

19) RÉPUBLIQUE FRANÇAISE  
**INSTITUT NATIONAL  
 DE LA PROPRIÉTÉ INDUSTRIELLE**  
 PARIS

11) N° de publication : **2 880 762**  
 (à n'utiliser que pour les  
 commandes de reproduction)

21) N° d'enregistrement national : **05 00180**

51) Int Cl<sup>8</sup> : H 04 Q 7/32 (2006.01), H 04 M 3/56, H 04 N 7/15

12)

**DEMANDE DE BREVET D'INVENTION**

**A1**

22) Date de dépôt : 07.01.05.

30) Priorité :

43) Date de mise à la disposition du public de la demande : 14.07.06 Bulletin 06/28.

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) : FRANCE TELECOM Société anonyme — FR.

72) Inventeur(s) : MARTIN ALEXIS, DAMLAMIAN JEAN JACQUES et AIRIAU ROLAND.

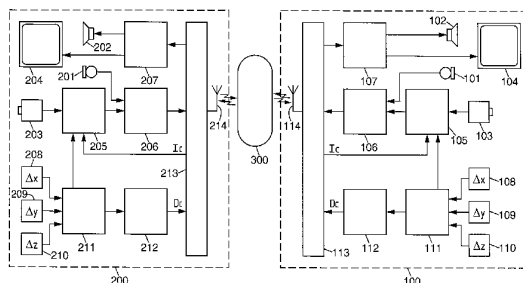
73) Titulaire(s) :

74) Mandataire(s) : CABINET PLASSERAUD.

54) **TERMINAL DE VISIOPHONIE AVEC REGLAGES INTUITIFS.**

57) L'invention permet un réglage intuitif du cadrage d'un terminal 100 par un correspondant distant utilisant un terminal portable 200 pendant une communication visiophonique. Le terminal 100 comporte une caméra 103, un moyen de cadrage 105, un moyen d'encodage vidéo 106 et un moyen de multiplexage et de communication 113. Le terminal 200 comporte des capteurs de mouvements 208-210, des moyens de mise en forme 211-212 et un moyen de multiplexage et de communication 213. En réponse à un mouvement du terminal 200, celui-ci élabore des données de cadrage  $D_c$  à partir d'informations de mouvement  $\Delta x$ ,  $\Delta y$  et  $\Delta z$  provenant des capteurs de mouvement 208-210 et envoie lesdites données de cadrage  $D_c$  au terminal

100. En réponse à des informations de cadrage  $I_c$  correspondant aux données de cadrage  $D_c$ , le moyen de cadrage 105 extrait des portions d'image à partir d'images capturées par la caméra 103.



FR 2 880 762 - A1



## TERMINAL DE VISIOPHONIE AVEC REGLAGES INTUITIFS

L'invention concerne les téléphones portables à haut débit de données capables de capturer et de visualiser des images pour des communications en visiophonie. Ces téléphones sont également appelés terminaux de visiophonie. Plus particulièrement l'invention se rapporte au réglage des moyens de capture  
5 d'images.

L'arrivée du haut débit dans les télécommunications permet de mettre à disposition du grand public des services de visiophonie. Plus particulièrement, les réseaux de radiotéléphonie de troisième génération, tel que par exemple l'UMTS (de l'anglais Universal Mobile Telecommunication  
10 System), permettent des applications de visiophonie avec des téléphones ou terminaux portables.

La visiophonie permet à deux personnes de communiquer à distance tout en se voyant mutuellement. A cet effet, le terminal de chaque personne dispose d'un écran de visualisation et d'une caméra. Les terminaux portables  
15 disposent d'un écran de petite dimension ayant une faible résolution, et il est important d'avoir un bon cadrage en gros plan si l'on souhaite voir les traits de son interlocuteur.

Un inconvénient provient du fait qu'un interlocuteur A communicant avec un interlocuteur B ne voit que ce que la caméra de A lui transmet. L'interlocuteur B n'ayant pas le contrôle du cadrage, c'est l'interlocuteur A qui  
20 doit veiller à contrôler la prise de vue de sa caméra. Ce contrôle de prise de vue peut être réalisé à l'aide d'une vignette de contrôle dans un coin de son écran. Chaque interlocuteur doit alors veiller à ce que son image reste bien au centre de la vignette de contrôle.

Un tel système de cadrage n'est pas très pratique pour de nombreuses  
25 raisons. Il réduit la zone utile de l'écran de visualisation du terminal qui est déjà petite. Chaque interlocuteur doit faire très attention à son propre cadrage. Les mouvements de cadrage ne sont pas des mouvements naturels car la vignette montre une image filmée avec une inversion entre droite et gauche.

30 Pour les systèmes fixes de visioconférence, il est connu d'avoir recours

- 2 -

à une commande à distance de la caméra. Ainsi, l'interlocuteur A peut régler la caméra de B et réciproquement. Chaque utilisateur dispose d'une télécommande lui permettant d'envoyer les différentes commandes de zoom, et de déplacement. Un tel système ne peut pas être mis en place sur un terminal portable car les caméras des terminaux portables ne sont  
5 généralement pas mobiles et, en outre, cela nécessiterait d'utiliser les touches du clavier du terminal pendant la communication. Les touches d'un terminal portable sont petites, et il est peu pratique de les utiliser tout en gardant son terminal dans une direction de cadrage donnée.

10 L'invention propose de remédier aux problèmes de cadrage évoqués précédemment. Chaque terminal portable est muni d'une caméra ayant une résolution supérieure à la résolution de l'image transmise. Le cadrage de l'image transmise se fait à l'aide d'informations de cadrage provenant d'un interlocuteur distant. Des capteurs de mouvements sont placés dans chaque  
15 terminal portable pour récupérer des informations de mouvement dudit terminal. Les informations de mouvement sont ensuite transformées en commandes de cadrage à envoyer à l'autre terminal.

Selon un premier aspect, l'invention est un terminal portable de visiophonie comprenant des moyens de communication, des capteurs de  
20 mouvements et des moyens de mise en forme. Les moyens de communication permettent de communiquer avec un autre terminal de visiophonie par l'intermédiaire d'un réseau de radiocommunication. Les capteurs de mouvement produisent des informations représentatives de mouvements du terminal. Les moyens de mise en forme permettent de transformer les  
25 informations de mouvement en données de cadrage sortantes à destination de l'autre terminal. Les moyens de communication sont agencés pour insérer les données de cadrage sortantes dans des données à émettre sur le réseau de radiocommunication à destination de l'autre terminal.

Les moyens de mise en forme incluent des moyens de filtrage et  
30 d'élaboration de commande permettant de comparer les informations de mouvement à un seuil de mouvement minimum et un seuil de mouvement maximum. Les données de cadrage sortantes sont produites en réponse à la

détection d'un mouvement compris entre le seuil de mouvement minimum et le seuil de mouvement maximum.

Selon un deuxième aspect, l'invention est un terminal portable de visiophonie comprenant des moyens de communication, une caméra, un  
5 moyen de cadrage et un moyen d'encodage vidéo. Les moyens de communication permettent de communiquer avec un autre terminal de visiophonie par l'intermédiaire d'un réseau de radiocommunication. La caméra capture des images ayant une première taille. Le moyen de cadrage extrait une  
10 portion d'image à partir d'une image capturée par la caméra. Le moyen de cadrage sélectionne la portion d'image en fonction d'informations de cadrage distantes provenant de l'autre terminal. Ladite portion d'image ayant une deuxième taille inférieure à la première taille. Le moyen d'encodage vidéo transforme un flux de portions d'image provenant du moyen de cadrage en données vidéo sortantes. Les moyens de communication sont agencés pour  
15 extraire les informations de cadrage distantes de données reçues du réseau de radiocommunication en provenance de l'autre terminal.

Ainsi, un déplacement intuitif du terminal selon le premier aspect permet à un utilisateur A de recadrer l'image filmée par le terminal selon le deuxième aspect de son correspondant B. Préférentiellement, les deux aspects  
20 sont mis en œuvre sur un même terminal.

Selon un troisième aspect, l'invention est un procédé de recadrage d'une image prise par une caméra d'un premier terminal portable de visiophonie à l'aide d'un deuxième terminal portable de visiophonie muni d'un  
25 écran et de capteurs de mouvement au cours d'une communication en visiophonie. Le procédé comprend une étape d'élaboration de données de cadrage, en réponse à un mouvement du deuxième terminal portable, à partir d'informations de mouvement provenant des capteurs de mouvement, et une étape d'envoi desdites données de cadrage au premier terminal portable.

En réponse à des informations de cadrage reçues par le premier  
30 terminal et correspondant aux données de cadrage envoyées par le deuxième terminal, le premier terminal portable extrait des portions d'image correspondant auxdites informations de cadrage à partir d'images capturées

par sa caméra et produit une séquence vidéo représentative d'une succession de portions d'image.

Selon un dernier aspect, l'invention est un signal transportant un flux de trames de visiophonie entre un premier terminal portable et un deuxième terminal portable. Au moins une trame envoyée par le premier terminal comporte des données audio, des données vidéo et des données de cadrage. Les données de cadrage indiquent la position et/ou le déplacement d'une portion d'image prise par une caméra du deuxième terminal. Ladite portion d'image correspond à une image à envoyer du deuxième terminal au premier terminal.

L'invention sera mieux comprise et d'autres particularités et avantages apparaîtront à la lecture de la description qui va suivre, la description faisant référence aux figures annexées parmi lesquelles :

la figure 1 est un schéma fonctionnel représentant deux terminaux portables en communication selon l'invention,

la figure 2 illustre des recadrages d'images réalisés selon l'invention,

la figure 3 montre une caractéristique de transfert entre une information de mouvement et une commande de cadrage,

la figure 4 montre un organigramme de fonctionnement pour le terminal effectuant une capture d'image selon l'invention.

La figure 1 représente deux terminaux portables 100 et 200 communiquant par l'intermédiaire d'un réseau de radiotéléphonie ou radiocommunication 300. Le réseau de radiotéléphonie 300 est un réseau de radiocommunication dit de troisième génération, par exemple conforme à la norme UMTS. Par réseau de radiotéléphonie de troisième génération, il faut comprendre un réseau de radiocommunication à haut débit permettant d'échanger des données audio, vidéo ou autres entre un terminal radio et le réseau.

La présente description s'intéresse à la gestion du cadrage lors d'une communication en visiophonie entre deux terminaux. Seuls les moyens mis en œuvre pour le cadrage sont détaillés. Les autres éléments constitutifs des

terminaux et du réseau sont bien connus de l'homme du métier.

Pour simplifier la description, les deux terminaux 100 et 200 sont identiques. Des références similaires 1xx et 2xx sont utilisées pour décrire des éléments similaires, le chiffre des centaines permettant de différencier le terminal. Ainsi, ce qui est décrit en référence aux éléments du terminal 100 est applicable au terminal 200 et réciproquement.

Le premier terminal 100, de type téléphone portable, comporte :

- un microphone 101 pour capturer du son,
- un haut-parleur 102 pour reproduire du son,
- 10 - une caméra 103 pour capturer des images,
- un écran de visualisation 104 pour reproduire des images,
- un moyen de cadrage 105 relié à la caméra 103 pour extraire une portion d'image à partir d'une image capturée par la caméra 103, le moyen de cadrage 105 sélectionnant la portion d'image en fonction d'informations de cadrage  $I_C$  provenant d'un autre terminal distant,
- 15 - un moyen d'encodage audio et vidéo 106 relié au microphone 101 et au moyen de cadrage 105 pour transformer du son capturé par le micro 101 en données audio sortantes et un flux de portions d'image provenant des moyens de cadrage 105 en données vidéo sortantes, les données vidéo sortantes étant par exemple une séquence vidéo comprimée selon un algorithme de compression d'image,
- 20 - un moyen de décodage audio et vidéo 107 relié au haut-parleur 102 et à l'écran de visualisation 104 pour transformer des données audio entrantes en signal de pilotage du haut-parleur 102, et des données vidéo entrantes en signal d'image à reproduire sur l'écran de visualisation 104,
- 25 - des capteurs de mouvement 108 à 110 pour produire des informations de mouvements du terminal,
- un moyen de filtrage 111 relié aux capteurs de mouvement 108 à 110 pour filtrer les informations de mouvement,
- 30 - un moyen d'élaboration de commande 112 relié au moyen de filtrage 111 qui transforme les informations filtrées de mouvement en données de cadrage sortantes  $D_C$  à destination d'un autre terminal,

- 6 -

- un moyen de multiplexage et de communication 113 relié aux moyens d'encodage 106, d'élaboration de commande 112, de décodage 107 et de cadrage 105 pour, d'une part, regrouper des données audio, vidéo et de cadrage  $D_C$  à émettre en paquets de données à destination du réseau, et pour, d'autre part, recevoir et séparer des paquets de données reçus en données audio, en données vidéo et en informations de cadrage  $I_C$ , et
- une antenne 114 reliée aux moyens de multiplexage et de communication 113 pour échanger avec le réseau 300 des signaux radio représentatifs des paquets de données émis et reçus par le terminal 100.

Classiquement, pour maximiser l'intégration des composants d'un terminal portable, celui-ci comporte principalement un processeur central, un processeur de traitement du signal et éventuellement un processeur de traitement d'image. Ces trois processeurs sont utilisés de manière microprogrammée pour traiter toutes les données et tous les signaux sous forme numérique. Ainsi, les moyens 105-107 et 111-112 décrits de manière fonctionnelle sur la figure 1 peuvent être réalisés en pratique par programmation de ces processeurs. Des convertisseurs analogique/numérique et numérique/analogique réalisent la liaison entre les processeurs et les différents éléments 101-104 et 108-110 auxquels ils sont connectés. Le moyen de multiplexage et de communication 113 est également réalisé à l'aide des processeurs du terminal mais il inclut en outre une interface radio reliée à l'antenne 114.

Lors d'une communication en visiophonie, le moyen de décodage audio et vidéo 107 reçoit les données audio et vidéo provenant du terminal distant 200 et les transforme en signaux de commande du haut-parleur 102 et de l'écran 104 afin de restituer à un utilisateur A du terminal 100 le son capturé par le micro 202 et l'image capturée par la caméra 203 du terminal 200 de son correspondant B. L'écran 104 est par exemple de type LCD (de l'anglais : Liquid Crystal Display) ou de type OLED (de l'anglais : Organic Light-Emitting Diode display) de petite dimension, par exemple d'une diagonale inférieure à deux pouces. La résolution de cet écran 104 est par exemple inférieure à 200

pixels par coté. L'utilisateur A du terminal 100 peut ainsi avoir une image satisfaisante lui permettant de saisir les expressions d'un visage en gros plan. Cependant, une telle résolution ne permet pas de distinguer ces mêmes détails en plan plus large.

5           La caméra 103 du terminal 100 comporte une lentille grand angle et un capteur d'image, par exemple de type CCD (de l'anglais : Charge Coupled Device). La caméra 103 sert pour la visiophonie mais également pour d'autres applications parmi lesquelles la photographie numérique. Selon une technique connue, les photos prises par le terminal portable 100 peuvent être envoyées à  
10   travers le réseau 300 à un ordinateur. Ainsi la caméra 103 dispose généralement d'une résolution supérieure à la résolution de l'écran 104. A titre d'exemple, la résolution de la caméra 103 est d'au moins 640x480 pixels afin d'avoir une qualité minimum de visualisation sur un écran d'ordinateur.

          Lors d'une communication en visiophonie, le moyen de cadrage 105  
15   sert à extraire, de chaque image prise par la caméra 103, une portion d'image dont la forme et la résolution correspondent à l'écran 204 du terminal 200. Ces informations de forme et de résolution sont transmises lors de l'initialisation de la communication. Le moyen de cadrage comporte également des moyens pour effectuer un sur-échantillonnage et/ou un sous-échantillonnage de l'image  
20   permettant de réaliser une fonction de zoom électronique selon une technique connue. Le zoom électronique permet de transformer une portion d'image choisie de taille quelconque afin de l'adapter à la taille de l'écran 204 du terminal 200 du correspondant B.

          Le moyen de cadrage 105 comporte un stabilisateur d'image apte à  
25   compenser les déplacements de faible amplitude de l'image qui correspondent à d'éventuels tremblements d'un utilisateur. Le stabilisateur d'image, d'un type connu, est par exemple capable de détecter tout mouvement global et uniforme entre deux images provenant de la caméra et d'effectuer une translation correspondante en nombre de pixels dans la direction inverse du déplacement.  
30   La stabilisation de l'image se fait de préférence avant l'extraction de la portion d'image.

          Dans l'invention, le moyen de cadrage 105 déplace en outre la portion

- 8 -

d'image en fonction d'informations de cadrage  $I_C$  provenant du moyen de multiplexage 113. Les informations de cadrage  $I_C$  reçues par l'un des terminaux 100, respectivement 200, correspondent à des données de cadrage  $D_C$  émises par l'autre des terminaux 200, respectivement 100.

5            Selon l'invention, le cadrage est réalisé à distance par le correspondant qui est le mieux placé pour régler l'image visualisée. L'invention propose un contrôle intuitif du cadrage à distance. Lorsque l'utilisateur A du terminal 100 voit l'utilisateur B du terminal 200 sortir de son écran 104, le geste le plus intuitif est de déplacer son écran 104, tenu à la main, pour suivre le  
10 mouvement de l'utilisateur du terminal 200. Egalement, lorsqu'un détail attire particulièrement l'attention et que l'on désire le voir de plus près, un utilisateur rapproche naturellement son écran. Le contrôle intuitif proposé consiste à déplacer le terminal dans une direction naturelle pour déplacer la portion d'image vue sur l'écran.

15            La figure 2 représente les différentes modifications de cadrage en fonction d'un déplacement du terminal. La figure 2a montre une image 400 capturée par la caméra 203 du terminal 200 et une portion d'image 401 transmise et vue sur l'écran 104 du terminal 100. L'utilisateur du terminal 100 souhaite déplacer le cadrage de la portion d'image 401 pour obtenir une  
20 portion d'image 402 cadrée différemment. L'utilisateur déplace son terminal 100 d'un mouvement  $\Delta m$  dans la direction souhaitée. Ce mouvement  $\Delta m$  peut se décomposer en un mouvement  $\Delta x$  selon un premier axe parallèle à un premier coté de l'écran 104 et en un mouvement  $\Delta y$  selon un axe selon un deuxième axe parallèle à un deuxième coté de l'écran 104. Le mouvement est  
25 appliqué proportionnellement à la position du cadre de la portion d'image.

              La figure 2b montre une image 400 capturée par la caméra 203 du terminal 200 et une portion d'image 403 transmise et vue sur l'écran 104 du terminal 100. L'utilisateur du terminal 100 souhaite avoir une vision en plan plus large correspondant à la portion d'image 404. L'utilisateur déplace alors  
30 son terminal 100 en éloignant l'écran 104 d'un mouvement d'éloignement  $\Delta z$  qui provoque un élargissement du cadre. Un sous-échantillonnage est ensuite

réalisé pour adapter la portion d'image à la taille de l'écran 104, ce qui correspond à un zoom arrière.

Les données de cadrage  $D_C$  sont élaborées par le moyen d'élaboration de commande 112 en fonction d'informations de mouvement provenant des capteurs de mouvement 108 à 110 après filtrage dans le moyen de filtrage 111. Les capteurs de mouvement 108 à 110 sont par exemple constitués d'accéléromètres, de gyroscopes et/ou de magnétomètres électroniques capables de fournir des informations relatives au déplacement et à l'accélération en translation et en rotation selon trois axes perpendiculaires. Préférentiellement, deux des trois axes sont respectivement parallèles aux côtés de l'écran 104 et le troisième axe est perpendiculaire à l'écran 104. Les mouvements de translation et de rotation sont combinés entre eux pour obtenir des informations de mouvement  $\Delta x$ ,  $\Delta y$  et  $\Delta z$  représentatives d'un déplacement relatif de l'écran 104 pendant une durée prédéterminée selon l'un des trois axes.

La durée prédéterminée correspond par exemple à une durée d'échantillonnage des informations de mouvement  $\Delta x$ ,  $\Delta y$  et  $\Delta z$ . Le moyen de filtrage 111 réalise ensuite un filtrage des informations de mouvement  $\Delta x$ ,  $\Delta y$  et  $\Delta z$ . Les informations de mouvement  $\Delta x$ ,  $\Delta y$  et  $\Delta z$  sont représentatives d'une amplitude et d'une vitesse de déplacement. Pour éviter de prendre en compte des mouvements liés à des tremblements (mouvements de faible amplitude) ou à une fatigue (mouvements lents) de l'utilisateur, il convient de ne tenir compte que des informations de mouvement supérieures, en valeur absolue, à un seuil minimum  $S_m$ .

Par ailleurs, l'utilisateur peut bouger pendant la conversation en visiophonie sans vouloir modifier le cadrage. A cet effet, il convient de ne pas prendre en compte les informations de mouvement de grande amplitude, c'est-à-dire d'amplitude supérieure, en valeur absolue, à un seuil maximum  $S_M$ . Le filtre dispose de la fonction de transfert suivante pour les informations de mouvement  $\Delta x$ ,  $\Delta y$  et  $\Delta z$  :

$$\Delta u_f = \Delta u \text{ si } S_m < |\Delta u| < S_M, \text{ et}$$

- 10 -

$$\Delta u^f = 0 \text{ si } |\Delta u| \leq S_m \text{ ou si } |\Delta u| \geq S_M,$$

avec  $u$  remplaçant  $x$ ,  $y$  ou  $z$ , et  $\Delta x^f$ ,  $\Delta y^f$  et  $\Delta z^f$  correspondant aux informations de mouvement filtrées.

Le moyen d'élaboration de commande 112 transforme les informations  
 5 de mouvement filtrées  $\Delta x^f$ ,  $\Delta y^f$  et  $\Delta z^f$  en données de cadrage  $D_C$ . Les données de cadrage  $D_C$  peuvent être de différentes formes. A titre d'exemple, deux formes de données sont exposées ci-après.

Une première forme des données de cadrage est une forme  
 correspondant à des commandes de cadrage. Les commandes vont être  
 10 constituées de trois données représentatives de la modification du cadrage. Les informations filtrées  $\Delta x^f$  et  $\Delta y^f$  sont quantifiées pour transformer le mouvement en nombre (positif ou négatif) de pixels de déplacement du cadre définissant la portion d'image. L'information filtrée  $\Delta z^f$  est quantifiée pour  
 15 indiquer le nombre de pixels augmentant ou diminuant le cadre définissant la portion d'image. La position et la dimension du cadre de la portion d'image sont ensuite gérées par le moyen de cadrage 205 de l'autre terminal 200.

Un exemple de caractéristique de transfert entre une information de mouvement et une donnée de mouvement correspondante est représenté sur la figure 3. L'axe des abscisses correspond aux valeurs possibles d'une  
 20 information de mouvement, par exemple  $\Delta x$ . Trois zones inopérantes 410 à 411 correspondent à l'annulation de l'information de mouvement réalisée par le moyen de filtrage 111. Deux zones de quantification 413 et 414 correspondent à la commande de déplacement du cadre définissant la portion d'image. A titre d'exemple, il est possible de faire correspondre une valeur de 1 pixel de  
 25 déplacement lorsque l'information de mouvement correspond au seuil minimum  $S_m$  et une valeur de 20 pixels de déplacement lorsque l'information de mouvement correspond au seuil maximum  $S_M$ . Dans une réalisation simple, une extrapolation linéaire est réalisée entre les deux seuils  $S_m$  et  $S_M$  ; mais on notera qu'une courbe différente pourrait être utilisée.

30 Une deuxième forme des données de cadrage peut consister en une

position et une taille de cadre d'une portion d'image choisie. Dans ce cas, les modifications du cadre sont réalisées par le moyen d'élaboration de commande 112. Le moyen de cadrage 205 n'effectue que la sélection de la portion d'image accompagnée éventuellement d'un effet de zoom.

5           Quelle que soit la forme des données de cadrage  $D_C$ , celles-ci sont fournies au moyen de multiplexage et de communication 113. Le moyen de multiplexage et de communication 113 constitue des trames de données à émettre regroupant un ensemble de données à destination du terminal 200. La trame comporte des données audio et des données vidéo relatives à la  
10 communication visiophonique mais aussi les données de cadrage  $D_C$ . La trame est ensuite empaquetée avec des données de service destinées à identifier la trame et sa destination. Le paquet de données ainsi constitué est transformé en une salve qui s'insère dans un signal radio constitué d'un flux de trames à destination du réseau 300. Le réseau 300 aiguille ensuite les différentes trames  
15 pour les envoyer au terminal 200 sous forme d'un signal radio.

Le moyen de cadrage 105 dispose d'un stabilisateur d'image qui compense les mouvements de la caméra lorsque ceux-ci sont de faible amplitude. Cette compensation permet d'éviter que les déplacements opérés pour recadrer l'image de l'interlocuteur entraîne à son tour un besoin de  
20 recadrage réciproque. Si le seuil maximum  $S_M$  correspond à un déplacement compensable par le stabilisateur d'image, le mouvement réalisé pour modifier le cadrage de l'autre terminal 200 est automatiquement compensé.

Si le stabilisateur d'image n'a pas la capacité suffisante pour compenser le mouvement lié à un déplacement de recadrage, il est possible de  
25 relier le moyen de cadrage 105 au moyen de filtrage 111 afin que le déplacement réel du terminal 100 soit pris en compte pour déplacer sa fenêtre de cadrage. Le recadrage effectué dans le terminal 100 est de même nature mais d'amplitude et de signe différents par rapport au recadrage de l'autre terminal 200.

30           Avec les moyens précédemment décrits, il est possible de recadrer les images filmées par la caméra 103 du terminal 100 à partir du terminal 200 et réciproquement. Le recadrage se fait intuitivement en réponse à un

mouvement du terminal 200 qui entraîne l'élaboration de données de cadrage grâce aux capteurs de mouvement 208 à 210, au moyen de filtrage 211 et au moyen d'élaboration de commande 212. Ainsi la séquence vidéo produite par le terminal 100 est contrôlée par l'utilisateur B du terminal 200 pendant une  
5 communication en visiophonie.

Toutefois, un tel recadrage est utile pour un dialogue où les deux utilisateurs sont face à face et/ou quasi-immobiles. Lorsque l'un des utilisateurs déplace de manière conséquente son terminal, par exemple pour montrer quelque chose à son correspondant, le recadrage devient superflu. En outre, si  
10 le recadrage est réalisé de manière permanente pendant une longue conversation en visiophonie où des mouvements importants ont été réalisés, le cadre délimitant la portion d'image peut se trouver sur un bord de l'image filmée par la caméra et le recadrage peut devenir impossible dans une direction.

Selon une première amélioration, on recadre l'image filmée à l'aide du moyen de cadrage 105 du terminal 100 si ce terminal 100 est quasi immobile. La détection de la quasi-immobilité du terminal 100 peut se faire à l'aide du stabilisateur d'image qui détecte un mouvement homogène de faible amplitude de l'image. Il est également possible de détecter le mouvement du terminal 100  
15 à l'aide des capteurs de mouvement 108 à 110. Ainsi, si les mouvements sont, par exemple inférieurs au seuil maximum  $S_M$ , on peut considérer que le terminal est quasi-immobile.

Selon une deuxième amélioration, le moyen de cadrage 105 comporte des moyens de reconnaissance de forme aptes à reconnaître la forme d'un visage. La reconnaissance de forme se fait par exemple à l'aide d'une  
25 technique connue pour identifier qu'un visage est présent dans l'image. Ainsi, le terminal 100 peut établir que l'utilisateur est en train de dialoguer si un visage est détecté. Le moyen de cadrage prend alors en compte les informations de cadrage si un visage est détecté. Cela permet de contrôler le  
30 cadrage à distance même si l'arrière-plan filmé est mobile ou si le terminal est en mouvement.

De manière préférentielle, les deux améliorations sont combinées.

L'organigramme de la figure 4 illustre la mise en œuvre de ces deux améliorations dans le moyen de cadrage 105. Lors d'une étape 420 d'initialisation de communication visiophonique, le moyen de cadrage 105 reçoit les caractéristiques de l'écran 204 du terminal 200 du correspondant  
5 pour pouvoir déterminer le format de la portion d'image à envoyer.

Une fois la communication initialisée, le moyen de cadrage 105 fonctionne en image libre (étape 421). Au cours de cette étape 421, le moyen de cadrage découpe une portion d'image, par exemple centrée, indépendamment de toute information de cadrage  $I_C$  reçue.

10 De manière régulière, un test 422 est réalisé pour déterminer si des conditions permettent de passer en mode de contrôle à distance de l'image. Le test 422 consiste par exemple à vérifier si l'image peut être considérée comme immobile ou quasi-immobile ou si un visage est présent dans l'image filmée. Si l'une des deux conditions est réalisée, on passe à l'étape 423. Si aucune des  
15 deux conditions n'est réalisée, un test 424 est réalisé.

Le test 424 contrôle si la communication visiophonique est terminée. Si la communication n'est pas terminée, on retourne à l'étape 421 et le moyen de cadrage fonctionne en image libre.

L'étape 423 correspond au fonctionnement du moyen de cadrage 105  
20 avec prise en compte des informations de cadrage  $I_C$ . La portion d'image est déplacée, agrandie ou rétrécie en fonction des données de cadrage reçues. Le moyen de cadrage effectue, si besoin est, un sur-échantillonnage ou un sous-échantillonnage de la portion d'image cadrée pour la faire correspondre à la résolution de l'écran.

25 De manière régulière, un test 425 est réalisé pour déterminer si des conditions permettent de passer en mode de contrôle à distance de l'image. Le test 425 consiste par exemple à vérifier si l'image peut être considérée comme immobile ou quasi-immobile ou si un visage est présent dans l'image filmée. Si l'une des deux conditions est réalisée, on retourne à l'étape 423. Si aucune des  
30 deux conditions n'est réalisée, un test 426 est réalisé.

Le test 426 contrôle si la communication en visiophonie est terminée. Si la communication n'est pas terminée, la portion d'image est réinitialisée et

on retourne à l'étape 421 et le moyen de cadrage fonctionne en image libre.

L'invention décrite peut se décliner selon différentes variantes. Dans l'exemple décrit, les terminaux 100 et 200 sont identiques et tous deux capables d'émettre des informations de cadrage et de recevoir et prendre en  
5 compte des informations de cadrage. Cependant, l'homme du métier comprendra qu'un terminal pourrait élaborer et émettre des données de cadrage sans recevoir d'informations de cadrage. Réciproquement, un terminal peut prendre en compte des informations de cadrage sans lui-même élaborer de données de cadrage. Les terminaux mettant en œuvre l'invention peuvent  
10 donc se limiter aux moyens nécessaires à la mise en œuvre faite de l'invention sans reproduire tous les moyens compris dans les exemples décrits.

## REVENDICATIONS

1. Terminal portable de visiophonie (100, 200) comprenant :
  - des moyens de communication (113, 213) pour communiquer avec un autre terminal de visiophonie (200, 100) par l'intermédiaire d'un réseau de radiocommunication (300),
  - des capteurs de mouvement (108-110, 208-210) pour produire des informations ( $\Delta x$ ,  $\Delta y$ ,  $\Delta z$ ) représentatives de mouvements du terminal, et
  - des moyens de mise en forme (111, 112, 211, 212) pour transformer les informations de mouvement en données de cadrage ( $D_C$ ) sortantes à destination de l'autre terminal,les moyens de communication (113, 213) étant agencés pour insérer les données de cadrage ( $D_C$ ) sortantes dans des données à émettre sur le réseau de radiocommunication à destination de l'autre terminal.
2. Terminal selon la revendication 1, dans lequel les capteurs de mouvement (108-110, 208-210) comprennent des gyroscopes et/ou des accéléromètres et/ou des magnétomètres.
3. Terminal selon la revendication 1 ou 2, dans lequel les moyens de mise en forme (111, 112, 211, 212) sont agencés pour comparer les informations de mouvement à un seuil de mouvement minimum ( $S_m$ ) et un seuil de mouvement maximum ( $S_M$ ), et dans lequel les données de cadrage ( $D_C$ ) sortantes sont produites en réponse à la détection d'un mouvement compris entre le seuil de mouvement minimum ( $S_m$ ) et le seuil de mouvement maximum ( $S_M$ ),
4. Terminal selon l'une quelconque des revendications 1 à 3, comportant en outre :
  - une caméra (103, 203) pour capturer des images ayant une première taille,
  - un moyen de cadrage (105, 205) pour extraire une portion d'image (401-404) à partir d'une image capturée (400) par la caméra (103, 203),

- 16 -

le moyen de cadrage (105, 205) sélectionnant la portion d'image en fonction d'informations de cadrage ( $I_C$ ) distantes provenant de l'autre terminal, ladite portion d'image ayant une deuxième taille inférieure à la première taille,

- 5        - un moyen d'encodage vidéo (106, 206) pour transformer un flux de portions d'image provenant du moyen de cadrage (105, 205) en données vidéo sortantes, et

les moyens de communication (113, 213) étant agencés pour extraire les informations de cadrage ( $I_C$ ) distantes de données reçues du réseau de radiocommunication (300) en provenance de l'autre terminal (200, 100).  
10

5.        Terminal portable de visiophonie (100, 200) comprenant :

- des moyens de communication (113, 213) pour communiquer avec un autre terminal de visiophonie (200, 100) par l'intermédiaire d'un réseau de radiocommunication (300),  
15        - une caméra (103, 203) pour capturer des images ayant une première taille,  
- un moyen de cadrage (105, 205) pour extraire une portion d'image (401-404) à partir d'une image capturée (400) par la caméra (103, 203), le moyen de cadrage (105, 205) sélectionnant la portion d'image en  
20        fonction d'informations de cadrage ( $I_C$ ) distantes provenant de l'autre terminal, ladite portion d'image ayant une deuxième taille inférieure à la première taille,  
- un moyen d'encodage vidéo (106, 206) pour transformer un flux de portions d'image provenant du moyen de cadrage (105, 205) en  
25        données vidéo sortantes, et

les moyens de communication (113, 213) étant agencés pour extraire les informations de cadrage ( $I_C$ ) distantes de données reçues du réseau de radiocommunication (300) en provenance de l'autre terminal (200, 100).

6.        Terminal selon l'une des revendications 4 ou 5, dans lequel le  
30        moyen de cadrage (105, 205) est agencé pour prendre en compte les informations de cadrage ( $I_C$ ) distantes si l'image capturée est considérée comme immobile ou quasi-immobile.

7. Terminal portable selon l'une des revendications 4 à 6, dans lequel le moyen de cadrage (105, 205) comporte des moyens de reconnaissance de forme apte à reconnaître un visage, et dans lequel le moyen de cadrage (105, 205) est agencé pour prendre en compte les informations de cadrage ( $I_C$ )  
5 distantes si un visage est détecté.
8. Procédé de recadrage d'une image prise par une caméra (103) d'un premier terminal portable de visiophonie (100) à l'aide d'un deuxième terminal portable de visiophonie (200) muni d'un écran (204) et de capteurs de mouvement (208-210) au cours d'une communication en visiophonie, le  
10 procédé comprenant une étape d'élaboration de données de cadrage ( $D_C$ ), en réponse à un mouvement du deuxième terminal portable (200), à partir d'informations de mouvement ( $\Delta x$ ,  $\Delta y$ ,  $\Delta z$ ) provenant des capteurs de mouvement (208-210), et une étape d'envoi desdites données de cadrage ( $D_C$ ) au premier terminal portable (100).
- 15 9. Procédé selon la revendication 8, dans lequel le deuxième terminal portable (200) élabore les données de cadrage ( $D_C$ ) si les informations de mouvement ( $\Delta x$ ,  $\Delta y$ ,  $\Delta z$ ) sont supérieures à un seuil de mouvement minimum ( $S_m$ ) et si ces informations de mouvement sont inférieures à un seuil de mouvement maximum ( $S_M$ ).
- 20 10. Procédé selon l'une des revendications 8 ou 9, dans lequel, en réponse à des informations de cadrage ( $I_C$ ) reçues par le premier terminal (100) et correspondant aux données de cadrage ( $D_C$ ) envoyées par le deuxième terminal (200), le premier terminal portable (100) extrait des portions d'image correspondant auxdites informations de cadrage ( $I_C$ ) à partir d'images  
25 capturées par sa caméra (103) et produit une séquence vidéo représentative d'une succession de portions d'image.
11. Procédé selon la revendication 10, dans lequel les informations de cadrage ( $I_C$ ) sont prises en compte si l'image prise par la caméra du premier terminal portable (100) bouge faiblement ou si un visage est détecté dans  
30 l'image filmée.

- 18 -

12. Signal transportant un flux de trames de visiophonie entre un premier terminal portable (100) et un deuxième terminal portable (200), caractérisé en ce qu'au moins une trame envoyée par le premier terminal (100) comporte des données audio, des données vidéo et des données de cadrage (D<sub>C</sub>), les données de cadrage indiquant la position et/ou le déplacement d'une
- 5 portion d'image prise par une caméra (203) du deuxième terminal (200), ladite portion d'image correspondant à une image à envoyer du deuxième terminal (200) au premier terminal (100).

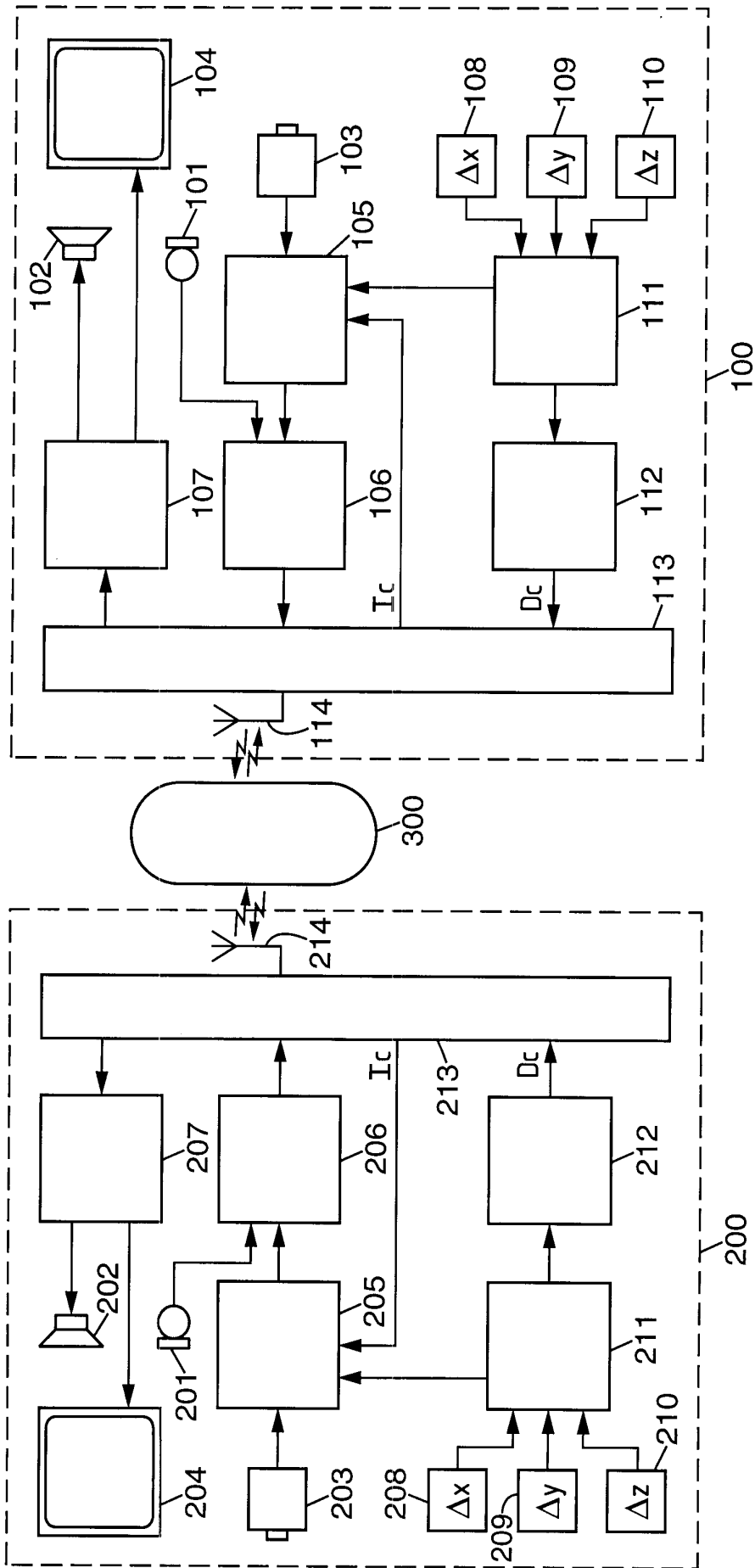
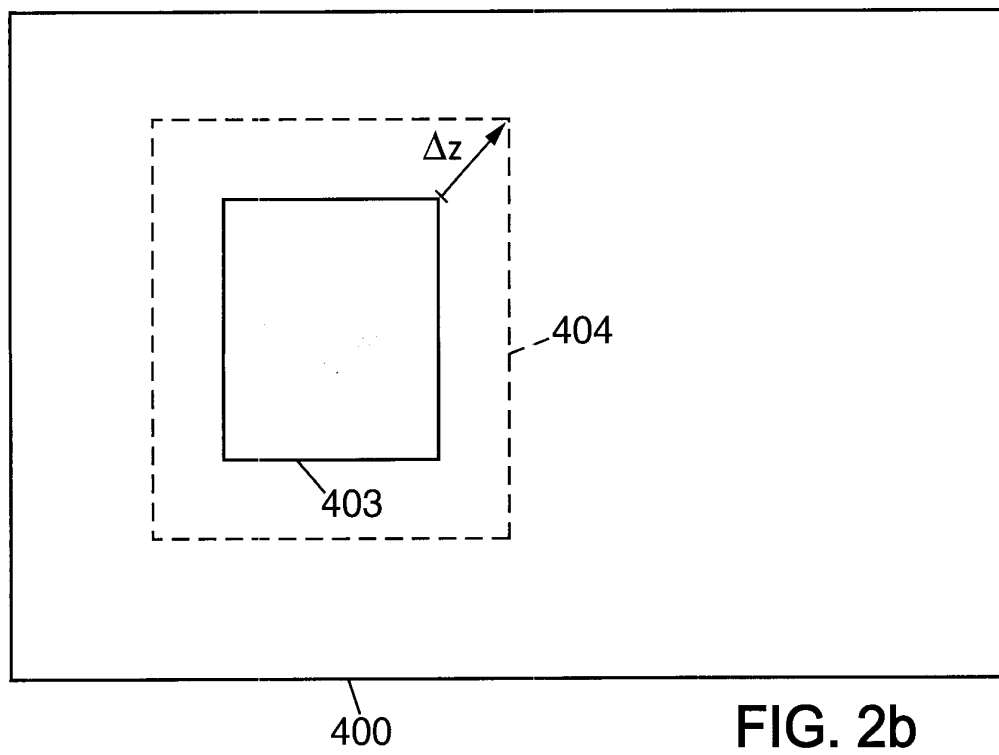
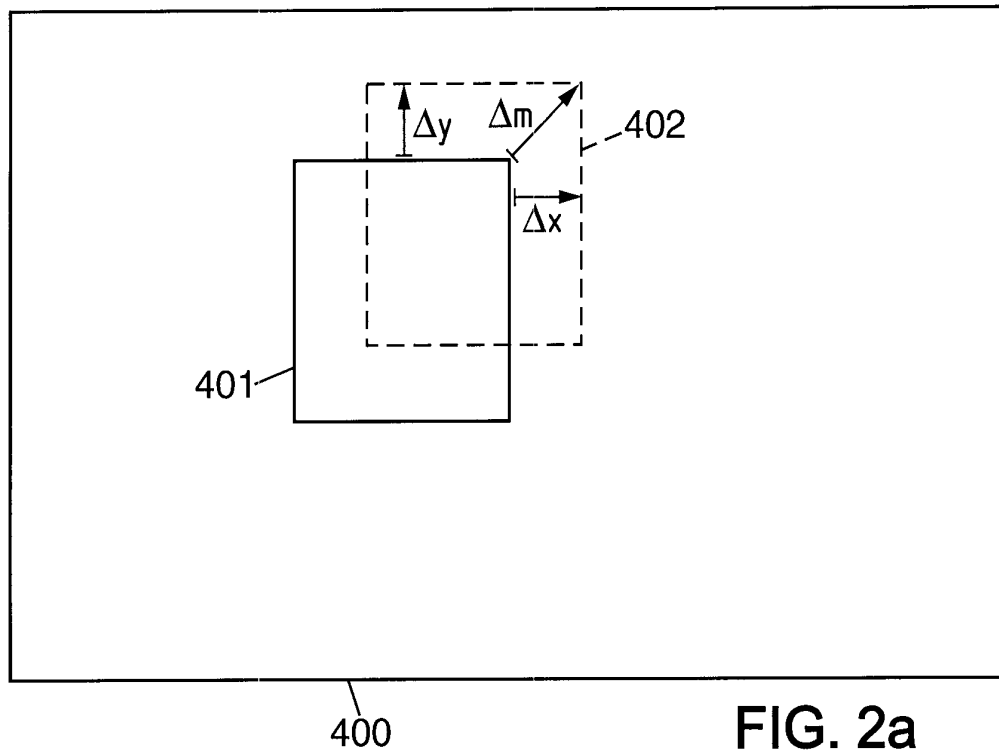
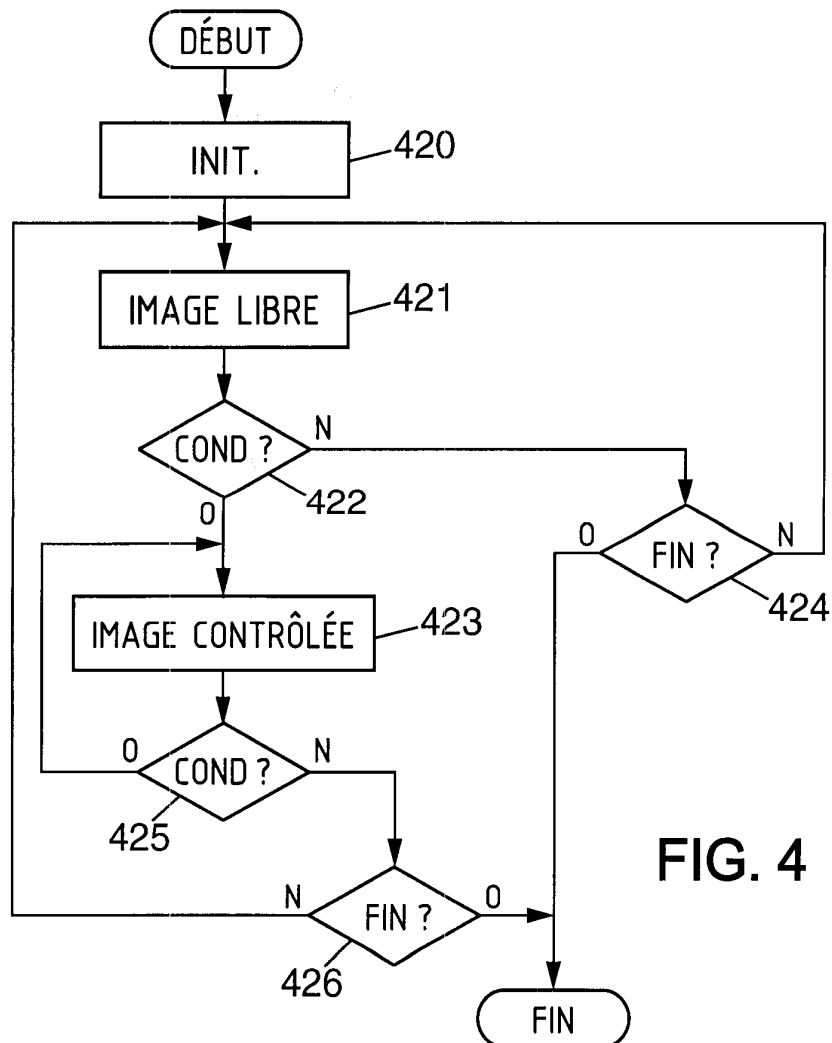
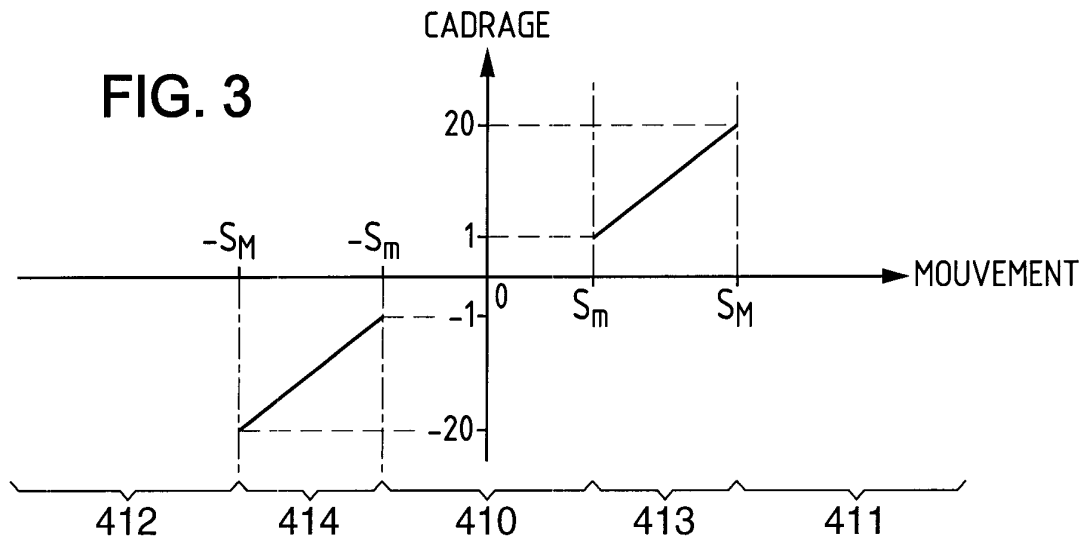


FIG. 1







**RAPPORT DE RECHERCHE  
PRÉLIMINAIRE**  
établi sur la base des dernières revendications  
déposées avant le commencement de la recherche

N° d'enregistrement  
national

FA 659672  
FR 0500180

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
X	EP 1 304 853 A (NOKIA CORPORATION) 23 avril 2003 (2003-04-23) * alinéa [0007] - alinéa [0008] * * alinéa [0012] - alinéa [0013] * * alinéa [0019] * * alinéa [0021] * * figure 3 *	1-5,8-12	H04Q7/32 H04M3/56 H04N7/15
A	EP 1 377 040 A (STMICROELECTRONICS S.R.L.) 2 janvier 2004 (2004-01-02) * alinéa [0004] * * alinéa [0036] * * alinéa [0053] - alinéa [0054] * * alinéa [0077] * * alinéa [0106] * * figures 5,6 *	2,9	DOMAINES TECHNIQUES RECHERCHÉS (Int.CL.7)  H04N
Date d'achèvement de la recherche		Examineur	
23 août 2005		Van der Zaal, R	
<p>CATÉGORIE DES DOCUMENTS CITÉS</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire</p>		<p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons ..... &amp; : membre de la même famille, document correspondant</p>	

**ANNEXE AU RAPPORT DE RECHERCHE PRÉLIMINAIRE  
RELATIF A LA DEMANDE DE BREVET FRANÇAIS NO. FR 0500180 FA 659672**

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche préliminaire visé ci-dessus.

Les dits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du 23-08-2005

Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets, ni de l'Administration française

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
EP 1304853 A	23-04-2003	US 2003076408 A1	24-04-2003
		EP 1304853 A2	23-04-2003
		JP 2003204466 A	18-07-2003
-----			
EP 1377040 A	02-01-2004	EP 1377040 A1	02-01-2004
		US 2004001147 A1	01-01-2004
-----			