lisible par un ordinateur ou par un microprocesseur. Ce moyen de stockage est intégré ou non au dispositif, et peut être amovible. Par exemple, il peut comporter une bande magnétique, une disquette ou un CD-ROM (disque compact à mémoire figée).

15 L'étape E1 est une transformation réversible des données, par exemple une transformation en ondelettes de l'image, pour transformer les coefficients spatiaux en d'autres coefficients dont les propriétés statistiques permettent une meilleure extraction ultérieure du message inséré. Cette étape est facultative.

20 L'étape suivante E2 est la segmentation des données en régions, par exemple en blocs adjacents. Le nombre et/ou la taille des régions peuvent être prédéterminés ou réglables par un utilisateur.

Le message M à insérer comporte L symboles, où L est un entier. Chaque symbole M_i , avec l'entier i variant de 1 à L est associé à au moins une région à l'étape suivante E3. Une région donnée est associée à un unique
25 symbole à insérer. Pour l'association des symboles aux régions, ces dernières sont parcourues selon un ordre prédéterminé.

L'étape suivante E4 est une initialisation pour considérer le premier symbole M_1 à insérer, ainsi que la première région dans laquelle ce symbole
30 est à insérer.

A l'étape suivante E5 une variable C1 représentant le rang du symbole courant est mise à la valeur 1 et une variable C2 est mise à la valeur 0. La variable C2 représente le nombre de fois où le symbole courant a déjà été inséré. Les variables C1 et C2 sont liées à la longueur du message M.

5 L'étape suivante E6 est la génération d'une clé K en fonction d'une clé initiale K_{init} et des variables C1 et C2. Deux exemples sont détaillés dans la suite.

L'étape suivante E7 est la génération d'une séquence pseudo-aléatoire en fonction de la clé K précédemment générée.

10 L'étape suivante E8 est la modulation du symbole M_i par la séquence pseudo-aléatoire précédemment générée, ce qui a pour résultat une seconde séquence pseudo-aléatoire.

L'étape suivante E9 est une pondération psycho-visuelle de la seconde séquence pseudo-aléatoire pour assurer son invisibilité dans l'image.

15 La séquence pseudo-aléatoire ainsi modifiée est alors additionnée à la région courante à l'étape suivante E10.

L'étape suivante E11 est un test pour déterminer si la région courante est la dernière pour le symbole courant. Si la réponse est négative, cela signifie qu'il reste au moins une région dans laquelle le symbole courant doit être inséré. L'étape E11 est alors suivie de l'étape E12. A l'étape E12, est considérée la région suivante dans laquelle le symbole M_i est à insérer, et la variable C2 est incrémentée de une unité.

20

L'étape E12 est suivie de l'étape E6 précédemment décrite.

Lorsque la réponse est positive à l'étape E11, cela signifie que le
25 symbole courant a été inséré dans toutes les régions qui lui sont associées.

L'étape E11 est alors suivie de l'étape E13 qui est un test pour déterminer si le symbole courant est le dernier symbole à insérer. Si la réponse est négative, cela signifie qu'il reste au moins un symbole à insérer, et cette étape est suivie de l'étape E14 à laquelle le paramètre i est incrémenté de une
30 unité pour considérer le symbole suivant M_{i+1} et la première région qui lui est associée.

L'étape E14 est suivie de l'étape E5 précédemment décrite.

Lorsque la réponse est positive à l'étape E13, cela signifie que tous les symboles ont été insérés dans l'image. L'étape E13 est alors suivie de l'étape E15 à laquelle une transformation inverse est effectuée sur les données
5 traitées. La transformation inverse correspond à la transformation de l'étape E1 et a pour but de fournir les coefficients spatiaux d'une image dans laquelle le message M a été inséré. Bien entendu, si la transformation de l'étape E1 n'a pas été effectuée, la transformation inverse de l'étape E15 ne l'est pas non plus.

10 L'algorithme de la **figure 5** représente le fonctionnement général du dispositif d'extraction selon l'invention et comporte des étapes E20 à E22.

Cet algorithme peut être mémorisé en totalité ou en partie dans tout moyen de stockage d'information capable de coopérer avec le microprocesseur. Ce moyen de stockage est lisible par un ordinateur ou par un
15 microprocesseur. Ce moyen de stockage est intégré ou non au dispositif, et peut être amovible. Par exemple, il peut comporter une bande magnétique, une disquette ou un CD-ROM (disque compact à mémoire figée).

L'étape E20 est une transformation réversible des données dans lesquelles un message a été inséré, pour transformer les coefficients spatiaux.
20 Cette étape n'est effectuée que si l'insertion a été préalablement effectuée dans des données transformées. La transformation est ici identique à celle utilisée lors de l'insertion du message.

L'étape suivante E21 est la détection de la taille du message inséré. Cette étape sera détaillée dans la suite.

25 L'étape suivante E22 est l'extraction du message proprement dit. Cette étape est détaillée dans la suite.

Un premier mode de réalisation de détection de la taille du message inséré est maintenant détaillée en référence à la **figure 6a**, et comporte des étapes E210 à E223.

30 Dans ce mode de réalisation, le message inséré est composé de symboles qui sont des bits.

L'étape E210 est une segmentation des données dans lesquelles le message a été préalablement inséré. Cette étape est identique à l'étape E2 (insertion).

5 L'étape suivante E211 est une hypothèse sur la valeur de la longueur L du message que l'on cherche à extraire. La longueur L peut prendre des valeurs entre un et une valeur maximale $L_{\max}$, qui est fixée a priori ou qui dépend de la taille de l'image. Toutes ces valeurs seront successivement considérées.

10 L'étape suivante E212 permet de réaliser un bouclage sur tous les symboles M_1 à M_L du message. Pour chaque boucle, un symbole courant M_i est considéré.

L'étape suivante E213 est la sélection des régions correspondant au symbole courant M_i . Le mécanisme d'attribution d'une région à un symbole est identique à celui utilisé à l'étape E3 (insertion).

15 L'étape suivante E214 est la détermination des variables $C1$ et $C2$ pour le symbole courant et pour chacune des régions sélectionnées à l'étape précédente. Comme précédemment, la variable $C1$ représente le rang du symbole courant, et a donc une valeur unique pour toutes les régions sélectionnées. La variable $C2$ représente, pour chaque région, le nombre de
20 fois où le symbole courant a déjà été inséré.

L'étape suivante E215 est une génération de clés, pour le symbole courant et pour chaque région sélectionnée à l'étape E213. La génération de clé est identique à celle de l'étape E6 (insertion), et utilise notamment la clé initiale K_{init} .

25 L'étape suivante E216 est la génération de séquences pseudo-aléatoire en fonction de chacune des clés K précédemment générées.

L'étape suivante E217 est le calcul de la corrélation entre les séquences pseudo-aléatoires générées à l'étape précédente et les régions sélectionnées à l'étape E213.

30 L'étape suivante E218 calcule la valeur absolue de la corrélation calculée pour le bit courant et pour toutes les régions qui lui sont associées.

Si l'hypothèse (étape E211) sur la taille du message est fausse, alors la valeur calculée à l'étape E218 demeure faible. En revanche, si l'hypothèse sur la taille du message est vraie, alors la valeur calculée à l'étape E218 est élevée.

5 L'étape suivante E219 est la sommation de la valeur absolue de la corrélation calculée pour le symbole courant avec les valeurs absolues des corrélations précédemment calculées pour les autres symboles, pour la longueur de message considérée.

Pour une longueur donnée de message, les étapes E212 à E219
10 sont répétées pour tous les symboles du message.

Les étapes E211 à E219 sont répétées pour toutes les longueurs de message à considérer.

Chaque somme calculée à l'étape E219 (c'est-à-dire pour chaque longueur de message à tester) est mémorisée dans un tableau à l'étape
15 suivante E220. Lorsque toutes les longueurs de message ont été traitées, le tableau est complètement rempli et il faut déterminer le maximum local parmi les valeurs de ce tableau.

Pour cela, un filtrage passe-haut est effectué sur le tableau à l'étape E221. Par exemple, le filtrage passe-haut calcule l'écart entre la valeur de la
20 « case » courante du tableau et la moyenne de ses voisins. Le filtre passe-haut correspondant est le filtre $(-0.5, 1, -0.5)$.

Le résultat du filtrage est écrit dans un second tableau à l'étape E222.

L'étape suivante E223 est la détection de la valeur maximale
25 contenue dans le second tableau. Cette valeur maximale correspond à une longueur, qui est la taille du message inséré. La technique qui est utilisée ici est la détection de maximum local, qui permet d'extraire efficacement la taille des messages de faible longueur lorsque la gamme des tailles de message possibles est très large, par exemple de un jusqu'à plusieurs milliers de bits.

30 Un second mode de réalisation de détection de la taille du message inséré est maintenant détaillé en référence à la **figure 6b**, et comporte des

étapes E210 à E223 et E230. Les étapes E210 à E223 sont analogues à celles précédemment décrites (figure 6a).

Dans ce mode de réalisation, le message inséré est composé de symboles qui ne sont pas des bits, mais appartiennent à un dictionnaire de S
5 signes, avec S un entier supérieur à deux.

L'étape E230 est ajoutée entre les étapes E216 et E217 et les étapes E217 et E218 sont modifiées. Pour l'ensemble des régions correspondant à un même symbole inconnu, il faut déterminer quel est le symbole le plus probable. Pour cela, les séquences pseudo-aléatoires
10 générées à l'étape E216 modulent chacun des S symboles possibles à l'étape E230.

La corrélation entre les S séquences résultantes et les régions sélectionnées est ensuite calculée à l'étape E217.

A l'étape E218, la valeur maximale parmi les S valeurs de corrélation
15 est conservée. Cette valeur est ajoutée aux autres valeurs de corrélation correspondant aux autres symboles à l'étape E219.

La **figure 7a** représente une première variante de réalisation de détection de la taille du message inséré, dans le cas où le message est composé de bits. Ce mode de réalisation est représenté sous la forme d'un
20 algorithme qui comporte des étapes E310 à E324.

Les étapes E310 à E323 sont respectivement analogues aux étapes E210 à E223.

L'étape supplémentaire E324 permet d'accélérer les calculs lors de l'extraction de la longueur du message. Un facteur d'accélération F permet de
25 sélectionner F fois moins de pixels à l'étape E313 que lorsque le facteur vaut un.

En conséquence, l'étape E316 de génération de séquences pseudo-aléatoires génère F fois moins de valeurs. L'étape E317 de corrélation traite F fois moins de valeurs. Il est à noter cependant que les pixels sont sélectionnés
30 d'une manière spécifique, qui sera détaillée en référence à la figure 8, pour que la phase de détection de la longueur du message soit optimale.

La **figure 7b** représente une seconde variante de réalisation de détection de la taille du message inséré, dans le cas où le message est composé de symboles qui ne sont pas des bits. Ce mode de réalisation est représenté sous la forme d'un algorithme qui comporte des étapes E310 à
5 E330.

Les étapes E310 à E324 sont respectivement analogues aux étapes précédemment décrites.

L'étape E330 est ajoutée entre les étapes E316 et E317 et les étapes E317 et E318 sont modifiées. Pour l'ensemble des régions
10 correspondant à un même symbole inconnu, il faut déterminer quel est le symbole le plus probable. Pour cela, les séquences pseudo-aléatoires générées à l'étape E316 modulent chacun des S symboles possibles à l'étape E330.

La corrélation entre les S séquences résultantes et les régions
15 sélectionnées est ensuite calculée à l'étape E317.

A l'étape E318, la valeur maximale parmi les S valeurs de corrélation est conservée. Cette valeur est ajoutée aux autres valeurs de corrélation correspondant aux autres symboles à l'étape E319.

La **figure 8** illustre l'accélération des calculs, sous la forme d'un
20 algorithme comportant des étapes E30 à E39.

Le principe de l'accélération est de choisir un nombre de coefficients dans les données dans lesquelles un message a été inséré, puis de considérer le maximum de coefficients relativement au minimum de symboles recherchés. En d'autres termes, on choisit des coefficients correspondant à un même
25 symbole inséré, tant que cela est possible, puis on boucle sur les symboles, tant que le nombre de coefficients choisi au départ n'est pas atteint.

Ainsi, la corrélation qui est effectuée ultérieurement est plus efficace et la détection des symboles est améliorée.

A partir des données dans lesquelles un message a été inséré, le
30 nombre K de coefficients, ici de pixels, de ces données est déterminé à l'étape E30.

Le facteur d'accélération F est déterminé à l'étape E31, par exemple en le lisant dans une zone mémoire prédéterminée, et le nombre C de pixels à utiliser est déterminé à l'étape E32 par la formule : $C = \text{Ent}[K/F]$, où $\text{Ent}[\dots]$ dénote la partie entière. Le nombre C sera ensuite décrémenté à chaque fois
5 que des pixels sont sélectionnés.

A partir des données segmentées (E310) une boucle est parcourue pour chaque symbole du message de longueur L .

Cette boucle commence à l'étape E33 qui est une initialisation pour considérer le premier symbole M_i .

10 L'étape suivante E34 est la sélection des régions correspondant au symbole courant M_i . L'ensemble de ces régions comporte L_i coefficients.

L'étape suivante E35 est un test pour comparer la taille L_i et le nombre C de pixels à utiliser.

Si la taille L_i est inférieure au nombre C , alors cette étape est suivie
15 de l'étape E36 à laquelle tous les pixels des régions considérées sont sélectionnés et le nombre C vaut alors $C - L_i$.

Si la taille L_i est supérieure au nombre C , alors cette étape est suivie de l'étape E37 à laquelle les C premiers pixels sont sélectionnés dans les régions considérées et le nombre C est mis à la valeur zéro.

20 Les étapes E36 et E37 sont suivies de l'étape E38 qui est un test pour déterminer si le nombre C est nul. Si la réponse est positive, alors la sélection des pixels est terminée. Si la réponse est négative, alors l'étape E38 est suivie de l'étape E39 pour considérer un symbole suivant dans le message. L'étape E39 est suivie de l'étape E34 précédemment décrite.

25 La **figure 9a** illustre un premier mode de réalisation d'extraction du message proprement dit (étape E22), dans l'algorithme de la figure 5. Ce mode de réalisation correspond au cas où les symboles du message sont des bits.

L'extraction est représentée sous la forme d'un algorithme comportant des étapes E40 à E49.

30 L'étape E40 est une lecture en mémoire de la longueur L du message inséré. Cette longueur a été précédemment déterminée (étape E21).

L'étape suivante E41 est une initialisation d'une boucle pour considérer successivement tous les symboles du message.

Pour chaque symbole, l'étape E42 est une sélection des régions correspondant au symbole courant M_i .

5 L'étape suivante E43 est la détermination des valeurs des variables C1 et C2 respectivement pour chacune des régions précédemment sélectionnées. Les variables C1 et C2 sont déterminées comme précédemment exposé.

10 L'étape suivante E44 est le calcul de la clé correspondant à chacun des couples de valeurs (C1, C2).

L'étape suivante E45 est le calcul de la séquence pseudo-aléatoire correspondant à chacune des clés précédemment calculées.

15 L'étape suivante E46 est le calcul, pour chaque séquence pseudo-aléatoire précédemment calculée, de la corrélation entre la séquence pseudo-aléatoire et la région qui lui correspond.

L'étape suivante E47 est l'addition de toutes les valeurs de corrélation correspondant au symbole courant. Le signe de la somme est alors déterminé.

20 L'étape suivante E48 est une décision sur la valeur du symbole recherché. Si le signe précédemment déterminé est positif, alors le symbole est le bit un, et sinon, le symbole est le bit zéro.

La valeur du bit est ensuite mémorisée à l'étape suivante E49.

Lorsque tous les bits ont été déterminés, le message inséré est entièrement déterminé.

25 La **figure 9b** illustre un second mode de réalisation d'extraction du message proprement dit (étape E22), dans l'algorithme de la figure 5. Ce mode de réalisation correspond au cas où les symboles du message ne sont pas des bits mais appartiennent à un dictionnaire de S signes, avec S un entier supérieur à deux.

30 Ce mode de réalisation est représenté sous la forme d'un algorithme qui comporte des étapes E40 à E50.

Les étapes E40 à E47 et E49 sont respectivement analogues aux étapes portant les mêmes références et précédemment décrites. L'étape E48 est supprimée.

L'étape E50 est ajoutée entre les étapes E45 et E46 et les étapes E46 et E47 sont modifiées. Pour extraire un symbole inconnu, on considère l'ensemble des régions correspondant à ce symbole inconnu. Il faut alors déterminer quel est le symbole le plus probable parmi les S symboles possibles. Pour cela, les séquences pseudo-aléatoires générées à l'étape E45 modulent chacun des S symboles possibles à l'étape E50.

La corrélation entre les S séquences résultantes et les régions sélectionnées est ensuite calculée à l'étape E46.

A l'étape E47, la valeur maximale parmi les S valeurs de corrélation indique la valeur du symbole courant. Cette valeur est mémorisée à l'étape suivante E49.

La **figure 10** est un premier exemple d'attribution de clé selon l'invention. Ces clés servent de germes pour la génération des séquences pseudo-aléatoires. L'image a été découpée en huit régions rectangulaires R_0 à R_7 .

Le message à insérer comporte trois symboles S_0 , S_1 et S_2 . A chaque région est attribué un symbole de message.

Une clé est attribuée à chaque région. Cette clé dépend de l'indice du symbole enfoui dans la région considérée, et du nombre de fois où ce symbole a déjà été inséré précédemment.

Pour une région quelconque, on a la relation :

$$K = K_{\text{init}} + N.C2 + C1$$

Où N est un entier valant $L_{\text{max}} + 1$.

On rappelle que L_{max} est la valeur maximale que peut prendre la longueur du message.

La **figure 11** est un second exemple d'attribution de clé selon l'invention. Ces clés servent de germes pour la génération des séquences

pseudo-aléatoires. L'image a été découpée en huit régions rectangulaires R_0 à R_7 .

Le message à insérer comporte trois symboles S_0 , S_1 et S_2 . A chaque région est attribué un symbole de message.

- 5 Une clé est attribuée à chaque région. Cette clé dépend de l'indice du symbole enfoui dans la région considérée, et du nombre de fois où ce symbole a déjà été inséré précédemment.

Pour une région quelconque, on a la relation :

$$K = K_{\text{init}} + C2.$$

- 10 Bien entendu, la présente invention n'est nullement limitée aux modes de réalisation décrits et représentés, mais englobe, bien au contraire, toute variante à la portée de l'homme du métier.

REVENDEICATIONS

1. Procédé d'insertion d'un message dans des données numériques représentatives de grandeurs physiques, le message comportant des symboles ordonnés, comportant les étapes de :
- segmentation (E2) des données en régions,
 - association (E3) d'au moins une région à chaque symbole à insérer,
- caractérisé en ce qu'il comporte, pour chaque région dans laquelle un symbole considéré est à insérer, les étapes de :
- détermination (E7) d'une fonction pseudo-aléatoire, à partir d'une clé qui dépend :
 - d'une clé initiale, et
 - de la longueur du message,
 - modulation (E8) du symbole considéré par la fonction pseudo-aléatoire précédemment déterminée pour fournir une séquence pseudo-aléatoire,
 - addition (E10) de la séquence pseudo-aléatoire à la région considérée.
2. Procédé d'insertion selon la revendication 1, caractérisé en ce que la dépendance de la clé vis à vis de la longueur du message est assurée par la dépendance de la clé vis à vis :
- du nombre de fois où le symbole à insérer a déjà été inséré dans d'autres régions, et
 - du rang du symbole parmi les symboles ordonnés.
3. Procédé d'insertion selon la revendication 1 ou 2, caractérisé en ce qu'il comporte une étape préalable (E1) de transformation des données numériques par une transformation réversible.

4. Procédé d'extraction d'un message dans des données numériques représentatives de grandeurs physiques, le message comportant des symboles ordonnés, comportant les étapes de :

- segmentation (E210) des données en régions,
- 5 - extraction (E21) de la longueur du message inséré,
- extraction (E22) du message inséré.

5. Procédé d'extraction selon la revendication 4, caractérisé en ce que l'extraction de la longueur du message inséré comporte les étapes de :

- sélection (E211) d'un ensemble de valeurs de longueur, et
- 10 - calcul ((E217) d'une corrélation entre le message et les données numériques, pour chacune de ces valeurs,
- détermination (E223) d'un maximum local parmi les valeurs de corrélation.

6. Procédé d'extraction selon la revendication 4 ou 5, caractérisé en
15 ce que l'extraction de la longueur du message inséré est effectuée en traitant F fois moins de coefficients que n'en comportent les données numériques.

7. Procédé d'extraction selon la revendication 6, caractérisé en ce qu'il comporte les étapes de :

- détermination (E22) du nombre total de coefficients (C) à
20 considérer,
- sélection (E26, E27) d'un nombre maximum de coefficients correspondant à un même symbole inséré, puis, si le nombre total de coefficients à considérer n'est pas atteint,
- réitération (E29) de l'étape de sélection, pour un autre symbole.

25 8. Dispositif d'insertion d'un message dans des données numériques représentatives de grandeurs physiques, le message comportant des symboles ordonnés, comportant :

- des moyens (3) de segmentation des données en régions,
- des moyens (5) d'association d'au moins une région à chaque
30 symbole à insérer,

caractérisé en ce qu'il comporte, pour chaque région dans laquelle un symbole est à insérer :

- des moyens (7) de détermination d'une fonction pseudo-aléatoire, pour chaque région dans laquelle un symbole considéré est à insérer, à partir
- 5 d'une clé qui dépend :

- d'une clé initiale, et
- de la longueur du message,

- des moyens (8) de modulation du symbole considéré par la fonction pseudo-aléatoire précédemment déterminée pour fournir une
- 10 séquence pseudo-aléatoire,

- des moyens (5) d'addition de la séquence pseudo-aléatoire à la région considérée.

9. Dispositif d'insertion selon la revendication 8, caractérisé en ce que les moyens (7) de détermination d'une fonction pseudo-aléatoire sont
- 15 adaptés de sorte que la dépendance de la clé vis à vis de la longueur du message est assurée par la dépendance de la clé vis à vis :

- du nombre de fois où le symbole à insérer a déjà été inséré dans d'autres régions, et
- du rang du symbole parmi les symboles ordonnés.

- 20 10. Dispositif d'insertion selon la revendication 8 ou 9, caractérisé en ce qu'il comporte des moyens (2) de transformation préalable des données numériques par une transformation réversible.

11. Dispositif d'extraction d'un message dans des données numériques représentatives de grandeurs physiques, le message comportant
- 25 des symboles ordonnés, comportant :

- des moyens de segmentation des données en régions,
- des moyens (22) d'extraction de la longueur du message inséré,
- des moyens (23) d'extraction du message inséré.

12. Dispositif d'extraction selon la revendication 11, caractérisé en ce que les moyens (22) d'extraction de la longueur du message inséré comportent :

- des moyens de sélection d'un ensemble de valeurs de longueur, et
- 5 - des moyens de calcul d'une corrélation entre le message et les données numériques, pour chacune de ces valeurs,
- des moyens de détermination d'un maximum local parmi les valeurs de corrélation.

13. Dispositif d'extraction selon la revendication 11 ou 12, 10 caractérisé en ce que les moyens d'extraction de la longueur du message inséré sont adaptés à effectuer l'extraction en traitant F fois moins de coefficients que n'en comportent les données numériques.

14. Dispositif d'extraction selon la revendication 13, caractérisé en ce qu'il comporte :

- 15 - des moyens de détermination du nombre total de coefficients (C) à considérer,
- des moyens de sélection d'un nombre maximum de coefficients correspondant à un même symbole inséré, puis, si le nombre total de coefficients à considérer n'est pas atteint,
- 20 - des moyens de réitération de l'étape de sélection, pour un autre symbole.

15. Dispositif d'insertion selon l'une quelconque des revendications 8 à 10, caractérisé en ce que les moyens de segmentation, association, détermination, modulation et addition sont incorporés dans :

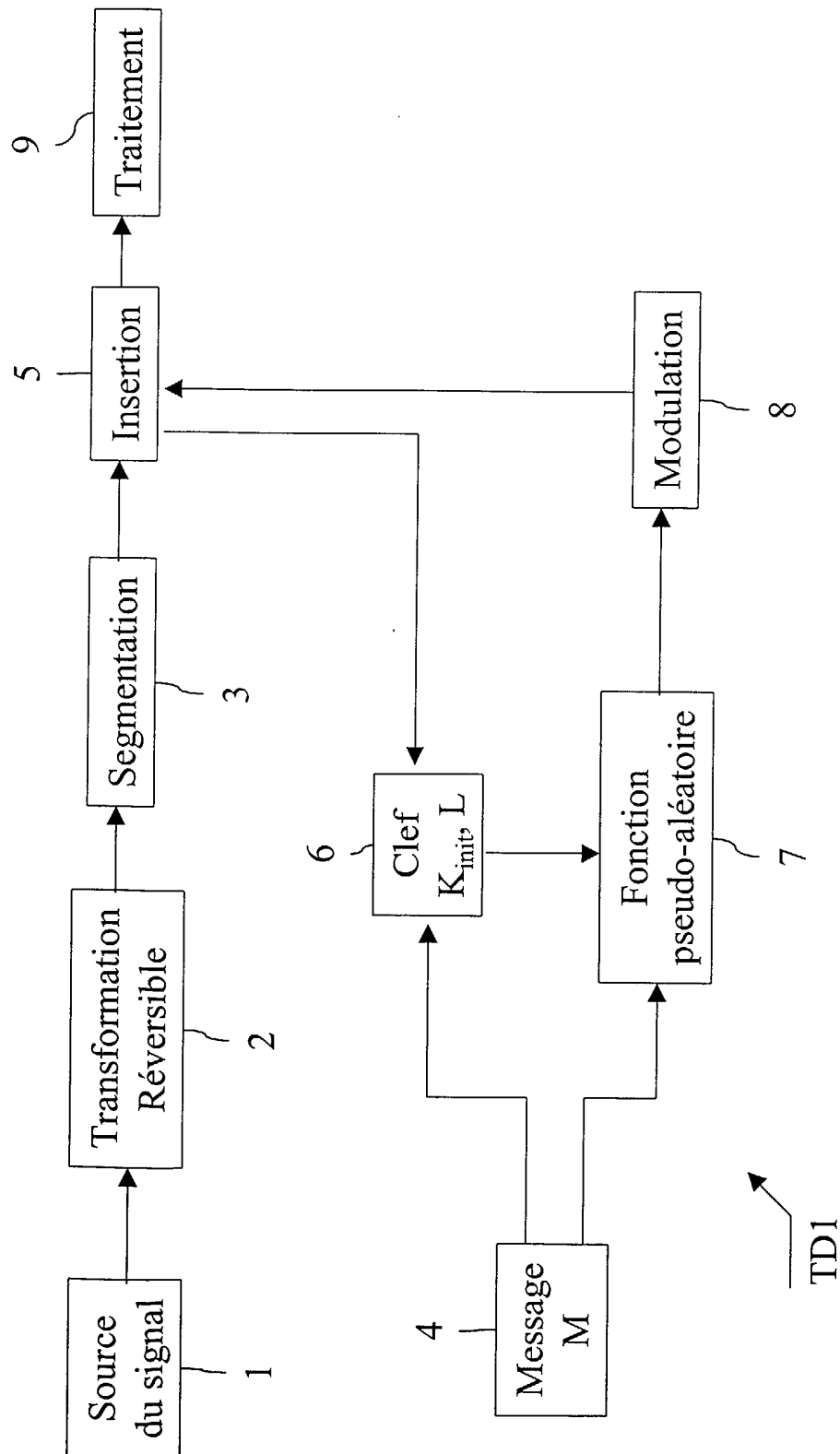
- 25 - un microprocesseur (100),
- une mémoire morte (102) comportant un programme pour traiter les données, et
- une mémoire vive (103) comportant des registres adaptés à enregistrer des variables modifiées au cours de l'exécution dudit programme.

16. Dispositif d'extraction selon l'une quelconque des revendications 11 à 14, caractérisé en ce que les moyens de segmentation et extraction sont incorporés dans :

- un microprocesseur (100),
- 5 - une mémoire morte (102) comportant un programme pour traiter les données, et
- une mémoire vive (103) comportant des registres adaptés à enregistrer des variables modifiées au cours de l'exécution dudit programme.

17. Appareil de traitement (10) d'une image numérique, caractérisé
10 en ce qu'il comporte des moyens adaptés à mettre en œuvre le procédé selon l'une quelconque des revendications 1 à 7.

18. Appareil de traitement (10) d'une image numérique, caractérisé en ce qu'il comporte le dispositif selon l'une quelconque des revendications 8 à 16.

*Figure 1*

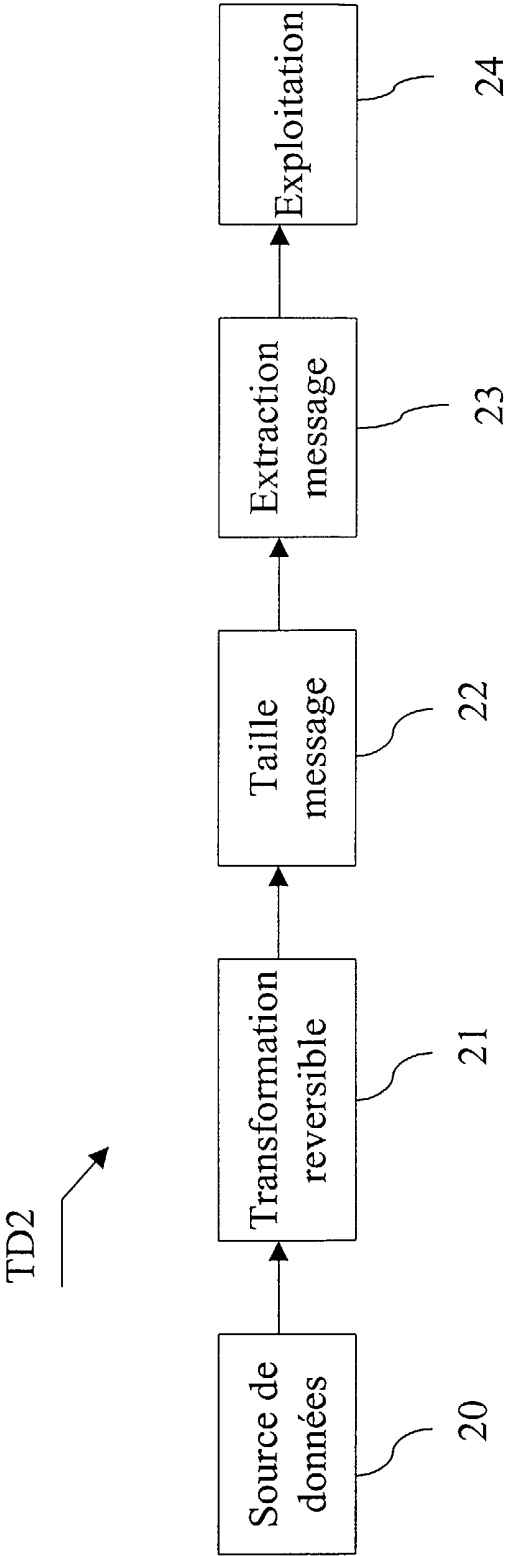
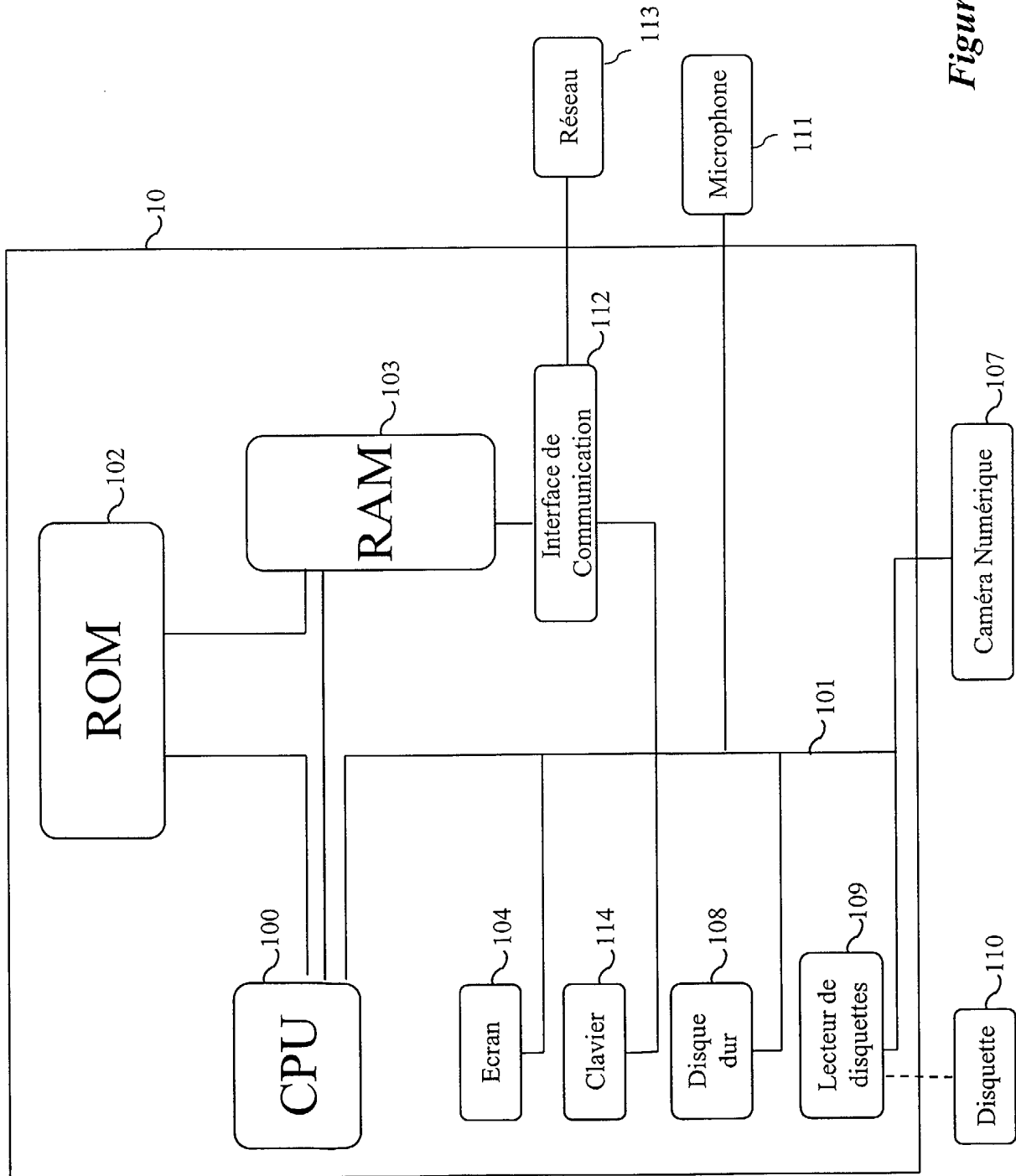
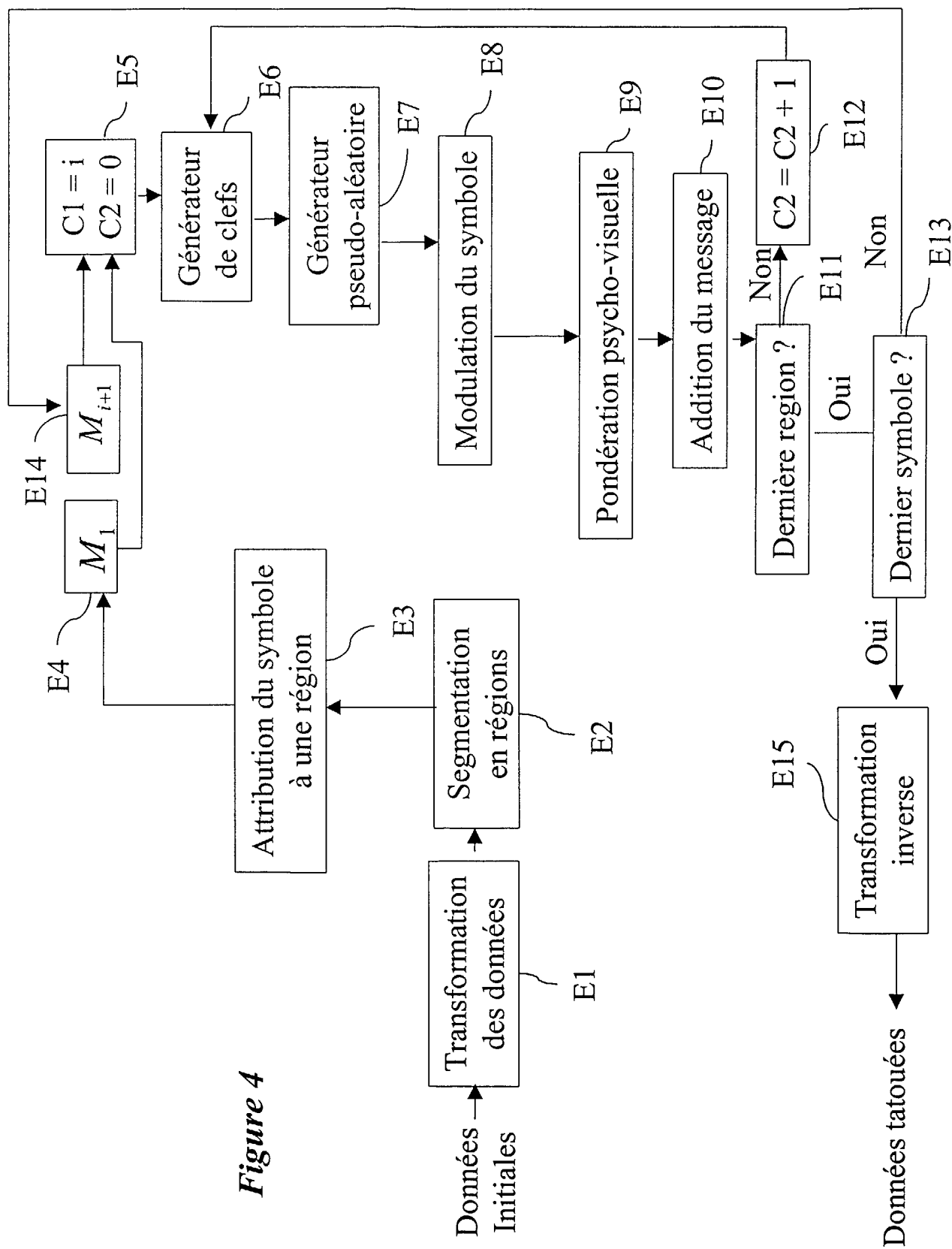


Figure 2

*Figure 3*



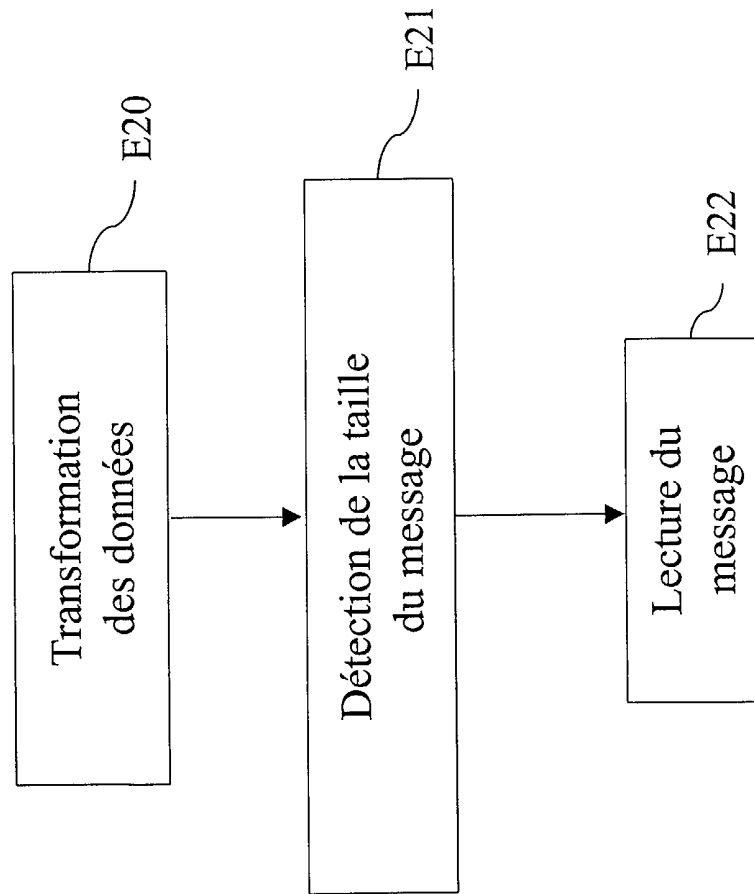


Figure 5

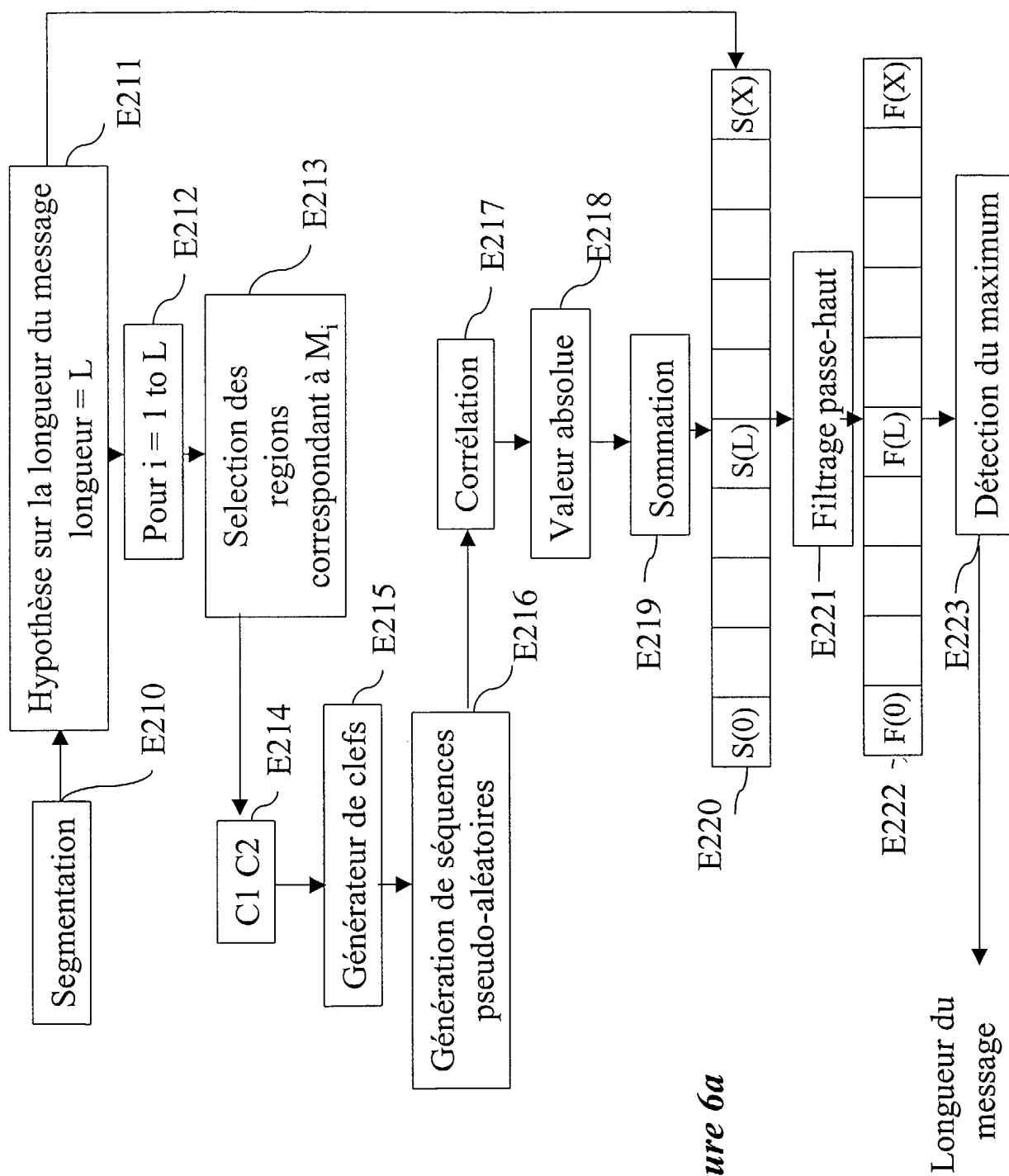
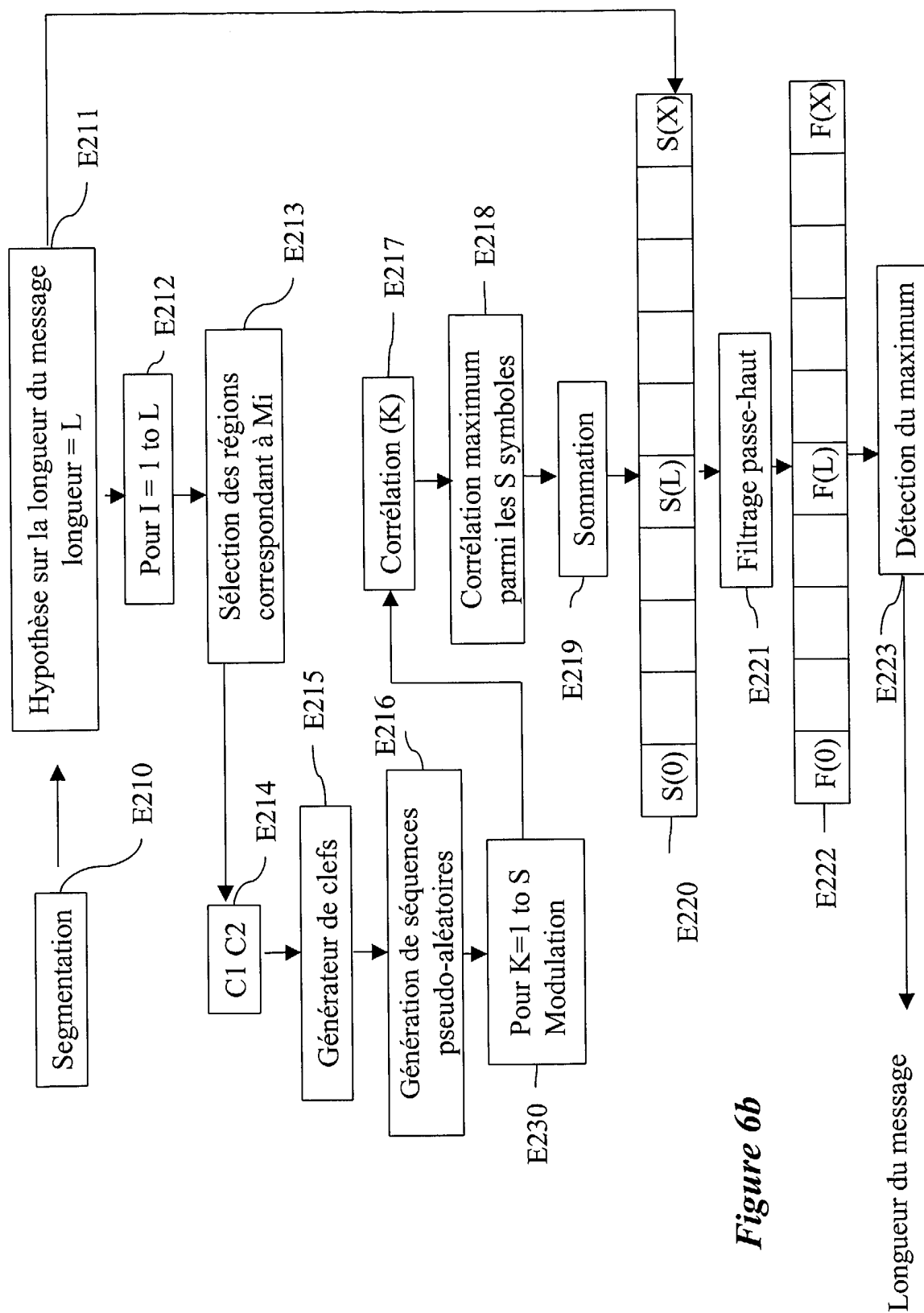


Figure 6a

**Figure 6b**

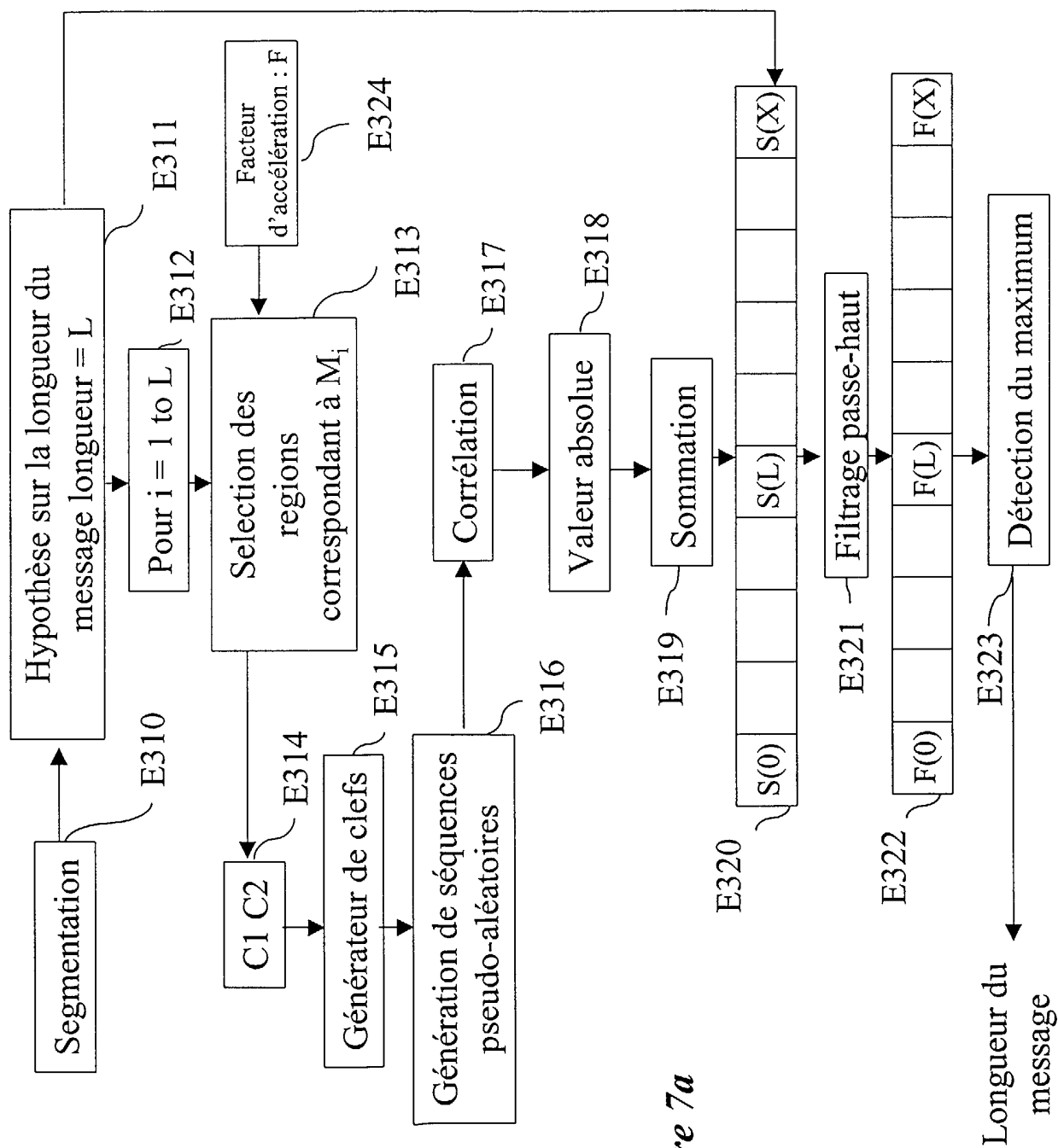


Figure 7a

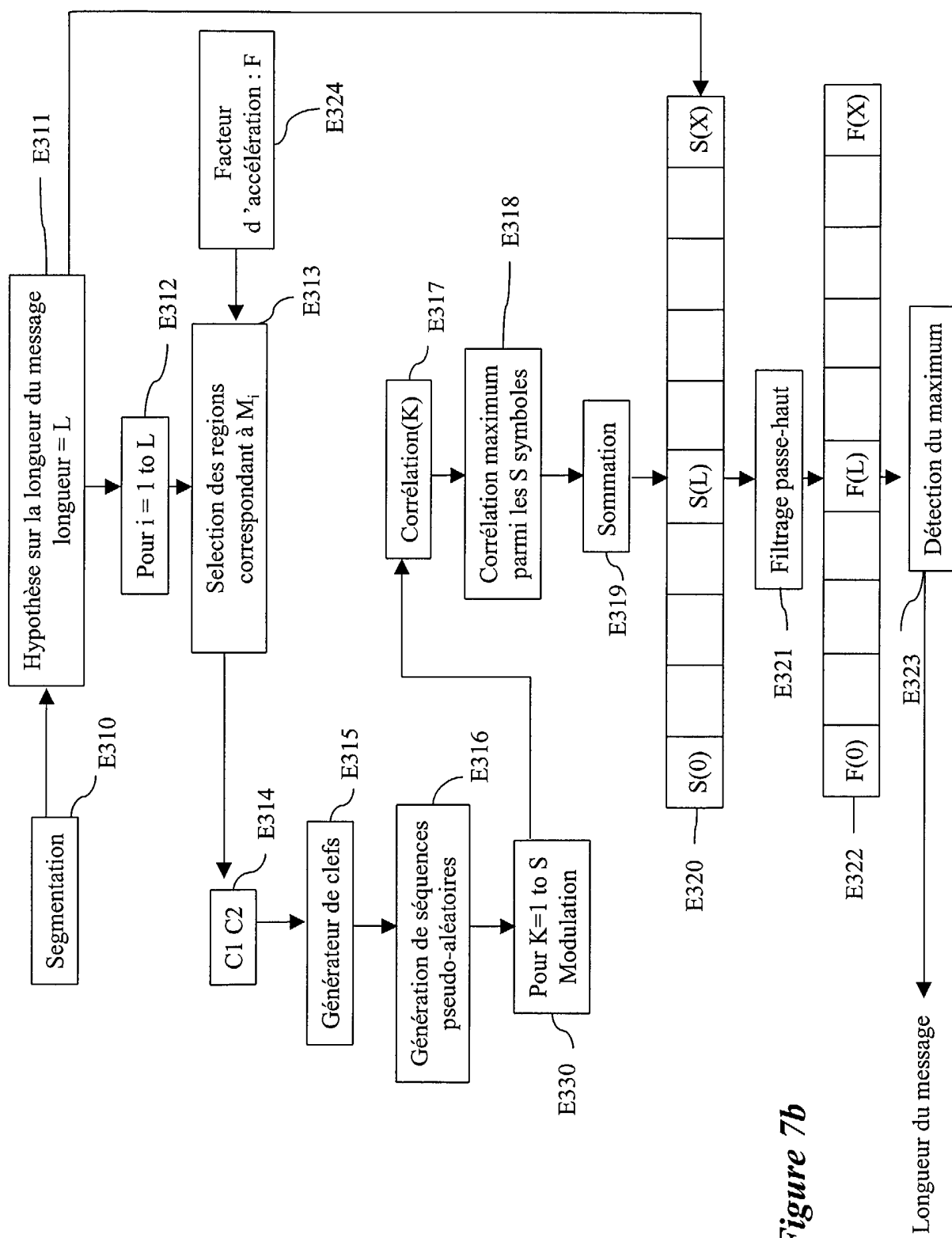


Figure 7b

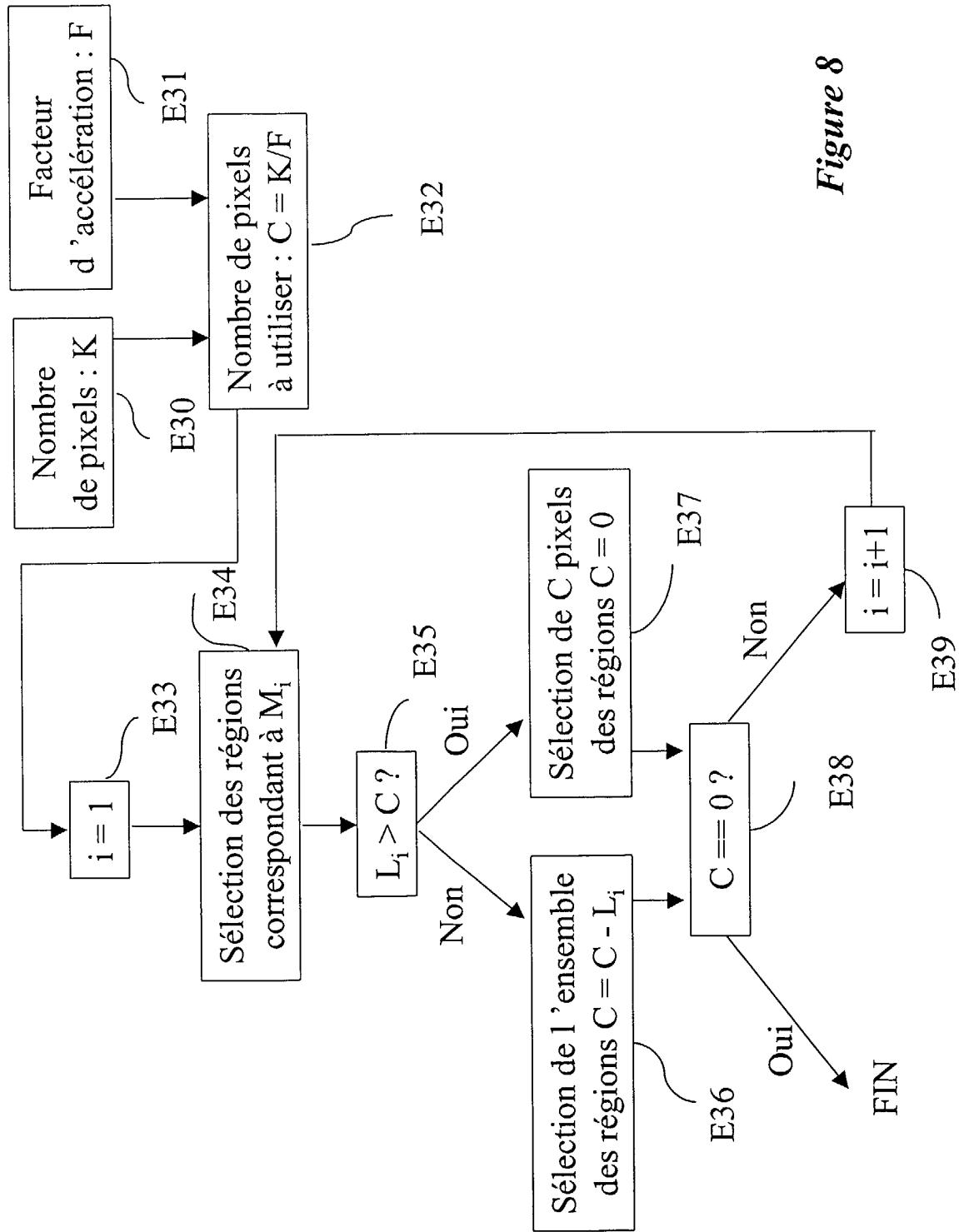
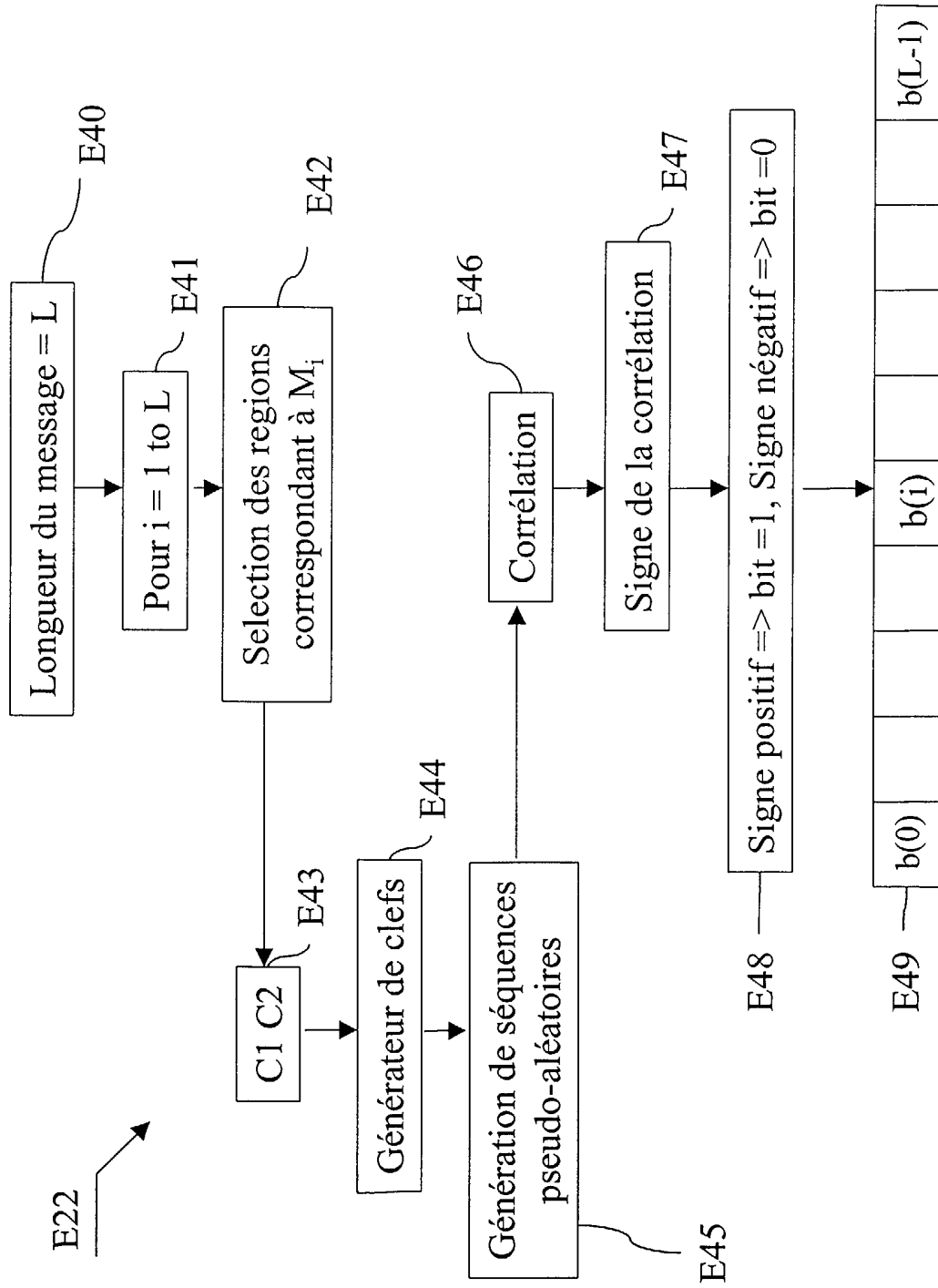


Figure 8

**Figure 9a**

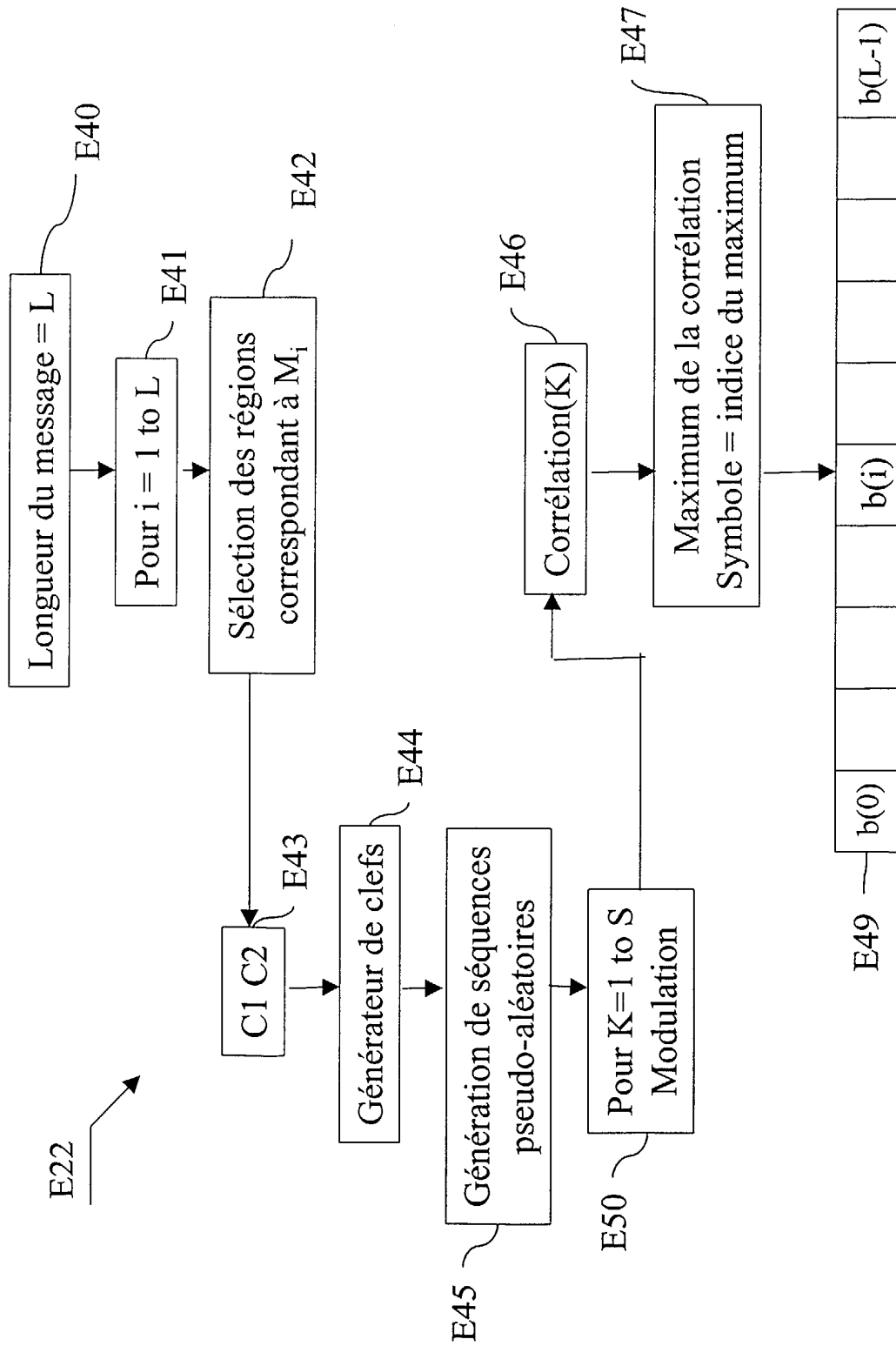


Figure 9b

S_0 R_0 K_{init}	S_1 R_1 $K_{init} + 1$	S_2 R_2 $K_{init} + 2$	S_0 R_3 $K_1 = K_{init} + N$
S_1 R_4 $K_1 + 1$ $=$ $K_{init} + N + 1$	S_2 R_5 $K_1 + 2$ $=$ $K_{init} + N + 2$	S_0 R_6 $K_2 = K_1 + N$ $=$ $K_{init} + 2N$	S_1 R_7 $K_2 + 1$ $=$ $K_{init} + 2N + 1$

Figure 10

S_0 R_0 K_{init}	S_1 R_1 K_{init}	S_2 R_2 K_{init}	S_0 R_3 $K_{init} + 1$
S_1 R_4 $K_{init} + 1$	S_2 R_5 $K_{init} + 1$	S_0 R_6 $K_{init} + 2$	S_1 R_7 $K_{init} + 2$

Figure 11



RAPPORT DE RECHERCHE PRÉLIMINAIRE

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

2812502

N° d'enregistrement
national

FA 592476

FR 0009727

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		
Y,D	EP 0 840 513 A (NIPPON ELECTRIC CO) 6 mai 1998 (1998-05-06) * abrégé * * colonne 1, ligne 3 - colonne 4, ligne 20 * * figures 1,3 *	1-18	H04N7/08 H04N1/32 H04N7/30
Y	--- SU, PO-CHYI; WANG, HOUNG-JYH M.; KUO, C.-C. J: "Digital Watermarking on EBCOT compressed images" PROC. SPIE, 'en ligne! vol. 3808, juillet 1999 (1999-07), pages 313-324, XP002165697 Denver Extrait de l'Internet: <URL:http:www.spie.org> 'extrait le 2001-04-20! * page 313 - page 324 *	1-18	
A	EP 0 891 071 A (MATSUSHITA ELECTRIC IND CO LTD) 13 janvier 1999 (1999-01-13) * abrégé *	1-18	DOMAINES TECHNIQUES RECHERCHÉS (Int.CL.7)
A	--- PODILCHUK C I ET AL: "DIGITAL IMAGE WATERMARKING USING VISUAL MODELS" PROCEEDINGS OF THE SPIE, 10 février 1997 (1997-02-10), XP000199957 * page 100 - page 105 *	1-18	H04N
A	--- MENG AND CHANG: "EMBEDDING VISIBLE VIDEO WATERMARKS IN THE COMPRESSED DOMAIN" INTERNATIONAL CONFERENCE ON IMAGE PROCESSING AND ITS APPLICATIONS,GB,LONDON, vol. 1, no. 1, 4 octobre 1998 (1998-10-04), pages 474-477, XP002124674 * page 474 - page 477 *	1-18	
Date d'achèvement de la recherche		Examineur	
9 mai 2001		Stoffers, C	
CATÉGORIE DES DOCUMENTS CITÉS 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 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			