

①⑨ RÉPUBLIQUE FRANÇAISE
INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE
PARIS

①⑪ N° de publication :
(à n'utiliser que pour les
commandes de reproduction)

2 889 761

②① N° d'enregistrement national : **05 52479**

⑤① Int Cl⁸ : G 06 T 17/00 (2006.01), H 04 N 7/15, 13/00

①②

DEMANDE DE BREVET D'INVENTION

A1

②② Date de dépôt : 09.08.05.

③⑦ Priorité :

④③ Date de mise à la disposition du public de la
demande : 16.02.07 Bulletin 07/07.

⑤⑥ Liste des documents cités dans le rapport de
recherche préliminaire : *Ce dernier n'a pas été
établi à la date de publication de la demande.*

⑥⑦ Références à d'autres documents nationaux
apparentés :

⑦① Demandeur(s) : TOTAL IMMERSION Société anonyme
— FR.

⑦② Inventeur(s) : LEFEVRE VALENTIN et PASSAMA
MARION.

⑦③ Titulaire(s) :

⑦④ Mandataire(s) : SANTARELLI.

⑤④ **SYSTEME PERMETTANT A UN UTILISATEUR DE LOCALISER UNE CAMERA AFIN DE POUVOIR INSERER, RAPIDEMENT DE MANIERE AJUSTEE, DES IMAGES D'ELEMENTS VIRTUELS DANS DES IMAGES VIDEO D'ELEMENTS REELS CAPTEES PAR LA CAMERA.**

⑤⑦ L'invention concerne un système permettant à un utilisateur de localiser une caméra afin de pouvoir insérer des images d'éléments virtuels dans des images vidéo d'éléments réels captées par la caméra, caractérisé en ce qu'il comprend:

- une caméra vidéo destinée à capter des images vidéo d'éléments réels,
- un équipement informatique connecté à la caméra vidéo,
- des premiers moyens de traitement informatique permettant d'afficher:
 - une zone de vidéo temps réel,
 - une zone d'images de synthèse,
 - des moyens de sélection, permettant de sélectionner des points clés dans la zone de vidéo temps réel et des points clés équivalents dans la zone d'images de synthèse,
 - des seconds moyens de traitement informatique destinés à apparier les points clés de la zone de vidéo temps réel et les points clés équivalents de la zone d'images de synthèse.

FR 2 889 761 - A1



Système permettant à un utilisateur de localiser une caméra afin de pouvoir insérer, rapidement de manière ajustée, des images d'éléments virtuels dans des images vidéo d'éléments réels captées par la caméra

Les technologies de « réalité augmentée » permettent de mixer, en temps réel, des images vidéo avec des images de synthèse.

Dans de nombreuses applications, l'utilisateur est confronté aux problèmes suivants :

- Il possède une modélisation 3D qu'il souhaite placer précisément par rapport à un endroit réel filmé par la caméra.
- Une partie de la modélisation 3D correspond à des éléments déjà présents sur l'endroit réel (la modélisation 3D pouvant aussi contenir des éléments futurs non présents dans l'endroit réel).
- Il souhaite installer sa caméra ou bon lui semble pour trouver le meilleur point de vue.
- Il souhaite placer rapidement les éléments virtuels dans l'image vidéo.

La présente invention décrit un outil de localisation rapide qui résout les contraintes énoncées ci-dessus.

L'invention concerne un système permettant à un utilisateur de localiser une caméra afin de pouvoir insérer rapidement de manière ajustée des images d'éléments virtuels dans des images vidéo d'éléments réels captées par la caméra, caractérisé en ce qu'il comprend :

- une caméra vidéo destinée à capter des images vidéo d'éléments réels,
- un équipement informatique, connecté à la caméra vidéo, notamment via une liaison vidéo composite, SVideo, DV, SDI ou HD-SDI, comportant des moyens de visualisation, notamment un écran,
- des premiers moyens de traitement informatique permettant d'afficher sur les moyens de visualisation :
 - une zone de vidéo temps réel, permettant à l'utilisateur de visualiser les images vidéo des éléments réels captés par la caméra vidéo,
 - une zone d'images de synthèse permettant à l'utilisateur de visualiser des images d'éléments virtuels issus d'une base de données d'éléments en trois dimensions,
- des moyens de sélection, notamment une souris associée à l'équipement informatique, permettant à l'utilisateur de sélectionner des points clés dans la zone de vidéo temps réel et des points clés équivalents dans la zone d'images de synthèse,

- des seconds moyens de traitement informatique destinés à apparier les points clés de la zone de vidéo temps réel et les points clés équivalents de la zone d'images de synthèse,
- des moyens de calcul de la pose de la caméra vidéo par rapport aux éléments virtuels issus de la base de données d'éléments en trois dimensions, en fonction des points clés ainsi appariés,
- des troisièmes moyens de traitement informatique destinés à afficher, en temps réel, des éléments virtuels superposés aux images vidéo captées par la une caméra vidéo.

Les applications concernées par la présente invention sont notamment les suivantes (liste non exhaustive) :

- Dans le domaine de la construction / bâtiments :
 - Sur un chantier, vérifier l'état d'avancement des travaux, en superposant les travaux théoriques (mémorisé sous forme de base de données de synthèse) sur les travaux réels filmés par la caméra.
 - Sur une maquette miniature réelle symbolisant le but à atteindre, ajouter une base de données de synthèse.
 - Implantation d'usines : visualiser des travaux non encore réalisés dans une usine existante, pour tester la viabilité du projet.
- Dans le domaine automobile :
 - Calage d'un cockpit virtuel sur un cockpit réel.
 - Calage d'un véhicule virtuel dans un environnement réel (exemple : showroom).

Description de l'outil

L'outil est composé des éléments suivants :

- Une caméra :
 - Par exemple une caméra robotisée en pan / tilt /zoom, éventuellement posée sur un trépied (par exemple une caméra Sony EVI D100 ou Sony EVI D100P).
 - Par exemple une caméra en plan fixe.
 - Par exemple une caméra associée à un capteur de mouvement.
- Un ordinateur PC, éventuellement un PC portable pour une plus grande mobilité.
- La caméra est branchée au PC avec deux types de connexions :
 - Une connexion vidéo (par exemple : composite, SVideo, DV, SDI ou HD-SDI)
 - Si besoin une connexion série (par exemple pour l'envoi en temps réel des paramètres de pan /tilt/zoom de la caméra EVI D100 vers le PC).
- L'ordinateur est équipé du logiciel D'FUSION configuré spécifiquement pour l'outil de localisation.

Procédure de localisation de la caméra

L'utilisateur emmène le système sur le lieu réel à observer.

Il est totalement libre de choisir le point de vue le plus « pertinent ».

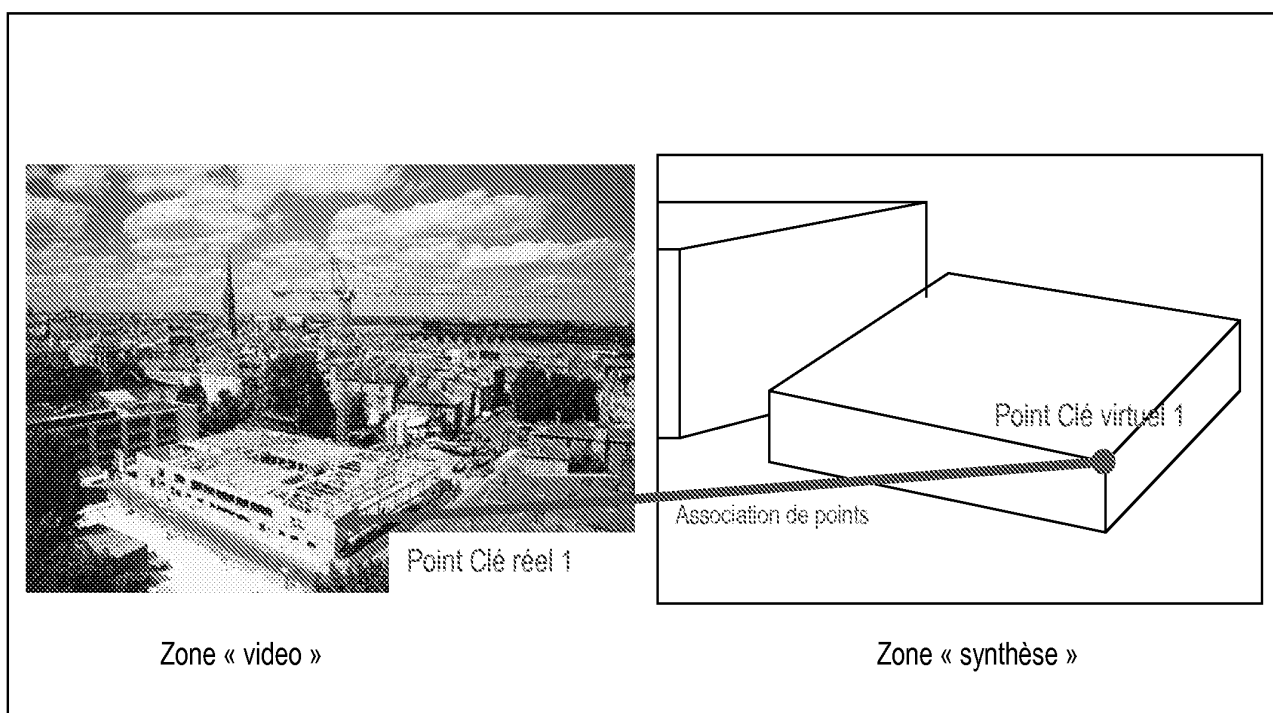
Il installe la caméra pan/tilt/zoom sur son trépied.

Localisation rapide de la caméra grâce à l'utilisation de traitements temps réels sur la vidéo, de rendu 3D temps réel et d'acquisitions de données en temps réel.

Cette procédure est indispensable pour obtenir un positionnement parfait des objets de synthèse sur la scène réelle.

Une fois que l'utilisateur a lancé le logiciel de localisation (D'FUSION configuré spécifiquement), une fenêtre apparaît, divisée en deux parties :

- Une zone de vidéo temps réel, où l'on visualise l'image en provenance de la caméra.
- Une zone « image de synthèse », qui affiche la base de données en 3 dimensions.



Pour localiser la caméra, l'utilisateur associe des points clé virtuels (dans la zone « synthèse ») à leur équivalent (Points clé réels) dans la zone « vidéo ». Ce processus doit être aussi précis et rapide que possible grâce aux fonctionnalités suivantes :

Sélection des points clé réels dans la zone « vidéo » :

- Dans le cas d'une caméra robotisée, on contrôle le mouvement de la caméra à la souris (pan / tilt contrôlé avec les axes X/Y de la souris, contrôle du zoom avec la molette de la souris).
- Dans le cas d'une caméra robotisée, on contrôle le zoom optique sur les points clé réels pour améliorer la précision (les points clés réels peuvent être sélectionnés à l'intérieur de l'image zoomée).
- Une fois sélectionné, un point clé réel reste affiché (avec son numéro d'index) dans l'image vidéo même si la caméra bouge en pan/tilt/zoom.
- L'utilisateur peut sélectionner N points clés dans la zone « vidéo », qui resteront affichés en temps réel avec leur index 1 à N.

Sélection des points clé virtuels dans la zone « synthèse » :

- L'utilisateur peut facilement déplacer à la souris le point de vue de la caméra virtuelle pour obtenir rapidement un point de vue virtuel « proche » du point de vue de la caméra réelle (la position et l'orientation de la caméra virtuelle sont modifiables comme dans un modeleur standard).
- Une fois le point de vue fixé dans la zone « synthèse », l'utilisateur peut sélectionner les N points clé virtuels, en cliquant avec la souris sur chaque point. Les points clés virtuels apparaissent avec leur index, et ils sont toujours bien placés, même si l'utilisateur change les paramètres de la caméra virtuelle.
- Grâce à l'algorithme de « picking », chaque point clé virtuel sélectionné est associé à 3 coordonnées (X, Y, Z) dans le repère de la base de donnée de synthèse.

Calcul de la pose de la caméra réelle par rapport à la base de donnée de synthèse :

Le logiciel possède en mémoire les informations suivantes :

- Les N points clés réels dans l'image réelle.
- Les N points clés virtuels dans l'image virtuelle, avec pour chaque point clé virtuel ses coordonnées (X, Y, Z) dans le repère de la base de données virtuelle.
- Pour calculer la pose de la caméra, on utilise l'algorithme POSIT (voir <http://www.cfar.umd.edu/~daniel/>). Cet algorithme est déjà publié par Daniel DeMenthon, la publication a pour titre : « D. DeMenthon and L.S. Davis, "Model-Based Object Pose in 25 Lines of Code", International Journal of Computer Vision, 15, pp. 123-141, June 1995 ».

Affichage de la base de donnée de synthèse superposée sur la vidéo réelle :

- Le mélange réel / virtuel est réalisé grâce :
 - o Dé-distortion temps réel de l'image caméra.
 - o Utilisation de la pose caméra calculée grâce à la méthode décrite dans le présent document.
 - o Si besoin, la base de donnée peut être affichée en transparence sur la scène réelle grâce aux techniques de blending mises en œuvre dans la technologie D'FUSION.

Avantages de la solution :

- Facilement transportable (un PC portable et une caméra).
- Peut fonctionner sur des maquettes ou à l'échelle 1.
- Peut fonctionner à l'intérieur ou à l'extérieur de bâtiments.
- Peut fonctionner à l'intérieur ou à l'extérieur de véhicules.
- Temps d'installation et de calcul de pose rapide grâce à l'interface décrite dans le présent document.
- Pas d'obligation d'utilisation de capteurs hardware.
- Les fonction pan/tilt/zoom de la caméra sont très utiles pour une utilisation dans des bâtiments : pour travailler à l'échelle 1 face à des buildings ou en intérieur de bâtiments (la plupart du temps, l'utilisateur ne dispose que de peu de recul, donc la scène réelle n'est vue que partiellement par la caméra).

Revendication

1. Système permettant à un utilisateur de localiser une caméra afin de pouvoir insérer rapidement de manière ajustée des images d'éléments virtuels dans des images vidéo d'éléments réels captées par la caméra, caractérisé en ce qu'il comprend :

- une caméra vidéo destinée à capter des images vidéo d'éléments réels,

- un équipement informatique, connecté à la caméra vidéo, notamment via une liaison vidéo composite, SVideo, DV, SDI ou HD-SDI, comportant des moyens de visualisation, notamment un écran,

- des premiers moyens de traitement informatique permettant d'afficher sur les moyens de visualisation :

- une zone de vidéo temps réel, permettant à l'utilisateur de visualiser les images vidéo des éléments réels captés par la caméra vidéo,

- une zone d'images de synthèse permettant à l'utilisateur de visualiser des images d'éléments virtuels issus d'une base de données d'éléments en trois dimensions,

- des moyens de sélection, notamment une souris associée à l'équipement informatique, permettant à l'utilisateur de sélectionner des points clés dans la zone de vidéo temps réel et des points clés équivalents dans la zone d'images de synthèse,

- des seconds moyens de traitement informatique destinés à apparier les points clés de la zone de vidéo temps réel et les points clés équivalents de la zone d'images de synthèse,

- des moyens de calcul de la pose de la caméra vidéo par rapport aux éléments virtuels issus de la base de données d'éléments en trois dimensions, en fonction des points clés ainsi appariés,

- des troisièmes moyens de traitement informatique destinés à afficher, en temps réel, des éléments virtuels superposés aux images vidéo captées par la caméra vidéo.