(19) RÉPUBLIQUE FRANÇAISE

INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE

PARIS

(11) N° de publication : **2 758 428**
(à n'utiliser que pour les commandes de reproduction)

(21) N° d'enregistrement national : **97 00423**

(51) Int Cl[6] : **H 04 N 5/222**

(12) DEMANDE DE BREVET D'INVENTION A1

(22) **Date de dépôt** : 16.01.97.

(30) **Priorité** :

(43) **Date de la mise à disposition du public de la demande :** 17.07.98 Bulletin 98/29.

(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) :** *TRANSCOM PRODUCTIONS LTD* — IE.

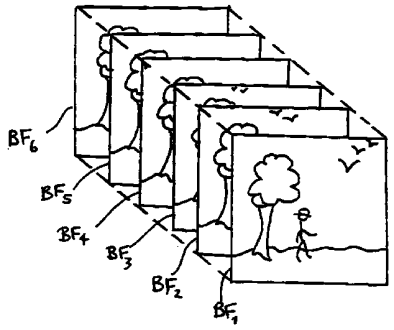
(72) **Inventeur(s) :** GOLDBERG MOZEST et MADRANE NABIL.

(73) **Titulaire(s) :** .

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

(54) INTERFACES VIDEO INTERACTIVES.

(57) Une interface interactive permettant d'interagir avec des informations vidéo crée une représentation visuelle affichée d'un quasi-object appelé une présentation visuelle racine: La présentation visuelle racine consiste en un ensemble d'images fondamentales sélectionnées parmi les informations vidéo et disposées afin que leurs axes x et y respectifs soient alignés avec les axes x et y de la présentation visuelle racine et que l'axe z de la présentation visuelle racine corresponde à l'axe des temps, de manière à ce que les images fondamentales soient espacées sur l'axe z de la présentation visuelle racine selon leur séparation temporelle. La représentation visuelle affichée de la présentation visuelle racine change selon la position désignée d'observation comme si la présentation visuelle racine était un objet tridimensionnel. L'utilisateur peut manipuler la représentation affichée en désignant des positions d'observations différentes, en choisissant des parties des informations vidéo à reproduire et en commandant des effets spéciaux tels que de couper le quasi-objet afin d'obtenir une meilleure vue d'une image en particulier. Un kit d'outils permet à un concepteur d'interfaces de concevoir de telles interfaces, notamment afin de limiter et de diriger la façon dont l'utilisateur final peut interagir avec l'interface.§.



FR 2 758 428 - A1

La présente invention concerne le domaine des interfaces permettant d'accéder à des informations vidéo. Plus précisément, la présente invention prévoit une interface interactive permettant d'accéder à des informations vidéo, et un kit d'outils destiné à être utilisé pour la création de telles interfaces interactives.

Des informations vidéo sont produites à une vitesse toujours croissante et des séquences vidéo, notamment des séquences courtes, sont utilisées de plus en plus, par exemple dans les serveurs du réseau Internet (site "web") et sur des disques CD-ROM, et elles sont créées par exemple par utilisation domestique de camescopes. Des outils permettant l'indexation et la manipulation des données vidéo et l'interaction avec celles-ci sont de plus en plus nécessaires. Il est en particulier nécessaire que des interfaces permettent à un utilisateur d'accéder sélectivement à des informations vidéo et d'interagir avec ces informations, notamment d'une façon non séquentielle.

De manière classique, les informations vidéo sont constituées d'une séquence d'images enregistrées à des intervalles fixes de temps ; dans le cas des signaux classiques de télévision par exemple, les informations vidéo comprennent vingt-cinq ou trente images par seconde. Chaque image est significative car elle correspond à une présentation visuelle qui peut être observée. Une image peut être formée d'un certain nombre de trames entrelacées, mais cette caractéristique n'est pas obligatoire, comme on peut le noter avec les formats vidéo les plus récemment proposés, tels que ceux qui sont destinés à la télévision à haute définition. Les images décrivent la décomposition des informations vidéo picturales dans le temps. Chaque image contient des informations de présentation visuelle structurées en lignes et en "éléments d'image" (pixels), qui représentent la décomposition spatiale des informations vidéo.

Dans le présent document, les expressions "informations vidéo" ou "séquences vidéo" désignent des données représentant une présentation visuelle enregistrée sur une période déterminée, sans référence à la longueur de cette période ou à la structure des informations enregistrées. Ainsi, le terme "séquence vidéo" est utilisé pour désigner une série quelconque d'images vidéo, indépendamment du fait que cette série correspond à une seule prise de vues par une caméra (enregistrée entre deux interruptions) ou plusieurs scènes ou prises de vues.

Traditionnellement, lorsqu'un utilisateur a voulu savoir quel était le contenu d'une séquence vidéo particulière, il a été obligé d'observer chaque image ou un sous-échantillon des images de la séquence lors de son affichage successif

au cours du temps. Cette solution est encore très répandue et, dans les applications dans lesquelles des données vidéo sont obtenues avec un ordinateur personnel, l'interface des informations vidéo est souvent constituée par affichage d'une fenêtre dans laquelle est contenue la séquence vidéo, avec un ensemble de commandes affichées analogues à celles qui se trouvent sur un magnétoscope (permettant une avance rapide, un réenroulement, etc.).

Les progrès réalisés dans les domaines de l'indexation vidéo et de l'édition vidéo ont conduit à d'autres formes d'interfaces pour consulter des informations vidéo.

Dans le domaine de l'indexation vidéo, il est nécessaire de coder des informations contenues dans une séquence vidéo pour pouvoir retrouver ultérieurement la séquence dans une base de données par référence à des mots-clés ou des concepts. Le contenu codé peut par exemple identifier les types d'objets présents dans la séquence vidéo, leurs propriétés ou leur déplacement, le type des mouvements de la caméra mis en oeuvre dans la séquence vidéo (panoramique, travelling suivant le sujet, variation de distance focale, etc.) et d'autres propriétés. Un "résumé" du document codé peut être préparé et peut comprendre certaines images représentatives tirées de la séquence, avec des informations sous forme de texte ou d'icônes indiquant comment la séquence a été codée. L'interface permettant l'interaction avec la base de données vidéo comprend habituellement un dispositif de saisie d'ordinateur permettant à l'utilisateur de spécifier les objets ou propriétés intéressants et, à la suite de cette demande, l'ordinateur détermine quelles séquences vidéo de la base de données correspondent aux termes saisis pour la recherche et affiche les "résumés" appropriés. L'utilisateur indique alors si une séquence vidéo particulière doit être reproduite ou non. Des exemples de produits mettant en oeuvre une telle solution sont décrits dans l'article "Advanced Imaging Product Survey : Photo, Document and Video" de la revue "Advanced Imaging", octobre 1994.

Dans certains schémas d'indexation vidéo, la séquence vidéo est divisée en séries plus courtes d'images en fonction des changements de scène ou du contenu sémantique des informations vidéo. Une structure hiérarchique peut être définie. Des "résumés" d'indexation peuvent être produits pour les différentes séries d'images correspondant aux noeuds de la structure hiérarchique. Dans un tel cas, au moment où la recherche est effectuée, le "résumé" correspondant à une séquence vidéo complète peut être retrouvé pour être affiché pour l'utilisateur qui peut alors demander l'affichage des "résumés" relatifs à des sous-sections de la séquence

vidéo qui se trouvent plus bas dans la structure hiérarchique. Si l'utilisateur le souhaite, une séquence ou sous-section choisie est reproduite sur le moniteur d'affichage. Ce schéma est décrit dans le document EP-A-0 555 028.

Un inconvénient de ces interfaces traditionnelles d'indexation et de recherche de séquences vidéo est que la qualité dynamique des informations vidéo est perdue.

Une autre solution, tirée du domaine de l'édition vidéo, est constituée de la "mosaïque numérique". La séquence vidéo est segmentée en scènes, et une ou plusieurs images représentatives de chaque scène est sélectionnée et affichée, habituellement avec des informations de texte, côte à côte avec des images représentatives d'autres segments. L'utilisateur dispose alors à la fois d'un résumé visuel de toutes les scènes et d'un accès visuel direct aux scènes individuelles. Chaque image représentative de la mosaïque peut être considérée comme une "icône". La sélection de l'icône par un dispositif de pointage (habituellement un curseur commandé par une souris) provoque la reproduction de la séquence ou sous-séquence vidéo associée. Des exemples de dispositions de mosaïques sont la disposition bidimensionnelle et une longue bande unidimensionnelle. Dans le premier cas, l'utilisateur parcourt les icônes de gauche à droite, ligne par ligne, alors que, dans le second cas, il doit déplacer la bande sur l'écran.

Les mosaïques numériques sont habituellement créées par un éditeur vidéo qui observe la séquence vidéo, segmente les données en scènes individuelles et place chaque scène, avec un commentaire descriptif, sur la mosaïque. De manière bien connue d'après la littérature technique, de nombreuses étapes de ce traitement peuvent être automatisées. Par exemple, différentes techniques de détection automatique de changement de scènes sont décrites dans les documents suivants :

"A Real-time neural approach to scene cut detection" de Ardizzone et al, IS&T/SPIE - Storage & Retrieval for Image and Video Databases IV, San Jose, Ca.,

"Digital Video Segmentation" de Hampapur et al, ACM Multimedia '94 Proceedings, ACM Press,

"Extraction of News Articles based on Scene Cut Detection using DCT Clustering" de Ariki et al, Conférence Internationale concernant le Traitement d'Images, septembre 1996, Lausanne, Suisse,

"Automatic partitioning of full-motion video" de HongJiang Zhang et al, Multimedia Systems (Springer-Verlag, 1993), 1, pages 10-28, et

EP-A-0 590 759.

Divers procédés de détection et de poursuite automatique de personnes et d'objets dans des séquences vidéo sont envisagées dans les documents :

"Modelling, Analysis and Visualisation of Nonrigid Object Motion", de T.S. Huang, Proc. of International Conf. on Pattern Recognition, Vol. 1, pages 361-364, Atlantic City, NJ, Juin 1990, et

Segmentation of People in Motion" de Shio et al, Proc. IEEE, vol. 79, pages 325. 332, 1991.

Des techniques de détection automatique de différents types de prises de vues par une caméra sont décrits dans les documents suivants :

"Global zoom/pan estimation and compensation for video compression" de Tse et al, Proc. ICASSP, Vol. 4, pages 2725-2728, mai 1991, et

"Differential estimation of the global motion parameters zoom and pan" de M. Hoetter, Signal Processing, Vol. 16, pages 249-265, 1989.

Dans le cas des mosaïques numériques, la qualité dynamique de la séquence vidéo est aussi souvent perdue ou cachée. Une certaine impression de mouvement propre à la séquence vidéo peut être préservée par sélection de quelques images destinées à représenter chaque scène, de préférence des images qui montrent le mouvement qui se produit dans cette scène. Cependant, les interfaces vidéo du type à mosaïque restent peu commodes à utiliser étant donné que de multiples actions sont nécessaires de la part de l'utilisateur lorsqu'il veut observer les données et avoir accès à celles-ci.

On a déjà essayé de créer une présentation visuelle unique qui représente à la fois le contenu de vues individuelles constituant une séquence vidéo et qui préserve le contexte, c'est-à-dire l'évolution au cours du temps des informations vidéo de présentation visuelle.

Une telle solution crée une "trace" constituée d'une seule image ayant des présentations visuelles superposées provenant de différentes images de la séquence vidéo, ces présentations visuelles étant décalées les unes par rapport aux autres du fait du mouvement qui se produit entre les différentes images dont les présentations visuelles sont tirées. Ainsi, dans le cas d'une séquence vidéo représentant par exemple un sprinter en action, la "trace" correspondante comprend plusieurs présentations visuelles du sprinter (qui se recouvrent probablement) séparées dans la direction de déplacement du sprinter. Une autre solution de ce type crée une présentation visuelle composite appelée "seuil saillant" représentative de la séquence vidéo, comme indiqué dans le document "Salient

Video Stills : Content and Context Preserved" de Teodosio et al., Proc. ACM Multimedia 93, Californie, 1-6 août 1993, pages 39 à 47.

Une autre solution de ce type général comprend la création d'une "icône vidéo" comme décrit dans les documents "Developing Power Tools for Video Indexing and Retrieval" de Zhang et al., SPIE, volume 2185, pages 140 à 149, et "Video Representation tools using a unified object and perspective based approach" des inventeurs, IS&T/SPIE Conference on Storage and Perusal for Image and Video Databases, San Jose, Californie, février 1995.

Dans une "icône vidéo", comme indiqué sur la figure 1a), la scène est représentée par un certain nombre d'images choisies dans la séquence et qui sont affichées comme si elles étaient empilées les unes derrière les autres dans une direction Z et vues en perspective. En d'autres termes, chaque image individuelle est représentée par un plan, et les plans sont placés les uns derrière les autres avec un léger décalage. Habituellement, la première image de l'empilement est affichée dans sa totalité alors que les images placées au-dessous sont partiellement cachées par les images se trouvant plus en avant dans le temps. L'enveloppe de l'empilement d'images a une forme parallélépipédique. L'utilisation d'un certain nombre d'images, même si elles sont partiellement cachées, donne à l'utilisateur une vue plus complète de la scène et, ainsi, une meilleure compréhension visuelle. En outre, pour certaines icônes de ce type, l'utilisateur peut avoir directement accès à une image quelconque représentée dans l'icône.

Deux types spéciaux d'icônes vidéo ont été proposés, les icônes vidéo "à base objet", et les icônes vidéo contenant une représentation du mouvement de la caméra. Dans une icône vidéo "à base objet" telle que représentée sur la figure 1b), des objets intéressants sont isolés dans les images individuelles et, pour certaines des images empilées au moins, la seule information de présentation visuelle contenue dans l'icône vidéo est l'information de présentation visuelle qui correspond à l'objet choisi. Dans cette icône vidéo, certaines des images individuelles au moins sont représentées comme si elles étaient transparentes, sauf dans les régions contenant l'objet choisi. Les icônes vidéo contenant une indication du mouvement de la caméra, comme indiqué dans l'exemple de la figure 1c), peuvent avoir une enveloppe de forme sinueuse correspondant au cas d'un mouvement de balayage latéral de la caméra.

Les icônes vidéo décrites précédemment présentent à l'utilisateur des informations qui concernent le contenu de la totalité de la séquence vidéo, et elles sont utilisées comme outil de sélection permettant à l'utilisateur d'avoir accès aux

images de la séquence vidéo, d'une manière qui ne correspond pas à l'ordre habituel. En d'autres termes, ces icônes permettent un accès non séquentiel à la séquence vidéo. Néanmoins, les façons dont l'utilisateur peut interagir avec les informations de séquence vidéo sont strictement limitées. L'utilisateur peut sélectionner des images à lire de manière non séquentielle, mais il n'a que des moyens faibles ou nuls d'obtention d'un niveau plus profond d'information concernant la séquence vidéo dans son ensemble, sauf s'il observe la reproduction de l'ensemble de la séquence.

La présente invention concerne un nouveau type d'interface d'accès à des informations vidéo qui permet à l'utilisateur d'avoir accès à des informations concernant une séquence vidéo d'une manière extrêmement souple. En particulier, des interfaces vidéo interactives selon la présente invention permettent à un utilisateur d'obtenir des niveaux plus profonds d'information concernant une séquence vidéo associée, ce à des positions de la séquence qui sont désignées par l'utilisateur comme étant intéressantes.

La présente invention concerne une interface permettant d'accéder à des informations concernant une séquence vidéo associée, l'interface comprenant :

– des informations délimitant une présentation visuelle tridimensionnelle racine, cette représentation racine étant constituée de plusieurs images fondamentales choisies parmi la séquence vidéo et/ou plusieurs parties d'images vidéo correspondant à des objets choisis représentés dans la séquence vidéo, les directions x et y de la représentation racine correspondant aux directions x et y des images vidéo, et la direction z de la présentation visuelle racine correspondant à l'axe des temps, si bien que les images fondamentales sont espacées les unes des autres dans la direction z de la représentation racine par des distances qui correspondent au temps séparant les images vidéo respectives,

– un dispositif d'affichage de vues de la présentation visuelle racine,

– un moyen destiné à désigner une position d'observation par rapport à la présentation visuelle racine, et

– un moyen destiné à calculer des données de présentation visuelle qui représentent la présentation visuelle racine tridimensionnelle observée depuis la position désignée d'observation et à transmettre les données calculées de présentation visuelle au dispositif d'affichage.

Selon la présente invention, des interfaces personnalisées peuvent être créées pour des séquences vidéo. Ces interfaces personalisées comprennent une présentation visuelle "racine" qui peut être affichée et qui représente directement le

contenu et le contexte des informations de présentation visuelle de la séquence vidéo. La présentation visuelle "racine" peut être manipulée, soit de façon automatique soit par l'utilisateur, pour l'affichage d'autres informations de présentation visuelle, par désignation d'une position d'observation par rapport à celle-ci, la représentation de la présentation visuelle affichée étant alors modifiée d'après les changements de la position désignée d'observation. Dans un mode de réalisation préféré de la présente invention, la représentation de la présentation visuelle affichée change selon la position désignée d'observation comme si la présentation visuelle racine était un objet tridimensionnel. Dans de tels modes de réalisation préférés, au fur et à mesure que la position désignée d'observation change, les données nécessaires afin de générer la représentation affichée (quasi-objet) de la présentation visuelle racine sont calculées afin de fournir une vue en perspective correcte compte tenu de l'angle de vue, de la distance entre la position d'observation et le quasi-objet affiché, et compte tenu du fait que la position d'observation se trouve au-dessus ou en-dessous du quasi-objet affiché.

Dans un mode de réalisation limité, l'invention prévoit des interfaces non interactives permettant d'accéder à des informations concernant une séquence vidéo, dans lesquelles les informations de présentation visuelle racine sont livrées accompagnées d'un script associé qui définit un procédé destiné à provoquer l'affichage automatique d'une série de représentations différentes de la présentation visuelle racine et la réalisation d'une suite de manipulations de la représentation affichée, aucune manipulation par l'utilisateur n'étant permise. Cependant, afin de bénéficier pleinement de tous les avantages fournis par la présente invention, il est préférable de mettre en oeuvre des interfaces interactives selon lesquelles la position d'observation de la présentation visuelle racine est désignée par l'utilisateur, comme suit. Lorsque l'utilisateur a d'abord accès à l'interface, une présentation visuelle affichée lui est d'abord présentée et qui correspond à la présentation visuelle racine observée d'un point de vue particulier (qui peut être un point de vue de référence). Lorsque l'utilisateur désigne des angles différents d'observation, la présentation visuelle affichée représente la présentation visuelle racine avec différentes perspectives. Lorsque l'utilisateur désigne des positions d'observation à des distances plus ou moins grandes de la présentation visuelle racine, la représentation affichée augmente ou réduit la taille et, de préférence, la résolution des informations affichées et donne accès le cas échéant aux données d'autres images vidéo.

Les interfaces interactives personnalisées selon la présente invention donnent des présentations visuelles affichées représentant des séquences vidéo respectives associées qui, d'une certaine manière, peuvent être considérées comme un environnement dans lequel on peut naviguer ou comme un objet qui peut être manipulé. Cet environnement ou cet objet est une entité quasi-tridimensionnelle. Les dimensions x et y de l'environnement-objet correspondent aux véritables dimensions dans l'espace (et aux directions x et y des images vidéo associées), alors que la dimension z de l'environnement-objet correspond à l'axe des temps. Ces interfaces peuvent être considérées comme constituant un perfectionnement des "icônes vidéo" décrites précédemment, qui sont alors rendues interactives et peuvent être manipulées par l'utilisateur.

Grâce aux interfaces réalisées selon la présente invention, l'utilisateur peut sélectionner des informations spatiales et temporelles dans une séquence vidéo, pour y avoir accès, par désignation d'une position d'observation vis-à-vis d'une icône vidéo représentant la séquence vidéo. Des "directions d'observation" obliques arbitrairement choisies peuvent être utilisées, si bien que l'utilisateur a simultanément accès à des informations correspondant à des parties d'un certain nombre d'images différentes de la séquence vidéo. Lorsque la position d'observation de l'utilisateur par rapport à l'icône vidéo varie, l'étendue d'une image donnée qui est visible par l'utilisateur et le nombre et la sélection des images qu'il peut observer changent de manière correspondante.

Comme indiqué précédemment, les interfaces vidéo interactives selon la présente invention utilisent une présentation visuelle racine qui comprend plusieurs images fondamentales disposées afin qu'elles forment un objet quasi-tridimensionnel. Il est préférable que les positions relatives des images fondamentales soient telles qu'elles indiquent visuellement un certain mouvement existant dans la séquence vidéo. Par exemple, si la séquence vidéo correspond à un travelling qui parcourt un couloir et tourne à un angle, l'enveloppe de l'ensemble des images fondamentales n'a pas une configuration parallélépipédique, mais forme au contraire un "tube" de section rectangulaire, courbé d'une manière qui correspond au déplacement de la caméra lors de la prise de la séquence vidéo.

Dans des modes de réalisation préférés d'interfaces vidéo interactives selon la présente invention, les images vidéo fondamentales constituant la présentation visuelle racine sont choisies en fonction de l'amplitude du mouvement ou du changement dans la séquence. Par exemple, dans le cas d'une séquence vidéo correspondant à un travelling dans lequel des informations du fond varient, il est

préférable que des images fondamentales successives comportent des informations de fond qui se recouvrent par exemple d'un facteur de 50 %.

Dans certains modes de réalisation de la présente invention, la présentation visuelle racine correspond à une "icône vidéo à base objet" ou, en d'autres termes, certaines des images fondamentales incorporées à la présentation visuelle racine ne sont pas contenues dans celle-ci en totalité, seules des parties correspondant à des objets choisis y étant incorporées. En outre, ou bien dans une variante, certaines images fondamentales peuvent être incorporées en totalité à la présentation visuelle racine, mais elles peuvent comprendre des "objets actifs", c'est-à-dire des représentations d'objets qui peuvent être sélectionnées par l'utilisateur. A la suite de la sélection de ces "objets actifs" par l'utilisateur, les images fondamentales correspondantes (et, éventuellement, les images supplémentaires) sont alors affichées comme si elles étaient transparentes dans leur totalité sauf dans une ou plusieurs parties dans lesquelles l'objet ou les objets sont affichés. La présence de ces objets, qui peuvent être choisis dans la présentation visuelle racine, permet à l'utilisateur d'isoler sélectivement les objets intéressants dans la séquence vidéo et d'obtenir d'un seul regard une impression visuelle de l'aspect et du mouvement des objets au cours de la séquence vidéo.

Les interfaces selon la présente invention permettent à l'utilisateur de sélectionner une partie arbitraire de la séquence vidéo pour la reproduire. L'utilisateur désigne une partie de la séquence vidéo qui l'intéresse en désignant une partie correspondante de la présentation visuelle affichée, faisant partie de l'interface de la séquence vidéo. Cette partie de la séquence vidéo est alors reproduite. L'interface peut générer un ensemble affiché de commande analogue à celui qui est présenté sur un magnétoscope, pour permettre à l'utilisateur de sélectionner différents modes pour cette reproduction, tels que l'avance rapide, le réenroulement, etc.

Dans des modes de réalisation préférés de l'interface selon l'invention, la présentation visuelle affichée faisant partie de l'interface reste visible alors que la partie désignée de la séquence est reproduite ou jouée. Ce résultat peut être obtenu d'un certain nombre de manières : par exemple par utilisation d'un second dispositif d'affichage sur lequel est réalisée la reproduction ou par désignation d'une "fenêtre de reproduction" sur l'écran d'affichage, cette fenêtre étant décalée par rapport à la région de l'écran utilisée par l'interface, ou par tout autre moyen convenable.

Des modes de réalisation préférés d'interface selon l'invention permettent aussi à l'utilisateur de désigner un objet intéressant et de sélectionner un mode de reproduction dans lequel seules les informations de la présentation visuelle qui concernent l'objet choisi soient incluses dans la reproduction. En outre, l'utilisateur peut sélectionner une image unique de la séquence vidéo pour l'afficher sous forme séparée de la présentation visuelle interactive affichée, créée par l'interface.

Dans des modes de réalisation préférés, les interfaces selon la présente invention permettent à l'utilisateur de créer une présentation visuelle affichée qui correspond à une distorsion de la représentation visuelle racine. Plus précisément, la présentation visuelle affichée peut correspondre à la présentation visuelle racine soumise à un "effet d'accordéon" selon lequel la présentation visuelle racine est, par exemple, "ouverte par pivotement", par rotation autour d'un axe de rotation afin que les images vidéo soient "écartées en éventail" à proximité du point d'ouverture, ou elle est modifiée par écartement des images vidéo au niveau d'un point considéré comme étant intéressant.

La présente invention donne à l'utilisateur des interfaces pour des séquences vidéo "à plusieurs fils", c'est-à-dire des séquences vidéo comprenant de nombreux segments relativement courts, qui existent par exemple dans un jeu vidéo lorsque les sélections par l'utilisateur modifient la scène qui est affichée. Les interfaces avec de telles séquences vidéo à plusieurs fils peuvent comprendre des images de différents segments vidéo dans la présentation visuelle racine, si bien que la présentation visuelle racine a une structure présentant des branchements. Dans une variante, une partie ou la totalité des fils différents peut ne pas être visible dans la présentation visuelle racine, mais peut devenir visible à la suite de la manipulation par l'utilisateur. Par exemple, si l'utilisateur exprime un intérêt pour une région particulière de la séquence vidéo, par désignation d'une partie de la présentation visuelle racine affichée à l'aide d'un dispositif de pointage (tel qu'une souris, ou par un contact avec un écran tactile, etc.), si plusieurs fils différents de la séquence commencent dans la zone désignée, les parties de la présentation visuelle correspondant à ces différents fils peuvent être ajoutées à la présentation visuelle affichée.

Dans des modes de réalisation préférés de l'interface selon la présente invention, la présentation visuelle racine de la séquence vidéo concernée est associée à des informations déterminant comment la présentation visuelle affichée correspondante change en fonction des différents types de manipulation par l'utilisateur. Par exemple, ces informations associées peuvent déterminer combien

d'images ou quelles images supplémentaires sont affichées lorsque l'utilisateur déplace la position d'observation en la rapprochant de la présentation visuelle racine. De même, les informations associées peuvent identifier quels objets de la scène sont des "objets actifs" et quelles données visuelles sont affichées pour ces objets actifs lors d'une activation par l'utilisateur.

En outre, il existe différentes possibilités pour la fourniture des composants de l'interface à l'utilisateur final. Dans une application dans laquelle des séquences vidéo sont transmises à un utilisateur par un trajet de télécommunications, par exemple par le réseau Internet, l'utilisateur qui est intéressé par une séquence vidéo particulière peut d'abord charger certains composants seulement de l'interface associée. D'abord, il charge des informations pour la création d'une représentation affichée de la présentation visuelle racine, avec un programme associé d'application (s'il n'a pas déjà un "moteur de gestion d'interface" convenable chargé dans son ordinateur). Le programme d'application téléchargé (ou déjà présent) contient des sous-programmes fondamentaux de changement de la perspective de la présentation visuelle affichée d'après les changements de position d'observation désignés par l'utilisateur. Le programme d'application est aussi adapté à une consultation de toute "information associée" (comme indiqué précédemment) qui fait partie de l'interface et prépare les conditions dans lesquelles la présentation visuelle affichée change à la suite de certaines manipulations prédéterminées par l'utilisateur (par exemple "agrandissement" et "activation de l'objet"). Lorsque l'interface ne contient pas une telle "information associée", le programme d'application utilise les paramètres préréglés par défaut.

La présentation visuelle racine correspond à un ensemble particulier d'images vidéo fondamentales et d'informations désignant les positions relatives de ces images. Les informations de présentation visuelle racine téléchargées vers l'utilisateur peuvent comprendre simplement les données nécessaires à la création d'une vue de référence de la présentation visuelle racine, ou peuvent comprendre les données de présentation visuelle de l'ensemble des images fondamentales (pour permettre le traitement des changements d'angle d'observation de l'utilisateur sans qu'il soit nécessaire de télécharger des informations supplémentaires). Dans le cas où l'utilisateur effectue une manipulation qui nécessite l'affichage d'informations vidéo qui ne sont pas présentes dans la présentation visuelle racine (par exemple lorsqu'il commande un "agrandissement", si bien que des données d'image supplémentaires sont nécessaires), ces informations supplémentaires peuvent être soit préalablement conditionnées et transmises avec les informations de la

présentation visuelle racine, soit téléchargées à partir du site du serveur hôte, le moment venu.

Des possibilités analogues existent dans le cas des interfaces livrées sur disque CD-ROM. En général, la présentation visuelle racine et les autres informations associées sont placées sur le disque CD-ROM en plus de la séquence vidéo complète. Cependant, il faut noter que, pour des raisons d'économie d'espace, les catalogues des séquences vidéo peuvent être réalisés uniquement avec les interfaces, sans les séquences vidéo complètes correspondantes.

En plus de la réalisation des interfaces elles-mêmes, la présente invention concerne aussi un appareil pour la création d'interfaces selon la présente invention. Il peut comprendre des éléments matériels spécialisés ou, de préférence, un système d'ordinateur programmé par un programme d'ordinateur spécialement réalisé.

Diverses étapes dans le procédé de création d'une interface personnalisée selon la présente invention peuvent être automatisées. Par exemple, la sélection des images fondamentales destinées à être incluses dans la "présentation visuelle racine" de l'interface peut être réalisée automatiquement à l'aide d'un algorithme choisi parmi un certain nombre d'algorithmes différents, tels que : sélection d'une image toutes les "n" images, sélection d'une image chaque fois que le mouvement de la caméra a déplacé le fond de "m" %, etc. De même, les positions relatives des images fondamentales dans la présentation visuelle racine peuvent être déterminées automatiquement en fonction du temps séparant ces images et, le cas échéant, d'autres facteurs tels que le mouvement de la caméra. De même, la présence d'objets ou de personnes dans la séquence vidéo peut être détectée automatiquement à l'aide de l'un des algorithmes connus (par exemple ceux qui sont décrits dans les documents précités), et une présentation visuelle racine "orientée objet" peut être créée automatiquement. Ainsi, dans certains cas, l'appareil de création d'interfaces selon la présente invention peut traiter automatiquement des informations de séquence vidéo pour créer une présentation visuelle racine. Ces modes de réalisation comprennent un dispositif destiné à associer à la présentation visuelle racine un ensemble normalisé de sous-programmes destiné à modifier la présentation visuelle affichée en fonction des manipulations par l'utilisateur.

Cependant, il est souvent préférable de concevoir activement les caractéristiques des interfaces interactives selon l'invention d'une manière qui limite ou canalise dans des directions préférées la façon selon laquelle l'utilisateur

final peut interagir avec les informations vidéo. Ceci s'applique en particulier au cas des séquences vidéo constituant des messages publicitaires ou se trouvant dans des logiciels éducatifs.

Ainsi, la présente invention concerne un kit d'outils destiné à être utilisé pour la création d'interfaces personnalisées. Dans des modes de réalisation préférés, le kit d'outils permet à un concepteur d'adapter la configuration et le contenu de la présentation visuelle racine ainsi que de spécifier les objets de la séquence vidéo qui sont des "objets actifs" et de commander le changement, par manipulation par l'utilisateur final, de la présentation visuelle affichée par l'interface. Ainsi, le kit d'outils permet notamment au concepteur de l'interface de déterminer quelles images de la séquence vidéo doivent être utilisées comme images fondamentales dans la présentation visuelle racine et combien d'images supplémentaires sont ajoutées à la présentation visuelle affichée lorsque l'utilisateur désigne une position d'observation proche de la présentation visuelle racine.

D'autres caractéristiques et avantages de la présente invention apparaîtront à la lecture de la description qui suit de modes de réalisation préférés, donnés à titre d'exemple et représentés sur les dessins annexés sur lesquels :

la figure 1 représente divers types d'icônes vidéo, la figure 1a) représentant une icône vidéo ordinaire, la figure 1b) une icône vidéo à base objet, et la figure 1c) une icône vidéo comprenant une représentation d'un mouvement de caméra ;

la figure 2 est un diagramme synoptique indiquant les composants d'une interface interactive dans un premier mode de réalisation de la présente invention ;

la figure 3 est un schéma représentant le contenu d'un fichier de données d'interface (FDI) utilisé dans le premier mode de réalisation de l'invention ;

la figure 4 est un schéma représentant une vue de référence d'une présentation visuelle racine et trois positions d'observation désignées par un utilisateur ;

la figure 5 est une présentation visuelle affichée dans le cas d'une présentation visuelle racine observée à partir des différentes positions d'observation de la figure 4, la figure 5a) étant la présentation visuelle affichée observée depuis la position A, la figure 5b) étant la présentation visuelle affichée observée depuis la position B, et la figure 5c) étant la présentation visuelle affichée observée depuis la position C ;

la figure 6 représente des présentations visuelles affichées mettant en oeuvre des présentations visuelles racines plus complexes selon la présente

invention, la figure 6a) étant dérivée d'une présentation visuelle racine représentant visuellement un mouvement, et la figure 6b) étant dérivée d'une présentation visuelle racine représentant visuellement un effet de zoom ;

la figure 7 représente l'effet de la sélection par un utilisateur d'un objet représenté dans la présentation visuelle affichée, dans un second mode de réalisation d'interface selon la présente invention ;

la figure 8 représente une manipulation par l'utilisateur d'une présentation visuelle racine pour la production d'un "effet d'accordéon" ;

la figure 9 est une présentation visuelle affichée qui correspond à une vue d'une présentation visuelle racine à arborescence représentant un scénario à plusieurs fils ; et

la figure 10 est un ordinogramme illustrant les étapes d'un procédé préféré de conception d'une interface selon la présente invention.

Les composants d'une interface interactive, dans un premier mode de réalisation préféré de la présente invention, sont maintenant décrits en référence à la figure 2. Dans cet exemple, une interface interactive selon l'invention est associée à des séquences vidéo enregistrées sur un disque CD-ROM.

Comme représenté sur la figure 2, un lecteur 1 de disques CD-ROM est raccordé à un système d'ordinateur comprenant une partie 2 à processeur central, un écran 3 d'affichage, et un dispositif de saisie qui peut être manoeuvré par l'utilisateur et qui, dans ce cas, comporte un clavier 4 et une souris 5. Lorsque l'utilisateur veut consulter les séquences vidéo enregistrées sur un disque CD-ROM 7, il place le disque 7 dans le lecteur et active le logiciel d'accès au disque CD-ROM placé dans la partie 2 du processeur central ou dans une unité/mémoire associée.

Dans ce premier mode de réalisation de l'invention, le disque CD-ROM porte sous forme enregistrée non seulement les informations vidéo de la présentation visuelle de la séquence vidéo (avec tout format commode), mais aussi un fichier respectif de données d'interface (FDI_i) 10 pour chaque séquence vidéo, avec un programme 11 d'application d'interface vidéo. Le contenu d'un exemple de fichier de données d'interface est représenté sur la figure 3. Des scripts respectifs 12 sont éventuellement associés aux fichiers de données d'interface. Lorsque les données du disque CD-ROM doivent être lues, le programme 11 d'application d'interface vidéo est utilisé par la partie 2 du processeur central du système d'ordinateur, et le fichier de données d'interface applicable à la séquence vidéo choisie par l'utilisateur est traité pour l'affichage d'une icône vidéo interactive (voir

par exemple les figures 4 et 5) sur l'écran d'affichage 3. L'utilisateur peut alors manipuler l'icône affichée, en utilisant les dispositifs de saisie à clavier ou à souris, pour explorer la séquence vidéo choisie.

On décrit maintenant, en référence aux figures 4 à 9, les types de manipulation de l'icône vidéo interactive qui sont disponibles pour l'utilisateur.

La figure 4 représente une simple icône vidéo interactive selon l'invention. En particulier, cette icône est représentée sur l'écran d'affichage par un ensemble d'images superposées disposées dans une enveloppe ayant la configuration d'un parallélépipède régulier. Chacune des images superposées correspond à une image vidéo choisie dans la séquence vidéo, mais ces images sont décalées les unes par rapport aux autres. On peut considérer que la présentation visuelle affichée correspond à un parallélépipède rectangle observé d'une position particulière (en haut et à droite dans cet exemple). Ce parallélépipède rectangle est une construction théorique formée de l'ensemble des images vidéo choisies disposées de manière à ce que leurs axes respectifs x et y correspondent aux axes x et y du parallélépipède rectangle et que l'axe z du parallélépipède rectangle corresponde à l'axe des temps. Ainsi, dans le parallélépipède rectangle de la construction théorique, les images choisies sont séparées dans la direction z en fonction des intervales temporels les séparant dans la séquence vidéo.

Lorsque l'utilisateur veut explorer la séquence vidéo à l'aide de l'icône vidéo interactive affichée sur l'écran d'affichage, l'une des opérations fondamentales qu'il doit effectuer est la désignation d'une position sur l'écran comme position d'observation par rapport à l'affichage affichée (par exemple par un "clic" de la souris de l'ordinateur). Sur la figure 4, trois positions d'observation ainsi désignées portent les références littérales A, B et C. A la suite de cette opération de l'utilisateur, la présentation visuelle affichée est changée et prend la forme indiquée sur la figure 5 : les figures 5a), b) et c) correspondent aux positions d'observation A, B et C respectivement de la figure 4.

On note que la présentation visuelle affichée change pour l'obtention d'une vue en perspective du parallélépipède rectangle théorique, observé suivant la position d'observation désignée par l'utilisateur.

Le parallélépipède rectangle précité est un cas particulier d'une "présentation visuelle racine" selon la présente invention. Cette "présentation visuelle racine" est dérivée de la séquence vidéo et transmet des informations concernant à la fois le contenu de la présentation visuelle du sous-ensemble choisi d'images (appelées dans la suite "images fondamentales") et la "position" relative

de ces informations de présentation visuelle dans le temps aussi bien que dans l'espace. Il faut noter que la "présentation visuelle racine" est définie par les informations contenues dans le fichier de données d'interface. La définition de la présentation visuelle racine spécifie quelles images vidéo sont des "images fondamentales" (par exemple par mémorisation des numéros correspondants d'images), et spécifie les positions des images fondamentales les unes par rapport aux autres dans la présentation visuelle racine.

La partie 2 du processeur central du système d'ordinateur calcule les données de présentation visuelle nécessaires pour la création de la présentation visuelle affichée à partir d'une part de la définition de la présentation visuelle racine contenue dans le fichier convenable de données d'interface, d'autre part des données de présentation visuelle des images fondamentales (et, le cas échéant, d'images supplémentaires), et enfin de la position d'observation désignée par l'utilisateur, à l'aide des techniques classiques de traçage de rayons. Les données nécessaires pour la création de la présentation visuelle affichée sont chargées dans un circuit de mémoire tampon vidéo et affichées sur l'écran d'affichage.

Selon la présente invention, il est préférable, lorsque l'utilisateur désigne une position d'observation proche de l'icône vidéo interactive, que les informations de présentation visuelle de la zone intéressante soient enrichies. Ce résultat est obtenu par incorporation, à la présentation visuelle affichée, de données de présentation visuelle relatives à des images vidéo supplémentaires, s'ajoutant aux images vidéo fondamentales. Ce cas est illustré sur la figure 5b) sur laquelle les images fondamentales BF5 et BF6 sont affichées avec les images supplémentaires AF1 et AF2. Lorsque la position d'observation désignée par l'utilisateur se rapproche de plus en plus de la présentation visuelle affichée, le programme d'application de l'interface vidéo provoque l'addition d'images supplémentaires, très rapprochées, à la présentation visuelle affichée. Finalement, les images vidéo successives de la séquence vidéo peuvent être incluses dans la présentation affichée.

Comme on peut le voir dans la figure 5 b), des informations visuelles correspondant à des parties de la présentation visuelle racine éloignées de la zone intéressante peuvent être omises de la représentation affichée "en gros plan".

De préférence, le fichier de données d'interface contient des données qui spécifient comment la sélection doit être réalisée sur des images supplémentaires qui doivent être ajoutées lorsque l'utilisateur "se rapproche" de la présentation visuelle affichée. De préférence, ces données déterminent des règles gouvernant le

choix de combien d'images supplémentaires et lesquelles doivent être utilisées pour enrichir la présentation visuelle affichée lors d'un changement de la position désignée d'observation. Ces règles peuvent, par exemple, définir une relation mathématique entre le nombre d'images affichées et la "distance" séparant la position désignée d'observation du quasi-objet affiché. Dans des modes de réalisation préférés de l'invention, le nombre d'images ajoutées à l'affichage, lorsque la position désignée d'observation s'approche du quasi-objet affiché, dépend du degré de mouvement ou de changement contenu dans la séquence vidéo à l'endroit considéré.

L'exemple représenté sur la figure 4 est une simplification dans laquelle la présentation visuelle affichée correspond à une présentation visuelle racine ayant une simple forme de parallélépipède rectangle. Cependant, selon la présente invention, la présentation visuelle racine peut avoir diverses formes différentes.

Par exemple, les positions relatives des images fondamentales peuvent être sélectionnées de manière à ce que l'enveloppe de la présentation visuelle racine ait une configuration qui reflète le mouvement de la séquence vidéo correspondante (par exemple le mouvement de la caméra, lors de prise de vues de poursuite ou analogue, ou le mouvement des objets représentés dans la séquence), comme indiqué par l'icône interactive correspondante représentée sur la figure 6 a). De même, les dimensions des images fondamentales de la présentation visuelle racine peuvent subir un changement d'échelle pour la présentation visuelle d'un effet de zoom qui se produit dans la séquence vidéo, comme indiqué par l'icône active correspondante représentée sur la figure 6b).

On note que l'icône interactive représentée sur la figure 6b) comprend certaines images dont une partie seulement de leurs informations visuelles a été affichée. Ceci correspond au cas où un objet particulièrement intéressant a été sélectionné. Cette sélection d'un objet peut être réalisée de diverses manières. Le cas échéant, la présentation visuelle racine peut être désignée de telle manière que, au lieu de contenir des images fondamentales complètes, seules des parties d'images qui représentent un objet particulier soient incorporées. Cette disposition implique qu'une sélection ait été réalisée au moment de la conception de la partie de présentation visuelle racine de l'interface, concernant les objets qui sont intéressants. Le concepteur peut décider, en outre ou à titre de variante, que la présentation visuelle racine contienne des images fondamentales complètes, mais aussi que certains objets représentés dans la séquence vidéo puissent être "choisis"

ou "extraits" à la demande de l'utilisateur. Cette propriété est maintenant décrite en référence à la figure 7.

La figure 7a) représente une vue initiale présentée à un utilisateur lorsqu'il consulte l'interface pour une séquence vidéo particulière choisie. Dans cette séquence, deux personnes marchent l'une vers l'autre et leurs trajets se recoupent. Le concepteur de l'interface a décidé que les deux personnes étaient des objets qui peuvent présenter un intérêt pour l'utilisateur final. Ainsi, il a incorporé au fichier de données d'interface des informations désignant ces objets comme pouvant "être extraits". Ces informations de désignation peuvent correspondre à des informations de plages de coordonnées x, y, identifiant la position de l'objet dans chaque image vidéo (ou dans chaque image d'un sous-ensemble d'images).

Lorsque l'utilisateur exprime un intérêt pour l'un des deux objets, par exemple par désignation d'une position de l'écran correspondant à l'un des objets (par exemple par un "clic" sur la personne gauche à l'aide du bouton droit de la souris), le programme d'application d'interface règle la présentation visuelle affichée de manière à ce que les parties étrangères des images affichées disparaissent de l'affichage et ne laissent qu'une représentation des deux personnes et de leurs mouvements, tel que représenté sur la figure 7b). Ainsi, les objets intéressants sont "extraits" du milieu environnant. Les parties transparentes ou "manquantes" des images affichées peuvent être rétablies dans la présentation visuelle affichée à la demande de l'utilisateur (par exemple par un autre "clic" du bouton de la souris).

Il faut noter que, selon la présente invention, des interfaces peuvent être réalisées de manière à ce que des objets particuliers "qui peuvent être extraits" puissent être extraits simultanément avec certains ou la totalité des autres objets qui peuvent être extraits, ou puissent être extraits individuellement. Des interfaces sophistiquées selon la présente invention peuvent inclure des sous-programmes d'extraction d'objets permettant à l'utilisateur de sélectionner à son gré des objets visibles dans la représentation affichée de la présentation visuelle racine, pour l'extraction de ceux-ci. Ainsi, par exemple, l'utilisateur peut utiliser un dispositif de pointage pour dessiner un cadre entourant un objet visible dans la vue affichée de la présentation visuelle racine et le programme d'application fournit alors des sous-programmes permettant d'identifier dans les autres images fondamentales de la présentation visuelle racine (et, éventuellement, dans les images supplémentaires) l'objet désigné, afin de produire un affichage de l'objet choisi comme si celui-ci se trouvait dans des images transparentes.

Il peut être souhaitable de permettre à l'utilisateur d'obtenir une vue rapprochée d'une partie déterminée d'une icône vidéo interactive d'une manière qui ne correspond pas strictement à une vue en perspective de la région concernée. Les modes de réalisation préférés d'interfaces selon l'invention donnent donc un effet dit "d'accordéon" comme indiqué sur la figure 8. Lorsque l'utilisateur manipule l'icône par un effet "d'accordéon" en un point particulier, les images fondamentales au voisinage de la région intéressante sont écartées afin qu'elles permettent une meilleure observation par l'utilisateur. En outre, la fonction d'affichage d'images supplémentaires pour augmenter la résolution dans le temps est de préférence inhibée pendant l'effet d'accordéon.

Dans le cas de séquences vidéo à plusieurs fils, par exemple qui existent dans les jeux vidéo sur ordinateur et les logiciels éducatifs et qui mettent en oeuvre des sous-séquences vidéo parallèles qui sont obtenues en alternance par sélection par l'utilisateur, ces séquences peuvent aussi être les sujets d'interfaces selon la présente invention. Dans ce cas, le concepteur de l'interface peut choisir d'incorporer à la présentation visuelle racine de l'interface des images provenant de sous-séquences vidéo parallèles différentes pour donner à l'utilisateur une idée des différentes lignes d'action qui sont disponibles dans la séquence vidéo. La figure 9 représente une icône vidéo interactive dérivée d'un exemple simple d'une telle présentation visuelle racine.

Dans une variante ou en outre, le concepteur peut créer des présentations visuelles racines secondaires pour les sous-séquences respectives, ces présentations visuelles racines secondaires étant utilisées pour la création de la présentation visuelle affichée uniquement lorsque l'utilisateur désigne une position d'observation proche de l'image vidéo à laquelle commence la sous-séquence. Dans le cas des interfaces des jeux sur ordinateur ou des logiciels éducatifs, ce choix est logique car c'est à l'endroit auquel la sous-séquence vidéo part de la séquence principale que les choix effectués par l'utilisateur, au cours de la partie changent le scénario exécuté.

Une autre manipulation qu'il est avantageux d'incorporer aux interfaces selon l'invention est l'ensemble traditionnel des commandes de magnétoscope qui sont affichées et qui permettent à l'utilisateur de reproduire la séquence vidéo à laquelle est associée l'icône vidéo affichée. En outre, l'utilisateur peut sélectionner des parties de reproduction ou images de la séquence, par exemple par un "clic" du bouton de la souris sur les images intéressantes qui sont affichées dans l'icône vidéo interactive. La reproduction vidéo peut être réalisée sur un écran séparé

d'affichage ou sur une fenêtre délimitée sur l'écran d'affichage qui représente déjà l'icône vidéo.

Comme indiqué précédemment, une séquence vidéo particulière peut être associée à un fichier de données d'interface et à un script. Le script est un sous-programme défini par le concepteur de l'interface, guidant l'utilisateur lors de l'utilisation de l'interface. Le script peut comprendre par exemple un sous-programme destiné à provoquer une démonstration automatique des différentes manipulations possibles du quasi-objet affiché. L'utilisateur peut modifier le déroulement du procédé définit dans le script de manière classique, par exemple en commandant une pause, un ralentissement, etc. Si désiré, le script peut inclure des informations sous forme de texte, de son ou des informations graphiques supplémentaires susceptibles d'être reproduites en liaison avec la vue affichée de la présentation visuelle racine, soit de façon automatique soit en réponse à des opérations effectuées par l'utilisateur final.

La figure 10 est un ordinogramme illustrant des exemples d'étapes au cours de la conception d'une interface selon la présente invention, dans le cas où un concepteur agit. Il faut comprendre que les interfaces selon la présente invention peuvent aussi être créées de manière totalement automatique. Il faut noter que les sélections effectuées par le concepteur ont un effet notable sur le contenu du fichier de données d'interface. Il faut aussi comprendre que toutes les étapes représentées sur la figure 10 ne sont pas obligatoirement nécessaires ; par exemple, les étapes concernant la création des présentations visuelles racines secondaires peuvent être omises dans le cas d'une séquence vidéo qui n'a pas plusieurs fils.

De façon similaire, il peut s'avérer utile d'inclure dans le procédé de conception d'interface certaines étapes supplémentaires qui ne sont pas indiquées sur la figure 10. Par exemple, il est souvent souhaitable d'inclure dans le fichier de données d'interface (comme il est illustré sur la figure 3) des informations concernant le mouvement de la caméra, le montage, etc... faisant partie de la séquence vidéo. Au cours de l'utilisation de l'interface, ces informations permettent, par exemple, d'incorporer à la présentation visuelle affichée, des images vidéo supplémentaires disposées afin de représenter visuellement les mouvements de la caméra. Au cours du processus de conception d'interface ces informations concernant les caractéristiques de la séquence vidéo peuvent être déterminées soit de façon automatique (en utilisant les méthodes connues de détection d'interruption et des procédés analogues) soit selon les choix effectués par le concepteur.

La présente invention prévoit un kit d'outils destiné à une utilisation par des concepteurs souhaitant concevoir une interface interactive selon la présente invention pour une séquence vidéo particulière. De préférence, ce kit d'outils est réalisé en forme de programme pour ordinateur. Ce kit d'outils fournit au concepteur des menus et des consignes affichés afin de l'amener à suivre un processus comportant des étapes telles que celle de la séquence typique illustrée sur la figure 10.

Tout d'abord, le concepteur désigne la séquence vidéo pour laquelle il souhaite créer une interface, par exemple en indiquant le nom du fichier stocké comportant les informations de séquences vidéo concernées. De préférence, le kit d'outils procède à la lecture de ces informations de séquence vidéo afin d'afficher la séquence vidéo dans une fenêtre sur l'écran de l'ordinateur et permettre ainsi au concepteur de se référer à la séquence vidéo au cours de la conception de l'interface. Selon de tels modes de réalisation préférés du kit d'outils, le concepteur peut effectuer sa sélection d'images fondamentales et/ou d'objets à inclure dans la présentation visuelle de racine, des objets capables d'être extraits, et autres, en parcourant lentement la séquence vidéo et, par exemple, utilisant une souris pour positionner un curseur sur des images ou des parties d'image qui présentent un intérêt pour le concepteur. Le kit d'outils enregistre les numéros d'image (et éventuellement les positions x,y des parties d'image) des images/parties d'image indiquées par le concepteur et crée une liaison entre ces informations de localisation et le paramètre approprié en cours de définition. De préférence, à la fin de la conception de l'interface, le concepteur se voit présenter d'une vue affichée de la présentation visuelle de racine qu'il peut manipuler afin de déterminer si des changements du fichier de données d'interface sont nécessaires.

Des versions différentes du programme d'application peuvent être associées au fichier de données d'interface (et au script, le cas échéant) selon les fonctions d'interface qui devraient être mises en oeuvre. Ainsi donc, dans le cas où le fichier des données d'interface n'est pas accompagné d'un script, il n'est pas nécessaire d'inclure dans le programme d'application un sous-programme permettant de gérer le déroulement d'un script. De même, dans le cas où le fichier de données d'interface ne permet pas à l'utilisateur de commander un effet d'accordéon, il n'est pas nécessaire d'inclure dans le programme d'application des sous-programmes permettant de calculer les données d'affichage demandées pour de tels effets. Bien entendu, si le concepteur d'interface sait que l'utilisateur final dispose d'un programme d'application convenable pour gérer des interfaces selon la présente

invention, il peut décider de ne pas accompagner le fichier de données d'interfaces d'un programme d'application ou, sinon, d'associer au fichier de données d'interface des informations identifiant une version du programme d'application qui convient à la gestion de cette interface précise.

La présente invention a été décrite précédemment en référence à des séquences vidéo conservées sur un disque CD-ROM. Il faut noter que la présente invention peut être mise en oeuvre dans de nombreuses autres applications. Le contenu du fichier de données d'interface et les composants de l'interface qui sont présents à l'emplacement de l'utilisateur final peuvent varier suivant l'application.

Par exemple, dans une application dans laquelle une séquence vidéo est fournie par un serveur de réseau, l'utilisateur peut d'abord télécharger, par sa connexion de télécommunications, uniquement le fichier de données d'interface qui s'applique à la séquence. Si l'utilisateur n'a pas déjà le logiciel convenant à la manipulation de l'icône vidéo interactive, il télécharge aussi le programme correspondant d'application. Lorsque l'utilisateur manipule l'icône vidéo interactive, des informations supplémentaires de présentation visuelle dont il peut avoir besoin et qui n'ont pas été déjà téléchargées peuvent être téléchargées de manière dynamique le moment venu.

Un autre exemple est le cas d'un catalogue sur disque CD-ROM, ne comprenant que des interfaces à la place des séquences vidéo associées, afin que l'espace soit économisé. Dans ce cas, au lieu de contenir un pointeur sur les informations de présentation visuelle des images fondamentales de la présentation visuelle racine, le fichier de données d'interface contient les informations de la présentation visuelle elles-mêmes. Certaines informations supplémentaires de la présentation visuelle peuvent aussi être présentes.

Bien que les modes de réalisation préférés précités de la présente invention indiquent certaines combinaisons caractéristiques, il faut noter que la présente invention n'est pas limitée aux détails de ces exemples particuliers. Premièrement, comme le traitement des présentations visuelles est réalisé sur des données de présentation visuelle sous forme numérique, il faut noter que, dans le cas où la séquence vidéo est formée de données de type analogique, un convertisseur analogique-numérique ou analogue est utilisé afin qu'il donne des données de présentation visuelle sous une forme convenant au traitement. Il faut noter aussi que la présente invention s'applique également à la création d'interfaces permettant d'interagir avec les séquences vidéo dans lesquelles les données vidéo sont présentes en forme comprimée, ou chiffrée, etc. Deuxièmement, les références à la

saisie par un utilisateur ou à la sélection par un utilisateur recouvrent aussi l'utilisation d'un dispositif quelconque de saisie qui peut être manoeuvré par un utilisateur et qui comprend, sans aucune limitation, un clavier, une souris (ou un autre dispositif de pointage), un écran ou panneau tactile, des dispositifs à saisie par un gant, des détecteurs des mouvements de l'oeil, des dispositifs commandés par la voix, etc. Troisièmement, les références à des "affichages" recouvrent l'utilisation de nombreux dispositifs différents, tels que, à titre non limitatif, les écrans de moniteurs classiques, les dispositifs d'affichage à cristaux liquides, etc.

En outre, afin d'aider la compréhension, la présentation de l'invention donnée ci-dessus décrit des interfaces selon la présente invention dans lesquelles les présentations visuelles racine respectives possèdent chacune une seule caractéristique intéressante, telle que, la création d'une représentation visuelle d'un mouvement, la création d'une représentation visuelle d'un effet de zoom, l'existence d'une structure arborescente, etc. Il faut noter que la présentation visuelle racine peut comporter plusieurs de ces caractéristiques à la fois, selon le choix du concepteur. De même, des effets spéciaux tels qu'un effet d'accordéon, etc., ont été décrits isolément. Il convient de noter qu'on peut concevoir les interfaces selon la présente invention afin qu'elles permettent une combinaison quelconque d'effets spéciaux.

REVENDICATIONS

1. Interface destinée à une séquence vidéo associée, l'interface comprenant :

– des informations déterminant une présentation visuelle racine tridimensionnelle, la présentation visuelle racine étant constituée de plusieurs images fondamentales choisies dans la séquence vidéo, et/ou plusieurs parties d'images vidéo correspondant à des objets choisis représentés dans la séquence vidéo, les directions x et y de la présentation visuelle racine correspondant aux directions x et y des images vidéo et la direction z de la présentation visuelle racine correspondant à l'axe des temps, si bien que les images fondamentales sont séparées les unes des autres dans la direction z de la présentation visuelle racine par des distances correspondant au temps séparant les images vidéo respectives,

– un dispositif d'affichage de vues de la présentation visuelle racine,

– un moyen destiné à désigner une position d'observation par rapport à la présentation visuelle racine, et

– un moyen destiné à calculer des données de présentation visuelle qui représentent une présentation visuelle racine tridimensionnelle observée depuis la position désignée d'observation, et à transmettre les données calculées de présentation visuelle au dispositif d'affichage.

2. Interface interactive selon la revendication 1, dans laquelle le moyen de désignation consiste en un moyen qui peut être commandé par l'utilisateur et qui est destiné à désigner la position d'observation par rapport à une vue de la présentation visuelle racine.

3. Interface interactive selon la revendication 1 ou 2, dans laquelle le moyen destiné à calculer des données de présentation visuelle à afficher est destiné à incorporer aux données calculées de présentation visuelle de sortie, en fonction de la position d'observation désigné, des données de présentation visuelle qui correspondent à des parties des images fondamentales qui ne sont pas visibles dans la vue de référence de la présentation visuelle racine.

4. Interface selon la revendication 1, 2 ou 3, dans laquelle le moyen de calcul des données de présentation visuelle à afficher est destiné à incorporer, aux données calculées de présentation visuelle, des données de présentation visuelle d'images de la séquence vidéo en plus des images fondamentales, selon la distance

entre la position d'observation désignée par l'utilisateur et la présentation visuelle racine.

5. Interface interactive selon la revendication 4, dans laquelle le moyen de calcul des données de présentation visuelle à afficher est destiné à sélectionner des images supplémentaires choisies d'après des critères spécifiés dans des informations supplémentaires mémorisées en association avec la définition de la présentation visuelle racine, pour être utilisées lors du calcul de la présentation visuelle à afficher.

6. Interface selon l'une quelconque des revendications précédentes, dans laquelle le moyen de calcul des données de présentation visuelle à afficher est destiné à calculer des données de présentation visuelle de sortie correspondant à un nombre d'images différente ou à une présentation visuelle affichée agrandie ou réduite, en fonction de la distance comprise entre la position d'observation désignée par l'utilisateur et la présentation visuelle racine.

7. Interface selon l'une quelconque des revendications précédentes permettant d'accéder à des informations concernant une séquence vidéo comprenant des données de présentation visuelle représentant un ou plusieurs objets choisis, et dans laquelle le moyen de calcul des données de présentation visuelle à afficher est destiné, pour chaque image affichée contenant un objet choisi respectif, à transmettre sélectivement des données de présentation visuelle provoquant l'affichage uniquement des données de présentation visuelle qui correspondent à l'objet ou aux objets choisis, le reste de l'image respective affichée paraissant transparent.

8. Interface selon la revendication 7, dans laquelle un dispositif qui peut être commandé par l'utilisateur est destiné à permettre la sélection d'objets représentés dans la présentation visuelle affichée, et dans laquelle le moyen de calcul des données de présentation visuelle à afficher est destiné à transmettre des données de présentation visuelle provoquant la mise sous forme transparente de parties d'images en fonction de la sélection des objets par l'utilisateur.

9. Interface selon l'une quelconque des revendications précédentes destinée à une séquence vidéo contenant une séquence principale d'images vidéo et au moins

une sous-séquence supplémentaire d'images vidéo, constituant un autre trajet à partir de, ou menant à, une image vidéo particulière de la séquence principale, dans laquelle l'utilisateur peut avoir accès aux informations de présentation visuelle relatives à une autre sous-séquence par désignation d'une position d'observation proche d'un point de la présentation visuelle racine qui correspond à l'image vidéo particulière, le moyen de calcul des données de présentation visuelle à afficher étant destiné à greffer à la vue affichée de la présentation visuelle racine, au point de branchement, une présentation visuelle racine secondaire représentant l'autre sous-séquence.

10. Interface interactive selon la revendication 9, dans laquelle la commande du moyen désignant la position d'observation permet à l'utilisateur de naviguer parmi les présentations visuelles racines et les présentations visuelles racines secondaires correspondant à différentes sous-séquences possibles contenues dans la séquence vidéo.

11. Appareil de création d'une interface pour une séquence vidéo selon l'une quelconque des revendications 1 à 10, l'appareil comprenant :

- un moyen permettant d'accéder à des informations de présentation visuelle sous forme numérique représentant une séquence vidéo,

- un moyen de création d'une présentation visuelle racine représentant la séquence vidéo, ce moyen de création de la présentation visuelle racine comprenant :

d'une part un moyen de sélection d'un sous-ensemble d'images provenant de la séquence vidéo ou de parties du sous-ensemble correspondant à des objets représentés dans la séquence vidéo afin qu'elles soient utilisées comme images fondamentales de la présentation visuelle racine, et

d'autre part un moyen de réglage des positions relatives des images fondamentales dans la présentation visuelle racine, et

- un moyen d'association, à la présentation visuelle racine, de sous-programmes destinés à changer la vue affichée de la présentation visuelle racine en fonction de la position d'observation désignée par rapport à la présentation visuelle racine.

12. Appareil selon la revendication 11, comprenant en outre un moyen d'identification d'objets représentés par les informations de présentation visuelle de

la séquence vidéo et de désignation d'objets qui peuvent être sélectionnés par un utilisateur final.

13. Appareil selon la revendication 11 ou 12, dans lequel le moyen de réglage des positions relatives des images fondamentales dans la présentation visuelle racine comporte un moyen donnant accès aux informations mémorisées représentant plusieurs gabarits et un dispositif de saisie d'informations de sélection désignant l'un des gabarits mémorisés.

14. Appareil selon la revendication 11, 12 ou 13, dans lequel le moyen de réglage des positions relatives des images fondamentales dans la présentation visuelle racine comporte un moyen de détection d'un mouvement dans la séquence vidéo et un moyen destiné à placer des images fondamentales dans la présentation visuelle racine à des positions relatives qui donnent une présentation visuelle du mouvement.

15. Appareil selon la revendication 14, dans lequel le moyen destiné à placer les images fondamentales dans la présentation visuelle racine est destiné à effectuer un changement progressif des dimensions des images fondamentales dans la présentation visuelle racine pour représenter visuellement une opération d'agrandissement ou de rétrécissement.

16. Appareil selon l'une des revendications 11 à 15, dans lequel le moyen de sélection d'un sous-ensemble d'images de la séquence vidéo, destinées à être utilisées comme images fondamentales de la présentation visuelle racine, est destiné à la sélection d'images en fonction de la vitesse de variation des informations du fond de la présentation visuelle.

17. Appareil selon l'une quelconque des revendications 11 à 16, comprenant un moyen de saisie de paramètres fixant des contraintes au changement possible de la vue affichée de la présentation visuelle racine en fonction d'une position d'observation désignée par un utilisateur, les paramètres de contraintes étant prises en compte par les sous-programmes associés par le moyen d'association à la présentation visuelle racine.

18. Appareil selon la revendication 16, dans lequel le moyen de saisie des paramètres de contraintes est destiné à saisir des données identifiant la vitesse à laquelle les images supplémentaires doivent être incorporées à une vue affichée de la présentation visuelle racine lorsqu'une position d'observation désignée par un utilisateur se rapproche de la présentation visuelle racine.

19. Appareil selon l'une des revendications 11 à 18, comprenant un moyen de création de présentations visuelles racines secondaires correspondant à des sous-séquences supplémentaires d'images vidéo formant d'autres trajets allant de, ou vers, une image vidéo particulière de la séquence vidéo principale.

FIGURE 1 a)

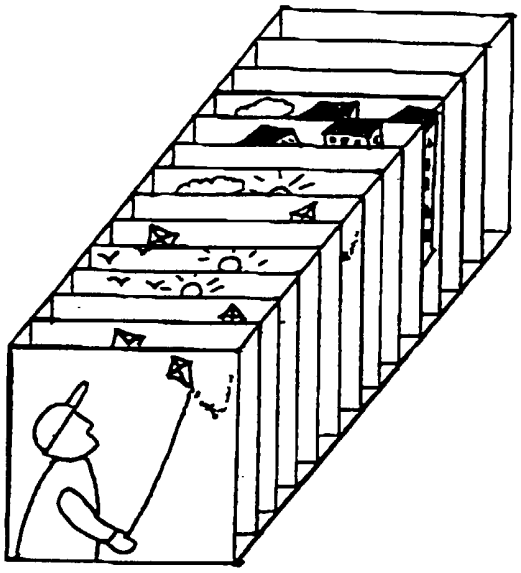


FIGURE 1 b)

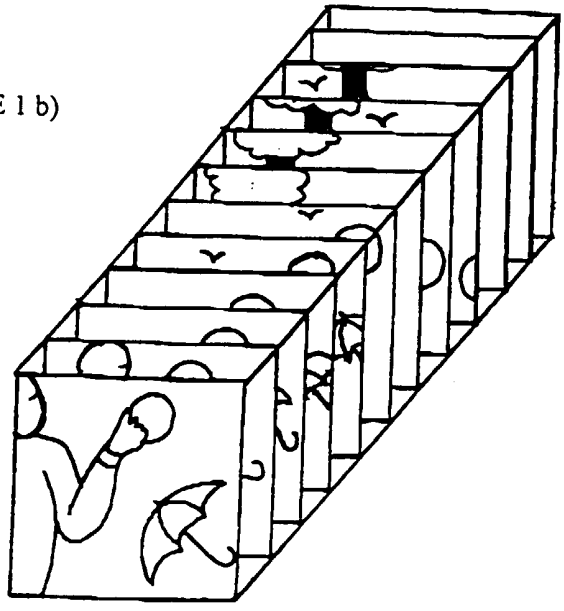


FIGURE 1 c)

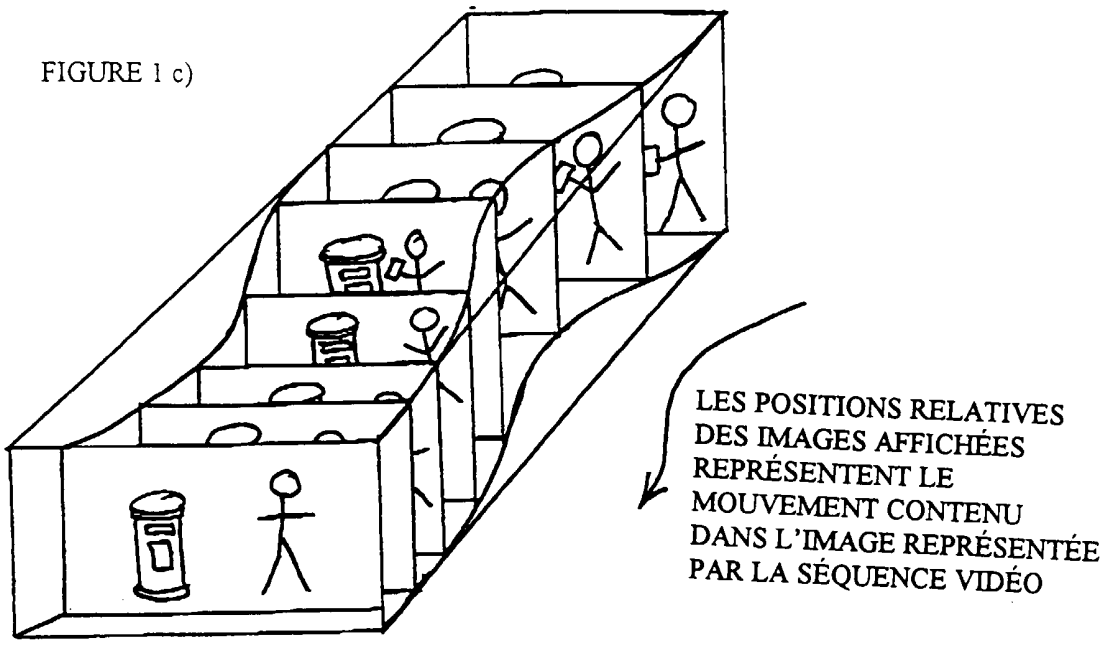


FIGURE 4

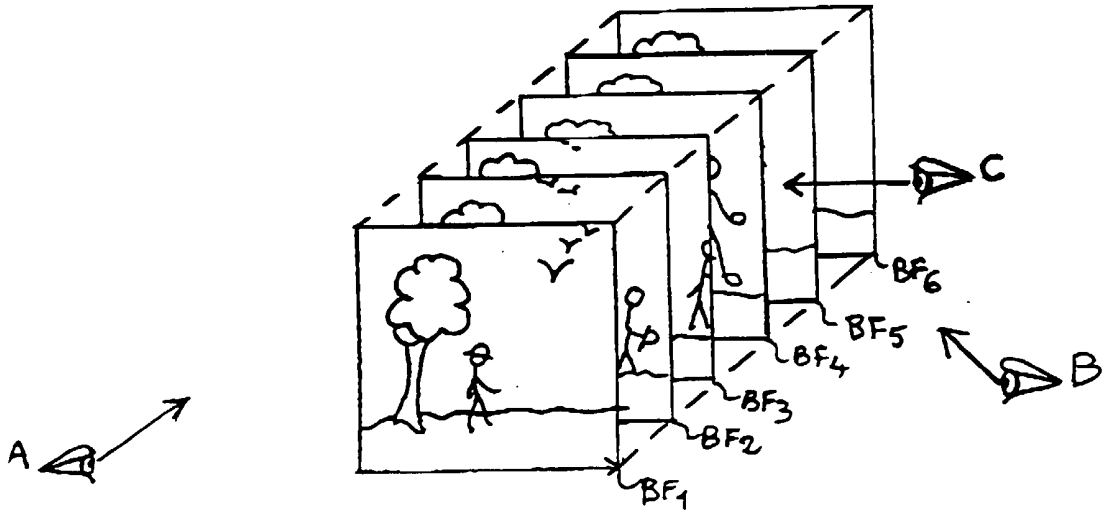


FIGURE 2

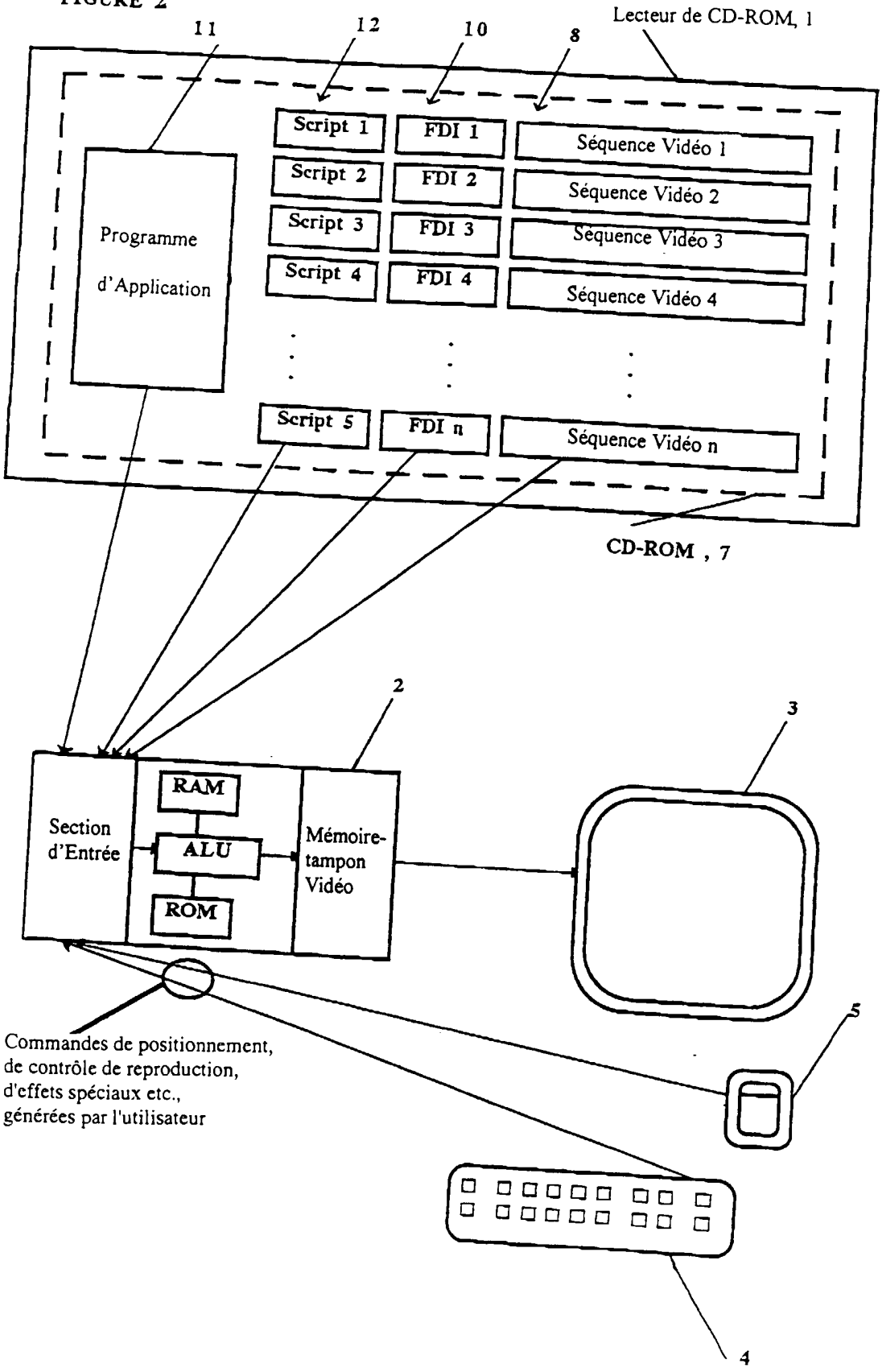

11
12
10
8
Lecteur de CD-ROM, 1
Programme d'Application
Script 1
FDI 1
Séquence Vidéo 1
Script 2
FDI 2
Séquence Vidéo 2
Script 3
FDI 3
Séquence Vidéo 3
Script 4
FDI 4
Séquence Vidéo 4
Script 5
FDI n
Séquence Vidéo n
CD-ROM , 7
2
3
Section d'Entrée
RAM
ALU
ROM
Mémoire-tampon Vidéo
Commandes de positionnement, de contrôle de reproduction, d'effets spéciaux etc., générées par l'utilisateur
5
4

FIGURE 3

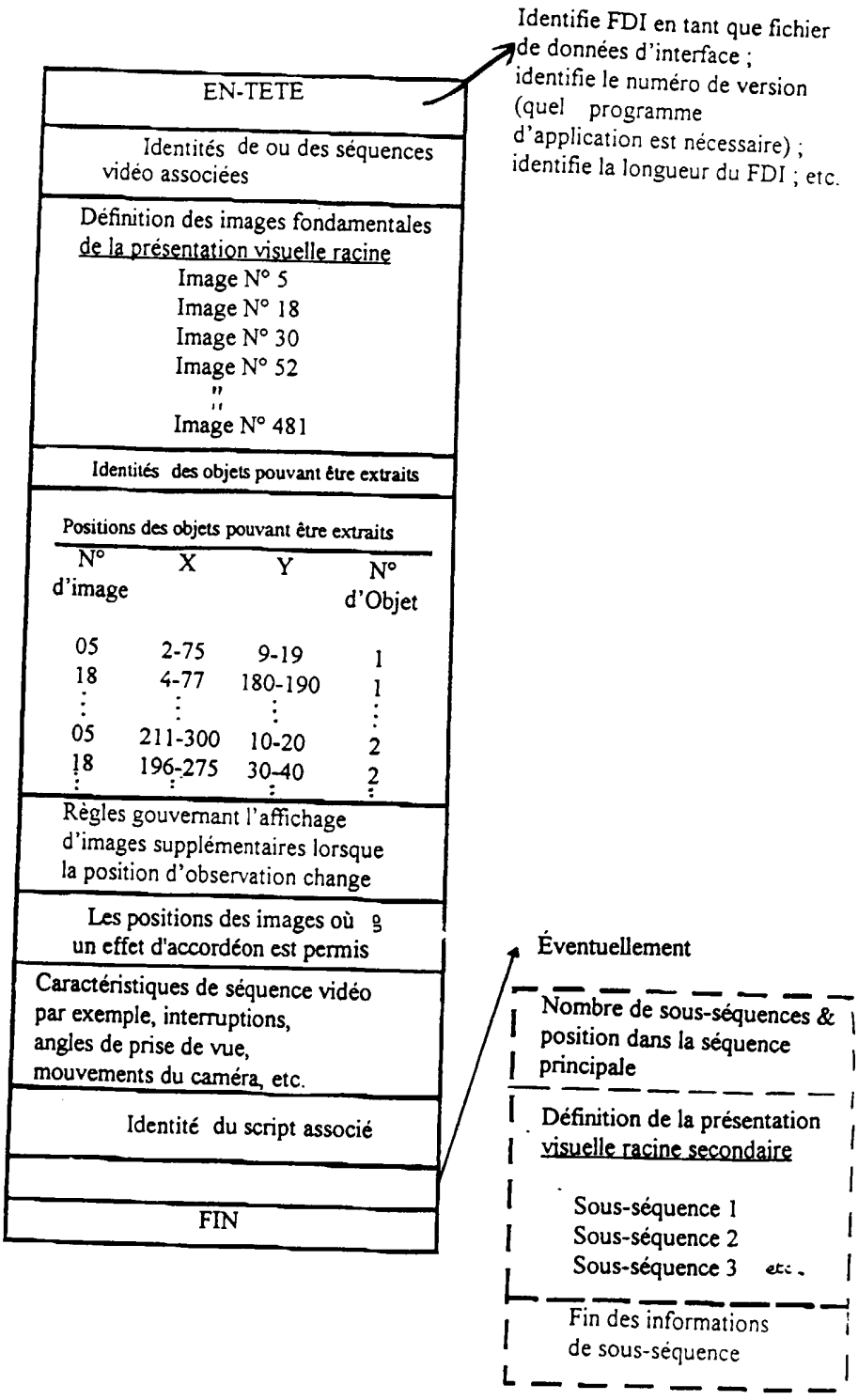

Identifie FDI en tant que fichier de données d'interface ; identifie le numéro de version (quel programme d'application est nécessaire) ; identifie la longueur du FDI ; etc.
EN-TETE
Identités de ou des séquences vidéo associées
Définition des images fondamentales de la présentation visuelle racine
Image N° 5
Image N° 18
Image N° 30
Image N° 52
"
Image N° 481
Identités des objets pouvant être extraits
Positions des objets pouvant être extraits
N° d'image
X
Y
N° d'Objet
05 2-75 9-19 1
18 4-77 180-190 1
05 211-300 10-20 2
18 196-275 30-40 2
Règles gouvernant l'affichage d'images supplémentaires lorsque la position d'observation change
Les positions des images où un effet d'accordéon est permis
Caractéristiques de séquence vidéo par exemple, interruptions, angles de prise de vue, mouvements du caméra, etc.
Identité du script associé
FIN
Éventuellement
Nombre de sous-séquences & position dans la séquence principale
Définition de la présentation visuelle racine secondaire
Sous-séquence 1
Sous-séquence 2
Sous-séquence 3 etc.
Fin des informations de sous-séquence

FIGURE 5 a)

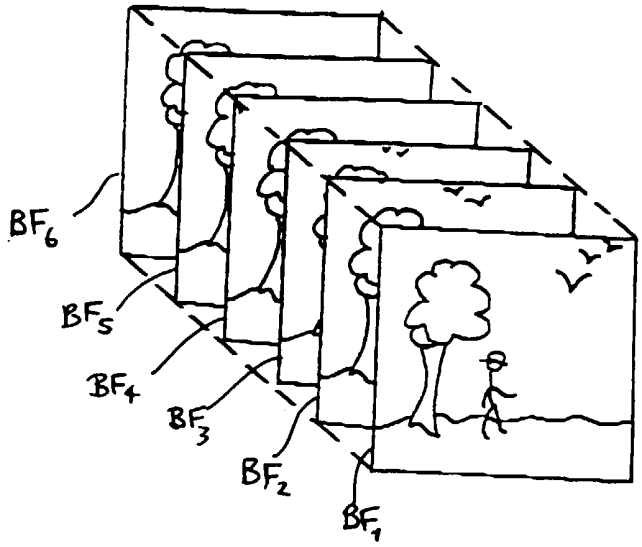


FIGURE 5 b)

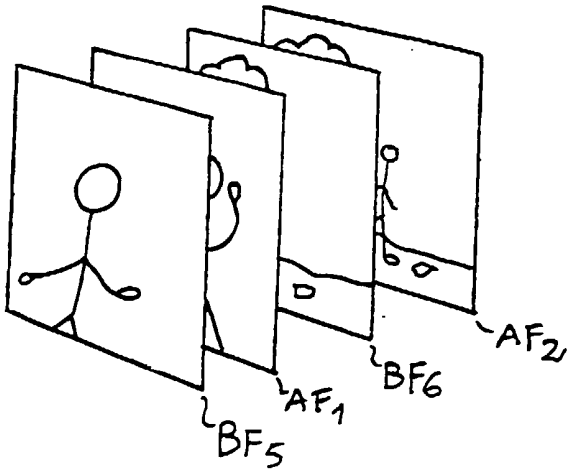


FIGURE 5c)

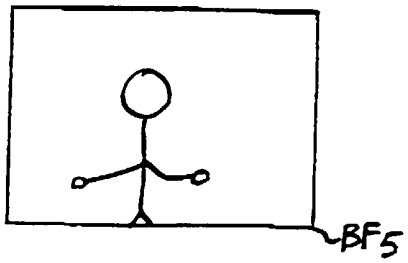
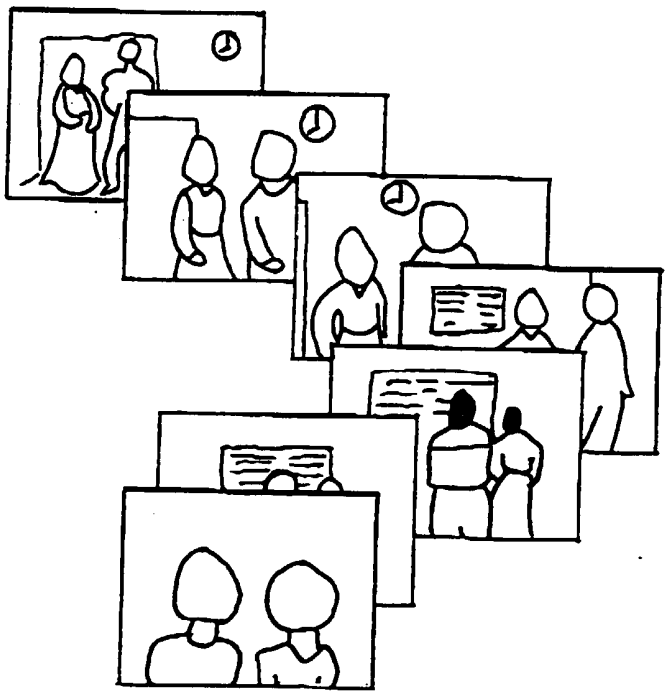


FIGURE 6 a)



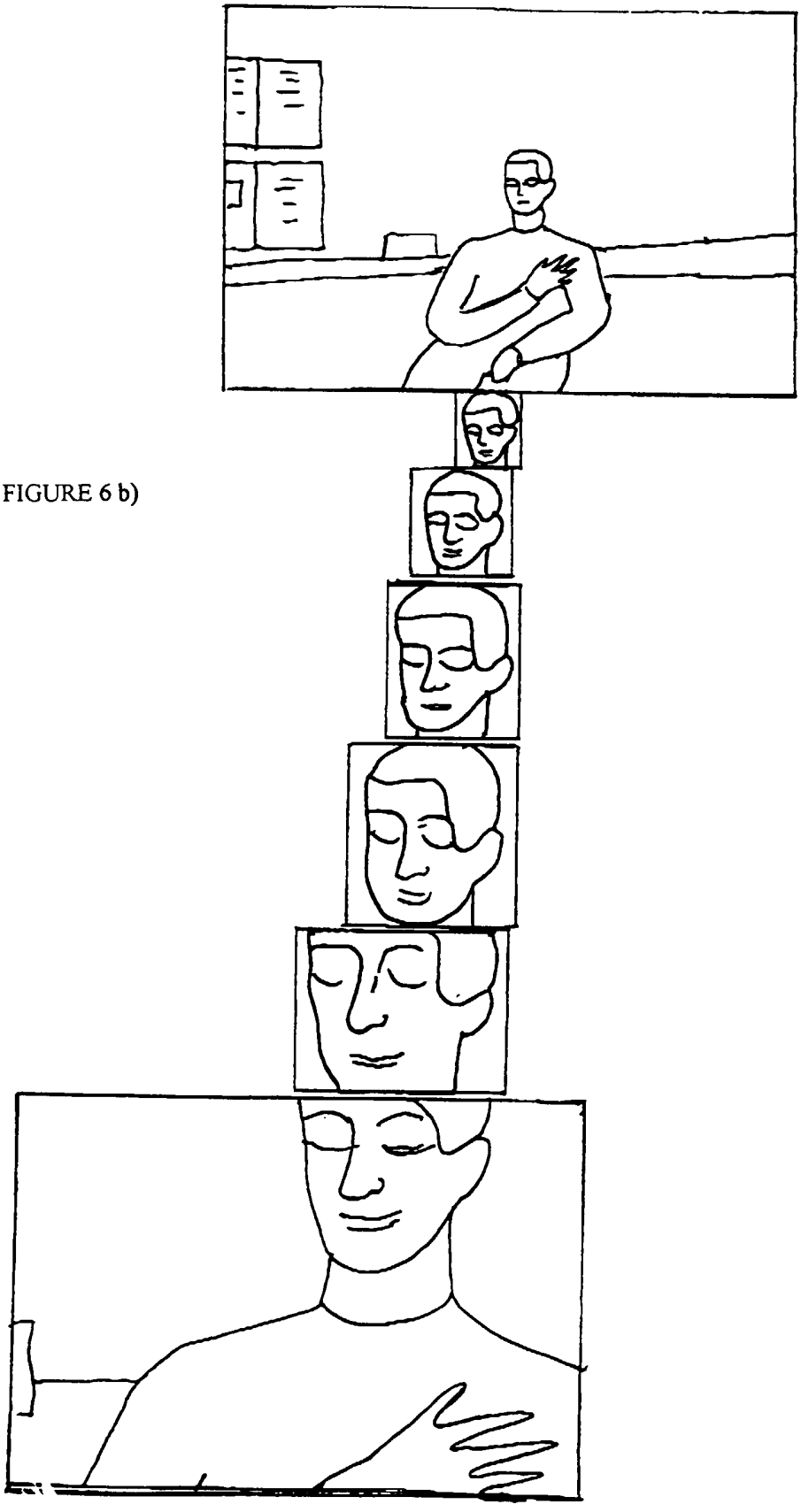


FIGURE 6 b)

FIGURE 7 a)

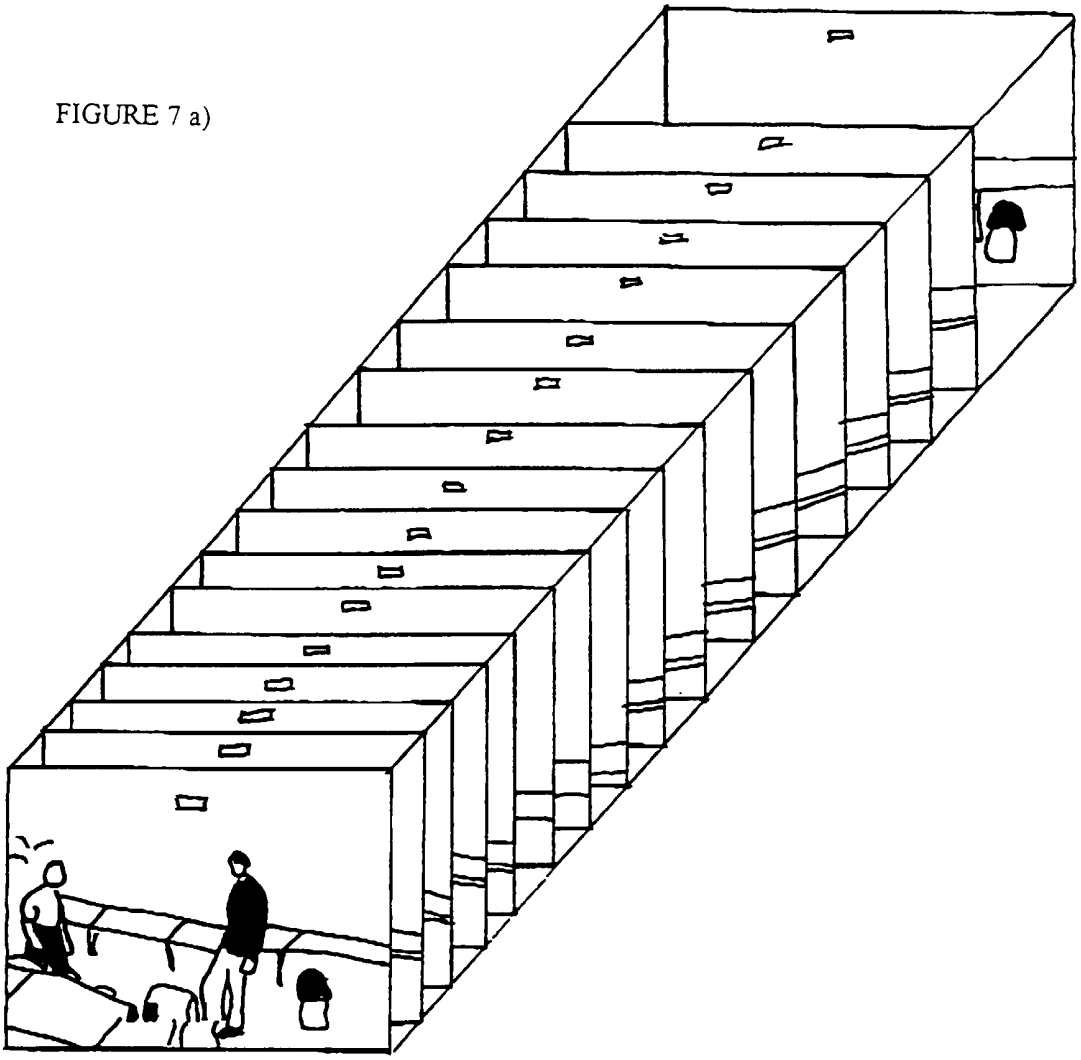


FIGURE 7 b)

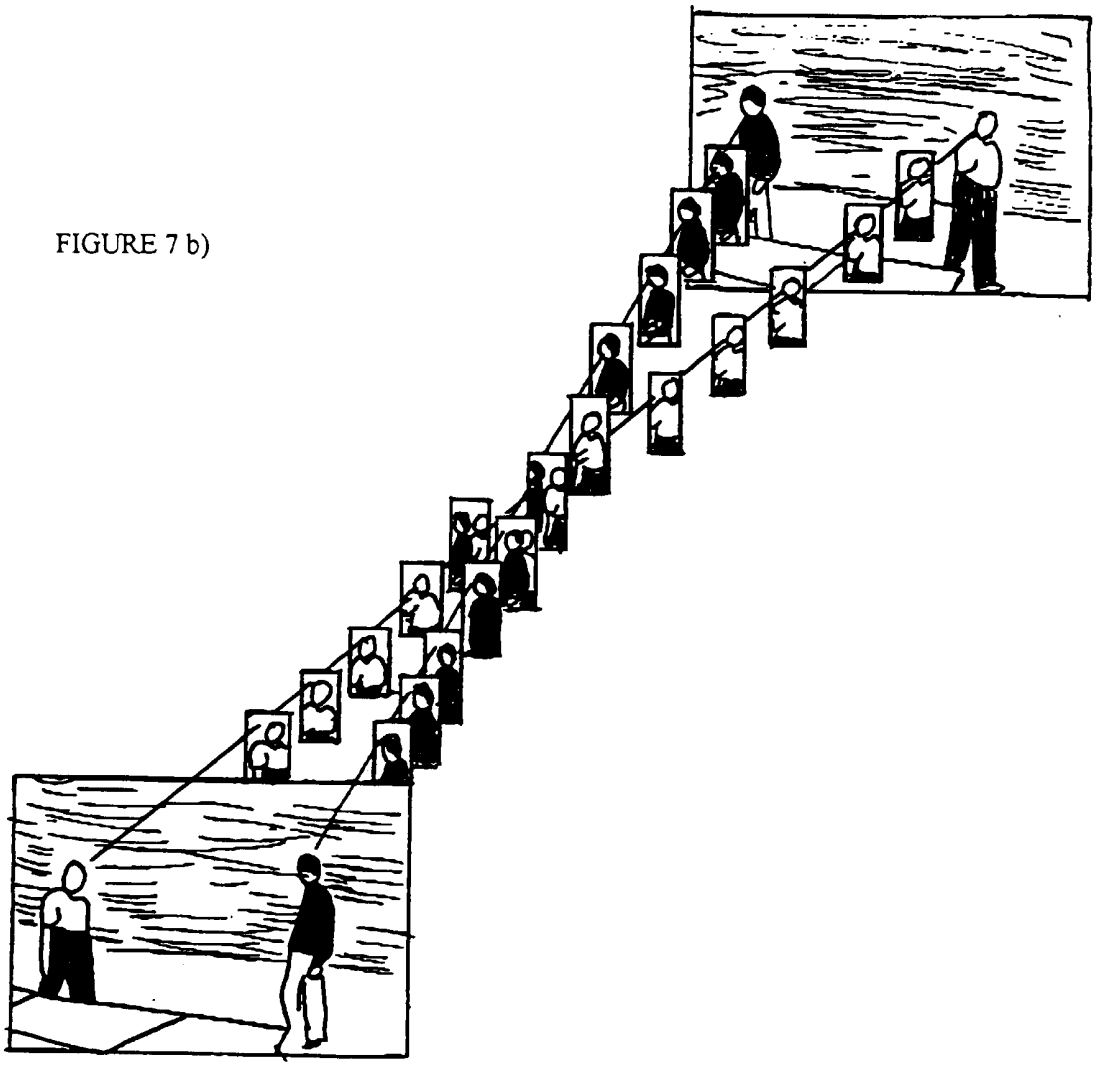


FIGURE 8

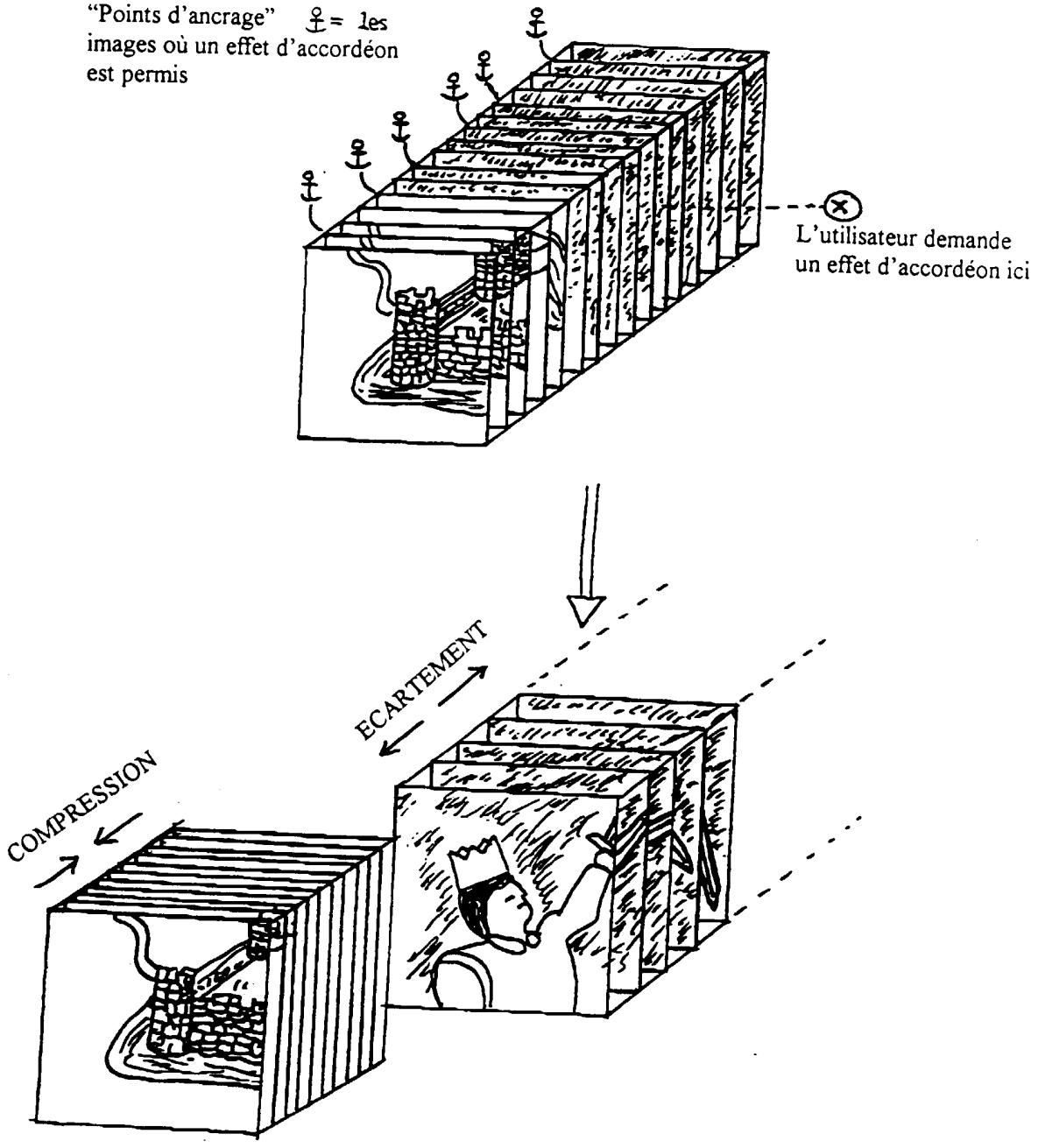

"Points d'ancrage" ⚓ = les images où un effet d'accordéon est permis
L'utilisateur demande un effet d'accordéon ici
ECARTEMENT
COMPRESSION

FIGURE 9

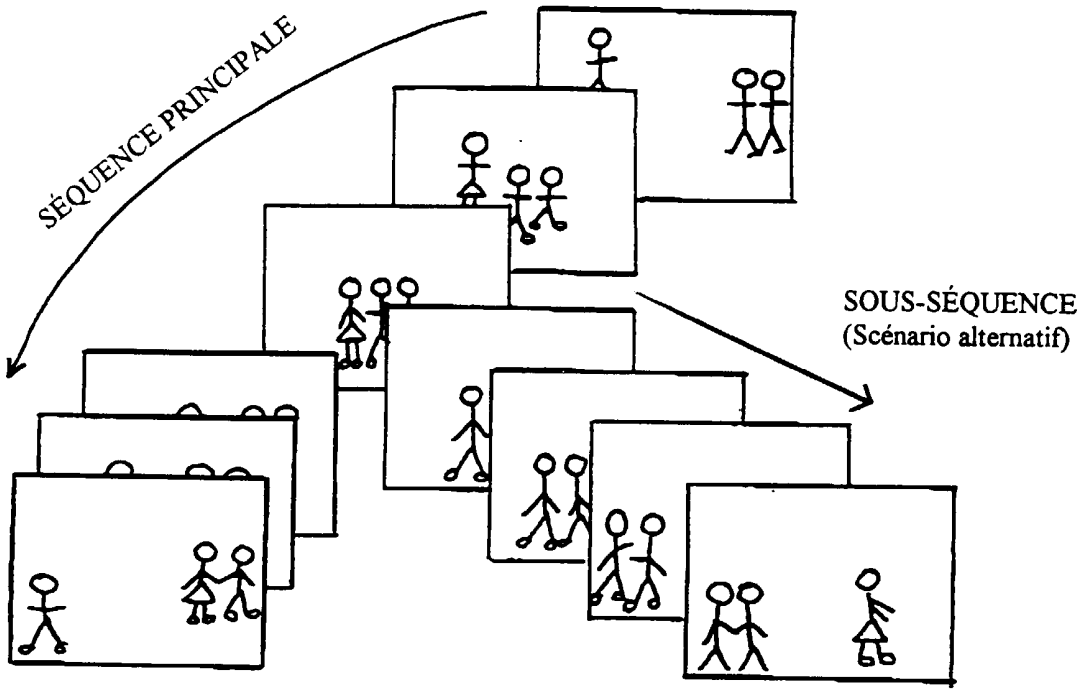
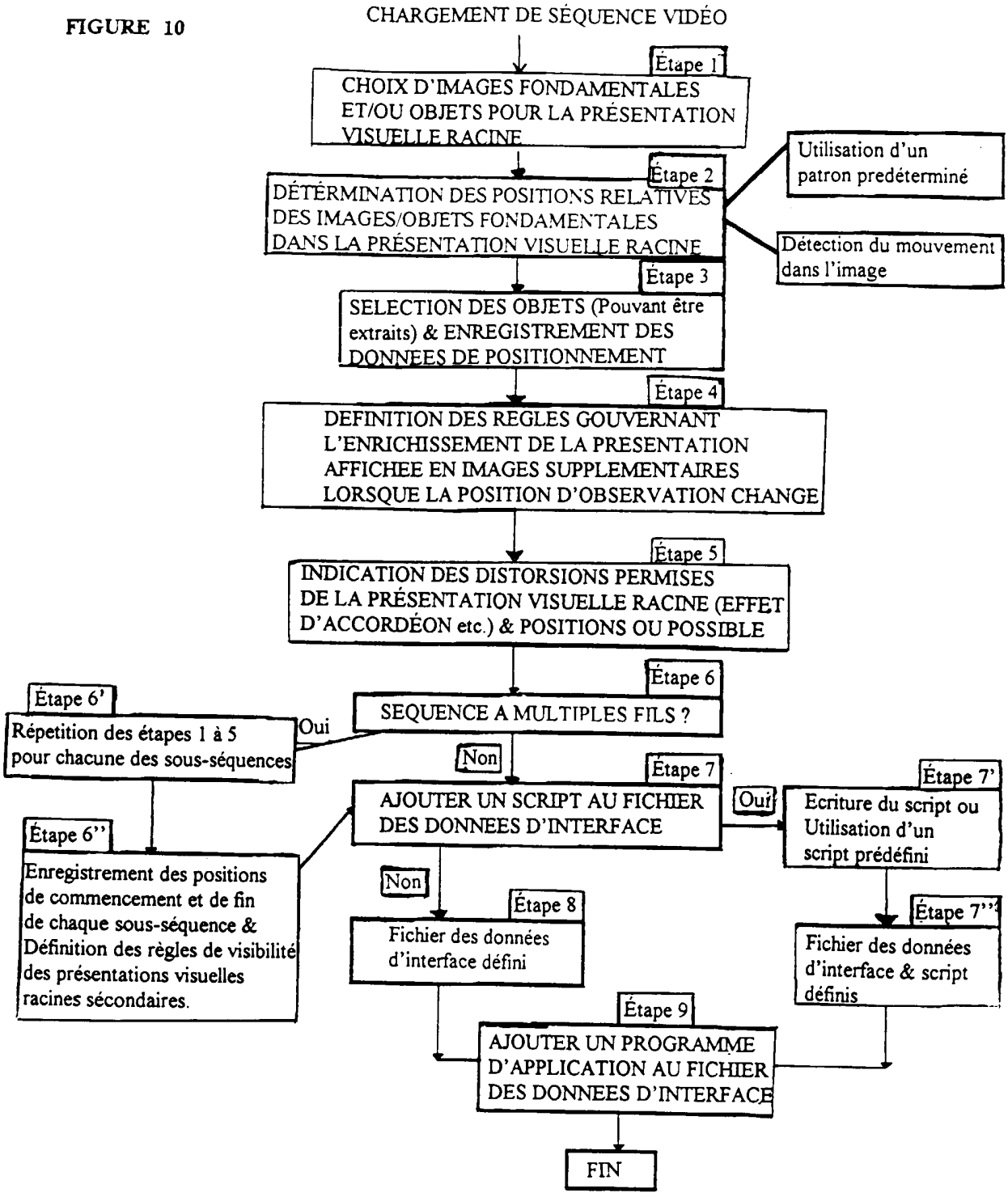


FIGURE 10


CHARGEMENT DE SÉQUENCE VIDÉO
Étape 1
CHOIX D'IMAGES FONDAMENTALES ET/OU OBJETS POUR LA PRÉSENTATION VISUELLE RACINE
Étape 2
DÉTERMINATION DES POSITIONS RELATIVES DES IMAGES/OBJETS FONDAMENTALES DANS LA PRÉSENTATION VISUELLE RACINE
Utilisation d'un patron predéterminé
Détection du mouvement dans l'image
Étape 3
SELECTION DES OBJETS (Pouvant être extraits) & ENREGISTREMENT DES DONNEES DE POSITIONNEMENT
Étape 4
DEFINITION DES REGLES GOUVERNANT L'ENRICHISSEMENT DE LA PRESENTATION AFFICHEE EN IMAGES SUPPLEMENTAIRES LORSQUE LA POSITION D'OBSERVATION CHANGE
Étape 5
INDICATION DES DISTORSIONS PERMISES DE LA PRÉSENTATION VISUELLE RACINE (EFFET D'ACCORDÉON etc.) & POSITIONS OU POSSIBLE
Étape 6
SEQUENCE A MULTIPLES FILS ?
Oui
Étape 6'
Répetition des étapes 1 à 5 pour chacune des sous-séquences
Non
Étape 7
AJOUTER UN SCRIPT AU FICHIER DES DONNEES D'INTERFACE
Oui
Étape 7'
Ecriture du script ou Utilisation d'un script prédéfini
Étape 6''
Enregistrement des positions de commencement et de fin de chaque sous-séquence & Définition des règles de visibilité des présentations visuelles racines sécondaires.
Non
Étape 8
Fichier des données d'interface défini
Étape 7''
Fichier des données d'interface & script définis
Étape 9
AJOUTER UN PROGRAMME D'APPLICATION AU FICHIER DES DONNEES D'INTERFACE
FIN

REPUBLIQUE FRANÇAISE

INSTITUT NATIONAL
de la
PROPRIETE INDUSTRIELLE

RAPPORT DE RECHERCHE PRELIMINAIRE

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

N° d'enregistrement national

FA 539708
FR 9700423

DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendications concernées de la demande examinée
A	HIROTADA UEDA ET AL: "AUTOMATIC STRUCTURE VISUALIZATION FOR VIDEO EDITING" BRIDGES BETWEEN WORLDS, AMSTERDAM, APR. 24 - 29, 1993, no. -, 24 avril 1993, ASHLUND S;MULLET K; HENDERSON A; HOLLNAGEL E; WHITE T, pages 137-141, XP000570441 * page 138, colonne de gauche, ligne 17 - page 139, colonne de gauche, ligne 16 * -----	1,7,8, 11,12

DOMAINES TECHNIQUES RECHERCHES (Int.CL.6)

Date d'achèvement de la recherche	Examinateur
29 septembre 1997	Fournier, C

CATEGORIE DES DOCUMENTS CITES

X : particulièrement pertinent à lui seul
Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie
A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général
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

2

EPO FORM 1503 03.82 (P04C13)

REPUBLIQUE FRANÇAISE

INSTITUT NATIONAL
de la
PROPRIETE INDUSTRIELLE

RAPPORT DE RECHERCHE PRELIMINAIRE

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

N° d'enregistrement national

FA 539708
FR 9700423

DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendications concernées de la demande examinée
D,X	MADRANE N ET AL: "VIDEO REPRESENTATION TOOLS USING A UNIFIED OBJECT AND PERSPECTIVE BASED APPROACH" PROCEEDINGS OF THE SPIE, vol. 2420, 9 février 1995, pages 152-163, XP000571900	1,11
A	* page 160, alinéa 4.2.3 - page 162, alinéa 5; figures 8,10,12-14 *	2,3,14
A	KATAOKA R ET AL: "ARCHITECTURE AND STORAGE STRUCTURE OF AN INTERACTIVE MULTIMEDIA INFORMATION SYSTEM" IEICE TRANSACTIONS ON INFORMATION AND SYSTEMS, vol. E78-D, no. 11, 1 novembre 1995, pages 1354-1361, XP000553522 * page 1356, colonne de droite, alinéa 4 - page 1359, colonne de droite, alinéa 5 *	1-3,11, 17
A	AKUTSU A ET AL: "VIDEO INDEXING USING MOTION VECTORS" PROCEEDINGS OF THE SPIE, vol. 1818, no. PART 03, 18 novembre 1992, pages 1522-1530, XP000671350 * page 1529, alinéa 5 - page 1530, ligne 15 *	1,11,14, 15
A	TONOMURA Y ET AL: "CONTENT ORIENTED VISUAL INTERFACE USING VIDEO ICONS FOR VISUAL DATABASE SYSTEMS" JOURNAL OF VISUAL LANGUAGES AND COMPUTING, vol. 1, 1 janvier 1990, pages 183-198, XP000195706 * le document en entier *	1,11

-/--

DOMAINES TECHNIQUES RECHERCHES (Int.CL.6)

G06F

Date d'achèvement de la recherche	Examinateur
29 septembre 1997	Fournier, C

CATEGORIE DES DOCUMENTS CITES

X : particulièrement pertinent à lui seul
Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie
A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général
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

EPO FORM 1503 03.82 (P04C13)