

**TITRE : Dispositif et procédé d'estimation de mouvement d'un capteur d'images
entre deux images, programme d'ordinateur associé**

5 La présente invention concerne un procédé d'estimation de mouvement d'un capteur d'images entre deux images capturées par ce capteur.

 Dans le domaine du traitement vidéo, de nombreuses méthodes de stabilisation numérique d'images, de détections d'objets mobiles en caméra embarquée, de construction de bandeaux, de détermination de trajectoire d'objets ... utilisent de tels
10 procédés d'estimation de mouvement entre deux images, comprenant généralement l'estimation d'un vecteur, dit vecteur de déplacement, définissant le mouvement de la caméra entre deux images consécutives ; un tel vecteur est par exemple composé des paramètres d'un modèle de déplacement entre les deux images.

 Bien souvent, ces méthodes de calculs de flux de déplacement minimisent une
15 fonction d'erreur de manière itérative en partant d'un vecteur de déplacement de valeur d'initialisation donnée et proposent ainsi, au fil des itérations, des vecteurs de déplacement de plus en plus proches d'un minimum de cette fonction d'erreurs (en supposant que ce minimum local corresponde au vecteur de déplacement vrai). Les méthodes itératives utilisées sont typiquement celles de type Newton ou Gauss-Newton
20 ou encore de descente de gradient.

 Ces méthodes présentent plusieurs problèmes. Si le vecteur initial de déplacement est trop éloigné, au sens de la norme euclidienne, du déplacement vrai, la méthode ne converge pas et produit un résultat inexploitable. Un deuxième problème, conséquence directe du premier est que, comme le vecteur initial ne doit pas être trop éloigné du
25 vecteur vrai, les amplitudes des déplacements entre images qui peuvent être estimées s'en trouvent limitées. Un troisième problème est que rien ne garantit, dans le cas général, que la fonction d'erreur à minimiser soit convexe : l'algorithme peut donc converger vers un minimum local éloigné de la solution vraie. Enfin, la convergence de ces méthodes peut nécessiter de nombreuses itérations, coûteuses en temps (et
30 ressources) de calcul, ce qui est problématique dans des applications temps réel.

 Le but de l'invention est alors de proposer une solution qui contribue à réduire ces problèmes.

 A cet effet, suivant un premier aspect, l'invention propose un procédé d'estimation de mouvement entre deux images mis en œuvre par un dispositif de traitement
35 électronique, lesdites images étant capturées à l'aide d'un appareil de capture d'images, ledit procédé comprenant les étapes consistant à :

2

- recueillir les données définissant chacune des deux images ;
- déterminer un vecteur de déplacement d'initialisation caractérisant un déplacement entre lesdites deux images ;

5 - mettre en œuvre, à partir du vecteur de déplacement d'initialisation, un algorithme itératif délivrant à chaque itération un vecteur de déplacement actualisé en fonction d'une fonction d'erreur à minimiser entre l'une des deux images et la transformée de l'autre image par un déplacement défini par le vecteur de déplacement actualisé,

10 - ledit procédé étant caractérisé en ce qu'il comprend les étapes consistant à :
 - recueillir des informations définissant le mouvement de l'appareil de capture d'images entre les deux images ;

 - recueillir la valeur de paramètre(s) de capture de l'appareil de capture d'images utilisé(s) pour la capture des deux images parmi une ouverture angulaire, un facteur de zoom et la résolution des images ;

15 - déterminer le vecteur de déplacement d'initialisation par projection, dans le plan image de l'appareil de capture d'image, du déplacement entre les deux images calculés en fonction des informations recueillies définissant le mouvement de l'appareil de capture d'images entre les deux images et en fonction de la valeur de chaque paramètre de capture.

20 L'invention permet ainsi d'utiliser les informations de paramétrage de la caméra ainsi que les informations de position et d'orientation de la caméra issues d'une centrale inertielle couplée à la caméra afin de proposer un premier vecteur de déplacement X_0 proche du vecteur vrai. Cela permet d'améliorer globalement les performances en accroissant la probabilité de converger vers le déplacement vrai. En particulier, cela
 25 améliore la précision de la solution fournie et permet d'estimer des déplacements plus grands entre images. De plus, utiliser un premier vecteur de déplacement X_0 proche du vrai a pour conséquence une diminution du nombre d'itérations et accélérera aussi la convergence tout en réduisant naturellement les temps et ressources de calcul.

30 Dans des modes de réalisation, le procédé d'estimation de mouvement suivant l'invention comporte en outre une ou plusieurs des caractéristiques suivantes :

- les informations définissant le mouvement de l'appareil de capture d'images entre les deux images sont fournies par un dispositif de détermination de mouvement attaché à l'appareil de capture d'images ;
- le vecteur de déplacement d'initialisation est pris égal à

35 $X_0^{opt} = (a_1^{opt}, a_2^{opt}, a_3^{opt}, a_4^{opt}, a_5^{opt}, a_6^{opt})^T$, où

$$\left\{ \begin{array}{l} a_1^{opt} = \frac{\left(\frac{W-1}{2} \left(1 + \frac{\tan(\psi)}{\tan\left(\frac{\alpha_\psi}{2}\right)} - s \times \cos(\varphi) \right) + \frac{H-1}{2} s \times \sin(\varphi) \right)^3}{2^{L-1}} \\ a_2^{opt} = s \times \cos(\varphi) - 1 \\ a_3^{opt} = -s \times \sin(\varphi) \\ a_4^{opt} = \frac{\left(\tan(\theta) \times \sqrt{\frac{(H-1)^2}{4 \tan\left(\frac{\alpha_\theta}{2}\right)^2} + \frac{(W-1)^2}{4} \frac{\tan(\psi)^2}{\tan\left(\frac{\alpha_\psi}{2}\right)^2}} - \frac{W-1}{2} s \times \sin(\varphi) + \frac{H-1}{2} (1 - s \times \cos(\varphi)) \right)}{2^{L-1}} \\ a_5^{opt} = s \times \sin(\varphi) \\ a_6^{opt} = s \times \cos(\varphi) - 1 \end{array} \right.$$

où L est le nombre de niveaux d'une pyramide de réduction successive de résolution d'image dans le cas d'un procédé d'estimation de mouvement multi-résolution ou est égal à 1 sinon ; (ψ, θ, φ) sont les angles d'Euler indiqués par les informations recueillies, φ représentant en outre la rotation de la caméra autour de sa ligne de visée, les images ayant W pixels en largeur et H pixels en hauteur.

Suivant un deuxième aspect, la présente invention propose un dispositif électronique d'estimation de mouvement entre deux images comprenant une entrée pour recueillir les données définissant chacune des deux images et étant adapté pour déterminer un vecteur de déplacement d'initialisation caractérisant un déplacement entre lesdites deux images et pour mettre en œuvre, à partir du vecteur de déplacement d'initialisation, un algorithme itératif délivrant à chaque itération un vecteur de déplacement actualisé en fonction d'une fonction d'erreur à minimiser entre l'une des deux images et la transformée de l'autre image par un déplacement défini par le vecteur de déplacement actualisé, ledit dispositif étant caractérisé en ce qu'il est adapté pour recueillir des informations définissant le mouvement de l'appareil de capture d'images entre les deux images, pour recueillir la valeur de paramètre(s) de capture d'un appareil de capture d'images utilisé(s) pour la capture des deux images parmi une ouverture angulaire, un facteur de zoom et la résolution des images et pour déterminer le vecteur de déplacement d'initialisation par projection, dans le plan image de l'appareil de capture d'image, du déplacement entre les deux images calculé en fonction des informations recueillies définissant le mouvement de l'appareil de capture d'images entre les deux images et en fonction de la valeur de chaque paramètre de capture.

Dans un mode de réalisation, ledit dispositif est adapté pour prendre le vecteur de déplacement d'initialisation égal à $X_0^{opt} = (a_1^{opt}, a_2^{opt}, a_3^{opt}, a_4^{opt}, a_5^{opt}, a_6^{opt})^T$, où

$$\left\{ \begin{array}{l} a_1^{opt} = \frac{\left(\frac{W-1}{2} \left(1 + \frac{\tan(\psi)}{\tan\left(\frac{\alpha_\psi}{2}\right)} - s \times \cos(\varphi) \right) + \frac{H-1}{2} s \times \sin(\varphi) \right)^4}{2^{L-1}} \\ a_2^{opt} = s \times \cos(\varphi) - 1 \\ a_3^{opt} = -s \times \sin(\varphi) \\ a_4^{opt} = \frac{\left(\tan(\theta) \times \sqrt{\frac{(H-1)^2}{4 \tan\left(\frac{\alpha_\theta}{2}\right)^2} + \frac{(W-1)^2}{4} \frac{\tan(\psi)^2}{\tan\left(\frac{\alpha_\psi}{2}\right)^2}} - \frac{W-1}{2} s \times \sin(\varphi) + \frac{H-1}{2} (1 - s \times \cos(\varphi)) \right)^4}{2^{L-1}} \\ a_5^{opt} = s \times \sin(\varphi) \\ a_6^{opt} = s \times \cos(\varphi) - 1 \end{array} \right.$$

où L est le nombre de niveaux d'une pyramide de réduction successive de résolution d'image dans le cas d'un procédé d'estimation de mouvement multi-résolution ou est égal à 1 sinon ; (ψ , θ , φ ,) sont les angles d'Euler indiqués par les informations recueillies, φ représentant en outre la rotation de la caméra autour de sa ligne de visée, les images ayant W pixels en largeur et H pixels en hauteur.

Suivant un troisième aspect, la présente invention propose un programme d'ordinateur comportant des instructions logicielles qui, lorsqu'elles sont exécutées par un ordinateur, mettent en œuvre un procédé tel que défini ci-dessus

Ces caractéristiques et avantages de l'invention apparaîtront à la lecture de la description qui va suivre, donnée uniquement à titre d'exemple, et faite en référence aux dessins annexés, sur lesquels :

- [Fig 1] la figure 1 représente deux images successives I_1 , I_2 capturées par une caméra dans un mode de réalisation de l'invention ;
- [Fig 2] la figure 2 est une vue d'un système de traitement d'images dans un mode de réalisation de l'invention ;
- [Fig 3] la figure 3 est une vue de pyramides d'images dans un mode de réalisation en multi-résolution de l'invention ;
- [Fig 4] la figure 4 illustre les angles d'Euler utilisés dans un mode de réalisation de l'invention ;
- [Fig 5] les figures 5, [Fig 6] 6 et [Fig 7] 7 illustrent les transformations par rapport au plan image dans un mode de réalisation de l'invention ;
- [Fig 8] la figure 8 est un organigramme d'étapes mises en œuvre dans un mode de réalisation de l'invention.

La figure 2 est une vue schématique d'un système de traitement d'images 1 dans un mode de réalisation de l'invention. Le système de traitement d'images 1 comprend un dispositif électronique d'estimation de mouvement 10, un bloc de capture d'images 2 et un bloc de détermination de mouvement 3, adapté pour déterminer les mouvements de la

5

caméra. Dans le cas considéré, le bloc de capture d'images 2 est une caméra 2 et le bloc de détermination de mouvement 3 est une centrale inertielle 3 solidaire de la caméra 2 (par exemple une mini-centrale inertielle).

La caméra 2 comporte des moyens de réglage de paramètres de capture d'image, pour régler la valeur des paramètres pour la capture de chaque image (le réglage d'un paramètre donné à une valeur parmi plusieurs valeurs proposées est par exemple effectué automatiquement et/ou en fonction de valeurs saisies par un opérateur depuis une interface homme-machine de la caméra). Parmi ces paramètres figurent par exemple un ou des paramètres d'ouverture angulaire de caméra et/ou un facteur de zoom utilisé

15

pour la capture et la résolution des images. La caméra 2 est adaptée pour délivrer à destination du dispositif électronique d'estimation de mouvement 10 chaque image capturée, sous la forme de données d'images (par exemple une valeur d'intensité associée à chaque pixel d'une matrice de pixels), associée à la valeur prise par chaque paramètre de capture d'image lors de la

20

La centrale inertielle 3 comporte :

- trois gyromètres mesurant les trois composantes du vecteur vitesse angulaire (vitesses de variation des angles de roulis, de tangage et de lacet) ; et
- trois accéléromètres mesurant les trois composantes du vecteur force spécifique (la force spécifique est la somme des forces extérieures autres que gravitationnelles divisée par la masse)

25

La centrale inertielle 3 comporte en outre un calculateur (non représenté) adapté pour déterminer par intégration en temps réel, à partir des mesures successives de ces six capteurs, la position de la caméra 2 (par exemple la position d'un point particulier de

30

caméra 2) et son orientation (par exemple les angles d'Euler donnant l'attitude : roulis, tangage et cap) par rapport à un repère de référence. La centrale inertielle 3 est adaptée pour délivrer ces informations de position et orientation à destination du dispositif électronique d'estimation de mouvement 10.

35

Pour rappel, l'attitude d'un corps rigide, ici de la caméra 2, est définie par la façon dont ce dernier est orienté. En général, pour exprimer une attitude, on considère deux

repères : un repère de référence fixe, noté N, d'axes X , Y , Z , et un repère mobile, attaché au corps, noté B, d'axes X_G , Y_G , Z_G .

Pour exprimer convenablement l'attitude, les axes du repère B sont choisis alignés avec les axes principaux d'inertie du corps rigide définissant sa rotation. Un repère de référence fixe souvent utilisé, est celui qui est tangent à la surface de la terre et défini selon la convention NED (North, East, Down), selon laquelle les vecteurs unitaires (X , Y , Z) pointent respectivement vers les directions Nord, Est et la pesanteur (la gravité). L'orientation, correspondant au passage du repère N au repère B lié au corps, peut être décrite à l'aide d'une succession de trois rotations élémentaires autour des trois axes du repère, respectivement X , Y , Z et définies par les angles d'Euler respectivement ψ , θ , ϕ .

Dans le mode de réalisation considéré, la caméra 2, respectivement la centrale inertielle 3, est reliée au dispositif d'estimation de mouvement 10, par une liaison de télécommunication, par exemple une liaison de radiocommunication dans un mode de réalisation.

Le dispositif d'estimation de mouvement 10 est adapté pour estimer le déplacement de la caméra 3 entre deux images I_1 , I_2 telles que capturées par la caméra 2 et représentées en figure 1, en mettant en œuvre une méthode itérative.

On notera ici que les imprécisions des informations sur les informations de position et d'orientation fournies par la centrale inertielle 3 ne permettent pas de traiter le problème directement et que l'étape de traitement d'images itérative est nécessaire pour obtenir le niveau de précision souhaitée quant à l'estimation du déplacement, niveau de précision supérieur à celui fourni par les informations de position et d'orientation fournies par la centrale inertielle 3.

La méthode itérative d'estimation est par exemple (dans les solutions existantes) insérée dans un schéma multi-résolution où les dimensions des images sont réduites par 2×2 (on obtient une pyramide inversée d'images), à $L-1$ reprises, jusqu'à obtenir de petites images où le choix (encore une fois dans les méthodes standards) d'un premier vecteur de déplacement X_0 égal au vecteur nul est supposé se trouver proche du vecteur vrai, où L est un nombre entier supérieur ou égal à 1. La figure 3 illustre une pyramide d'images à 4 niveaux.

Dans le mode de réalisation considéré, le dispositif électronique d'estimation de mouvement 10 comporte une interface de collecte 11 et un bloc électronique de traitement 12.

L'interface de collecte 11 est adaptée pour collecter les données de chaque image, par exemple I_1 , I_2 , ainsi que pour chaque image, les valeurs des paramètres de la caméra 2 délivrées par la caméra 2 ainsi que les informations de position et orientation

délivrées par la centrale inertielle 3 correspondant au temps de capture de cette image. Sont ainsi fournis dans le cas considéré notamment les trois angles d'Euler (ψ, θ, φ) de déplacement de la caméra entre les prises des images I_1 et I_2 comme représenté en figure 4.

5 Dans un mode de réalisation, le dispositif d'estimation de mouvement 10 est adapté pour mettre en œuvre les étapes décrites en référence à la figure 8.

Dans un mode de réalisation, le bloc de traitement 12 comporte une mémoire et un microcalculateur (non représentés), la mémoire stockant des instructions logicielles qui, lorsqu'elles sont exécutées sur le microcalculateur, mettent en œuvre les étapes
10 décrites en référence à la figure 8.

Ainsi, lorsqu'une nouvelle image est collectée, le bloc de traitement 12 récupère les informations de la caméra 2 (ouvertures angulaires, facteur de zoom, ...) et les informations correspondantes de la centrale inertielle 3. Ces informations, comparées avec celles de l'image précédente permettent d'estimer quel est le déplacement de la
15 caméra entre les instants de capture des deux images, de projeter dans le plan image de la caméra 2 (et au niveau le plus bas de la pyramide inversée dans le cas d'un schéma multi-résolution) ce déplacement et d'obtenir ainsi un vecteur de déplacement d'initialisation, X_0 . Ces opérations sont décrites plus en détail ci-après.

En référence à la figure 8, dans une étape 100 de collecte, l'interface de collecte
20 11 collecte les données de l'image I_1 , puis I_2 , ainsi que pour chaque image, les valeurs des paramètres de la caméra 2 et les informations de position et orientation et les fournit au bloc de traitement 12.

Dans une étape 101 de mise en œuvre du schéma multi-résolution, deux pyramides gaussiennes, chacune de niveau maximum égal à L (L dans le cas considéré
25 égal à 4) vont donc être calculées pour les deux images I_1 et I_2 , comme le montre la figure 3. Le choix du niveau L dépend principalement de la résolution des deux images (identiques dans le cas présent).

Afin de comprendre l'apport de la solution proposée, commençons par rappeler comment fonctionne une méthode d'estimation de mouvement entre deux images (de
30 même résolution).

Soient donc deux image I_1 et I_2 , le but est d'estimer le déplacement entre ces deux images comme le montre la figure 1.

Pour estimer ce mouvement de déplacement, un modèle de déplacement ou transformation T entre ces deux images est utilisé et l'objectif est d'estimer les paramètres
35 de ce modèle de déplacement. Ce modèle de déplacement ou transformation va donner

8

pour chaque pixel (i_1, j_1) de l'image I_1 le vecteur de déplacement ou flux de déplacement (u, v) à appliquer pour atteindre le pixel correspondant (i_2, j_2) de l'image I_2 . Le modèle de déplacement utilisé est défini par :

$$\begin{pmatrix} u(i_1, j_1) \\ v(i_1, j_1) \end{pmatrix} = \begin{pmatrix} a_1 \\ a_4 \end{pmatrix} + \begin{pmatrix} a_2 & a_3 \\ a_5 & a_6 \end{pmatrix} \begin{pmatrix} i_1 \\ j_1 \end{pmatrix}. \quad \text{(Formule 1)}$$

Il s'agit d'une transformation affine de vecteur de coefficients $X = (a_1, a_2, a_3, a_4, a_5, a_6)^T$ permettant de prendre en compte des transformations entre images de type translation, zoom, rotation, ...

Les coordonnées du pixel (i_2, j_2) dans l'image I_2 correspondant au pixel (i_1, j_1) dans l'image I_1 sont obtenues par les formules suivantes.

$$\begin{cases} i_2 = i_1 + u(i_1, j_1) \\ j_2 = j_1 + v(i_1, j_1) \end{cases}$$

L'objectif va être à chaque niveau n considéré successivement en partant du niveau le plus grand (correspondant aux images les plus petites, i.e. à $n = L-1$) jusqu'au niveau le plus petit ($n = 0$), d'estimer les paramètres de transformation $X = (a_1, a_2, a_3, a_4, a_5, a_6)^T$ entre les deux images I_1, I_2 considérées dans leur résolution correspondant au niveau n , en minimisant (i.e. en faisant descendre sous un seuil pré-défini d'erreur maximum acceptable) une somme d'erreurs aux carrés de la forme suivante, où ρ est une certaine fonction pouvant être l'identité ou une fonction rendant l'estimation robuste aux données aberrantes, cette somme représentant une erreur globale entre le déplacement « vrai » et le déplacement estimé entre les deux images :

$$\sum_{(i_1, j_1)} \rho \left(I_1(i_1, j_1) - I_2(i_1 + u(i_1, j_1), j_1 + v(i_1, j_1)) \right)^2. \quad (1)$$

La résolution de ce problème de minimisation est faite en utilisant une méthode récursive, par exemple de type Gauss-Newton, à partir, selon l'invention, d'une valeur initiale du vecteur de coefficients. Cette valeur initiale du vecteur de coefficients, ou valeur d'initialisation, est notée X_0^{opt} pour le niveau n le plus grand ($n = L-1$) et notée $X_{0,n}$ pour chacun des autres niveaux n .

Un ensemble itératif 102 d'étapes est mis en œuvre, pour chaque niveau n de pyramide, en commençant au niveau le plus haut de la pyramide inversée (là où les

images sont les plus petites : i.e. on considère la plus petite image de la pyramide de I_1 et la plus petite image de la pyramide de I_2 , dans le cas présent correspondant au niveau $n=3$: pour chaque niveau n , les calculs récurrents, par exemple selon la méthode de Gauss Newton, vont ainsi être lancés, avec une estimée initiale respective pour le niveau n pour
 5 estimer le déplacement entre l'image de niveau n de la pyramide de I_1 et l'image de niveau n de la pyramide de I_2 .

Dans la mise en œuvre de l'ensemble 102 d'étape pour $n=L-1$, i.e. pour estimer le déplacement entre la plus petite image de la pyramide de I_1 et la plus petite image de la pyramide de I_2 , les paramètres de la caméra 2 et les informations de la centrale inertielle
 10 3 associée à celle-ci sont utilisés pour calculer la valeur d'initialisation du vecteur de coefficients X_0^{est} .

Il est également considéré dans le mode de réalisation particulier que le repère de la centrale inertielle 3 est calé avec l'axe de visée de la caméra 2 de telle sorte que la rotation φ corresponde à la rotation du plan image autour de son centre.

Considérons que les valeurs de paramètres fournies par la caméra 2 indiquent entre les deux images I_1 et I_2 un facteur de zoom noté s . Notons α_ψ l'ouverture angulaire en azimut de la caméra et α_θ l'ouverture en site, qui indiquent les variations d'ouverture entre les deux images. Ces informations d'ouvertures angulaires sont fournies par la caméra 2 à l'étape 100. Les images produites ont W pixels en largeur et H pixels en
 20 hauteur.

Les translations t_i et t_j engendrées dans le plan image PI par les rotations ψ et θ de la caméra 2 sont représentées en figures 5 à 7.

En référence à la figure 5, on a

$$t_i = \tan(\psi) \times f.$$

25 où f est la focale de la caméra 2. Or,

$$f = \frac{l}{\tan\left(\frac{\alpha_\psi}{2}\right)},$$

où l est la distance horizontale entre le centre de l'image et le bord de droite (voir figure

5)

30 d'où

$$t_i = l \times \frac{\tan(\psi)}{\tan\left(\frac{\alpha_\psi}{2}\right)},$$

10

en passant dans le plan image avec des longueurs en pixels, il est finalement obtenu :

$$t_i = \frac{W-1}{2} \times \frac{\tan(\psi)}{\tan\left(\frac{\alpha_\psi}{2}\right)}.$$

5

t_j , en référence aux figures 6 et 7, peut être calculé ainsi (voir également figure 5) :

$$t_j = \tan(\theta) \times f' = \tan(\theta) \sqrt{f^2 + t_i^2}.$$

Or

$$f = \frac{h}{\tan\left(\frac{\alpha_\theta}{2}\right)},$$

10

ainsi,

$$t_j = \tan(\theta) \sqrt{\frac{h^2}{\tan^2\left(\frac{\alpha_\theta}{2}\right)} + t_i^2}.$$

En passant dans le plan image avec des longueurs en pixels, il en résulte que :

15

$$t_j = \tan(\theta) \times \sqrt{\frac{(H-1)^2}{4 \tan^2\left(\frac{\alpha_\theta}{2}\right)^2} + \frac{(W-1)^2}{4} \frac{\tan^2(\psi)}{\tan^2\left(\frac{\alpha_\psi}{2}\right)}}.$$

Il sera détaillé par la suite comment cette translation générée par les angles ψ et θ est prise en compte dans le calcul de X_0^{opt} .

20

Regardons maintenant quelle est la transformation générée dans le plan image par le dernier angle d'Euler φ et le changement de zoom de la caméra 2.

La ligne de visée de la caméra 2 et la centrale inertielle 3 sont supposées calées de telle sorte que φ représente la rotation de la caméra 2 autour de sa ligne de visée. Dans ces conditions, la rotation dans le plan image engendrée par φ est représentée par la matrice suivante :

25

$$\begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}.$$

Le changement d'échelle créé par le zoom de la caméra 2 est quant à lui représenté par la matrice suivante :

$$\begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix}.$$

- 5 La transformation générée par φ et le changement de zoom s est représentée par la composition des deux précédentes matrices.

$$\begin{pmatrix} s \times \cos(\varphi) & -s \times \sin(\varphi) \\ s \times \sin(\varphi) & s \times \cos(\varphi) \end{pmatrix}.$$

- 10 Il s'agit ici de la transformation géométrique et non d'un flux de déplacement. Pour obtenir le flux de déplacement affine à partir de cette transformation, il faut soustraire la position i, j , qui est la position de chaque pixel considéré, de l'image I_1 avec i entier compris entre 0 et $W-1$ et j compris entre 0 et $H-1$. Le flux engendré par la transformation précédente qui est une composition d'une rotation et d'un zoom vaut alors :

$$\begin{pmatrix} \frac{W-1}{2} \\ \frac{H-1}{2} \end{pmatrix} + \begin{pmatrix} s \times \cos(\varphi) & -s \times \sin(\varphi) \\ s \times \sin(\varphi) & s \times \cos(\varphi) \end{pmatrix} \begin{pmatrix} i - \frac{W-1}{2} \\ j - \frac{H-1}{2} \end{pmatrix} - \begin{pmatrix} i \\ j \end{pmatrix}.$$

15

On ajoute $(W-1)/2$ et $(H-1)/2$ et on soustrait $(W-1)/2$ à i et $(H-1)/2$ à j dans les deux premiers termes car la rotation et le zoom sont exprimés par rapport au centre de l'image qui se trouve au point de coordonnées $((W-1)/2, (H-1)/2)$. En développant l'expression précédente, il vient :

20

$$\begin{pmatrix} \frac{W-1}{2} + i \times s \times \cos(\varphi) - \frac{W-1}{2} s \times \cos(\varphi) - j \times s \times \sin(\varphi) + \frac{H-1}{2} s \times \sin(\varphi) - i \\ \frac{H-1}{2} + i \times s \times \sin(\varphi) - \frac{W-1}{2} s \times \sin(\varphi) + j \times s \times \cos(\varphi) - \frac{H-1}{2} s \times \cos(\varphi) - j \end{pmatrix}$$

En réorganisant les termes, il vient alors :

$$\begin{pmatrix} \frac{W-1}{2} - s \times \cos(\varphi) \frac{W-1}{2} + s \times \sin(\varphi) \frac{H-1}{2} \\ \frac{H-1}{2} - s \times \sin(\varphi) \frac{W-1}{2} - s \times \cos(\varphi) \frac{H-1}{2} \end{pmatrix} + \begin{pmatrix} s \times \cos(\varphi) - 1 & -s \times \sin(\varphi) \\ s \times \sin(\varphi) & s \times \cos(\varphi) - 1 \end{pmatrix} \begin{pmatrix} i \\ j \end{pmatrix}$$

25

En ajoutant à ce flux de déplacement la translation engendrée par les deux angles de rotation ψ et θ , le flux de déplacement total suivant est obtenu.

$$\begin{pmatrix} t_i + \frac{W-1}{2} - s \times \cos(\varphi) \frac{W-1}{2} + s \times \sin(\varphi) \frac{H-1}{2} \\ t_j + \frac{H-1}{2} - s \times \sin(\varphi) \frac{W-1}{2} - s \times \cos(\varphi) \frac{H-1}{2} \end{pmatrix} + \begin{pmatrix} s \times \cos(\varphi) - 1 & -s \times \sin(\varphi) \\ s \times \sin(\varphi) & s \times \cos(\varphi) - 1 \end{pmatrix} \begin{pmatrix} i \\ j \end{pmatrix}.$$

Par identification et sans oublier la transformation P à appliquer à chaque changement de niveau dans la pyramide, on a alors :

$$\begin{cases} a_1^{opt} = \frac{\left(\frac{W-1}{2} \left(1 + \frac{\tan(\psi)}{\tan\left(\frac{\alpha_\psi}{2}\right)} - s \times \cos(\varphi) \right) + \frac{H-1}{2} s \times \sin(\varphi) \right)}{2^{L-1}} \\ a_2^{opt} = s \times \cos(\varphi) - 1 \\ a_3^{opt} = -s \times \sin(\varphi) \\ a_4^{opt} = \frac{\left(\tan(\theta) \times \sqrt{\frac{(H-1)^2}{4 \tan\left(\frac{\alpha_\theta}{2}\right)^2} + \frac{(W-1)^2}{4} \frac{\tan(\psi)^2}{\tan\left(\frac{\alpha_\psi}{2}\right)^2}} - \frac{W-1}{2} s \times \sin(\varphi) + \frac{H-1}{2} (1 - \dots \right)}{2^{L-1}} \\ a_5^{opt} = s \times \sin(\varphi) \\ a_6^{opt} = s \times \cos(\varphi) - 1 \end{cases}$$

5

(Formule 2)

Ces principes étant exposés, les étapes 102_1 et 102_2 de l'ensemble d'étapes 102 sont maintenant décrites.

10 Pour le niveau L-1, dans une sous-étape 102_1 de détermination de vecteur initial, le vecteur initial X_0^{opt} est pris égal à $X_0^{opt} = (a_1^{opt}, a_2^{opt}, a_3^{opt}, a_4^{opt}, a_5^{opt}, a_6^{opt})^T$ avec les valeurs de ces coefficients définies ci-dessus, en formule 2.

Dans une sous-étape 102_2, pour le niveau L-1, les calculs récursifs, par exemple selon Gauss Newton sont mis en œuvre, à partir du vecteur initial X_0^{opt} déterminé en étape 15 102_1.

Lorsque les calculs récursifs de Gauss-Newton sont terminés, l'algorithme dispose, à ce niveau n=L-1 de la pyramide, d'une estimation $\hat{X}^L$ du vecteur de déplacement X^L entre les deux images courantes des deux pyramides.

20 Ensuite, pour descendre dans les pyramides et ainsi affiner et converger vers une estimation plus précise de X , il faut transformer l'estimation $\hat{X}^L$ afin de la faire passer « au niveau inférieur » dans les pyramides.

13

La transformation P à appliquer à $\hat{X}^L = (\hat{a}_1^L, \hat{a}_2^L, \hat{a}_3^L, \hat{a}_4^L, \hat{a}_5^L, \hat{a}_6^L)^T$ est la suivante :

$$\hat{X}^{L-1} = P(\hat{X}^L) = \begin{pmatrix} 2 \times \hat{a}_1^L \\ \hat{a}_2^L \\ \hat{a}_3^L \\ 2 \times \hat{a}_4^L \\ \hat{a}_5^L \\ \hat{a}_6^L \end{pmatrix}.$$

Une fois cette transformation P faite, il faut recommencer au niveau inférieur, i.e.
 5 au niveau L-2 (avec donc des images plus résolues), l'ensemble d'étapes 102, c'est-à-dire les calculs récursifs de type Gauss Newton (étape 102_2 pour le niveau L-2), mais en prenant cette fois dans l'étape 102_1 pour le niveau L-2, pour estimée initiale $X_{0_L-2} = P(\hat{X}^L)$.

10 Ce processus (102_1 puis 102_2) est répété jusqu'au niveau le plus bas de la pyramide correspondant à $n=0$ (qui correspond aux images I_1 et I_2 originales), en prenant pour chaque niveau comme estimée initiale du vecteur de coefficients la transformée par P de l'estimation du vecteur de coefficients résultant du niveau précédent.

L'estimation finale du mouvement $\hat{X}$ entre les deux images est alors obtenue, en
 15 sortant de l'ensemble d'étapes 102 mis en œuvre pour le niveau 0.

La présente invention permet ainsi d'utiliser les paramètres de la caméra (ouvertures angulaires, facteur de zoom, ...) et les informations d'attitude d'une (par exemple mini-)centrale inertielle couplée à la caméra afin de proposer un premier vecteur
 20 de déplacement d'initialisation plus proche de la solution vraie.

Dans un autre mode de réalisation, le bloc électronique de traitement 12 est réalisé sous forme d'un composant logique programmable, tel qu'un FPGA (de l'anglais *Field Programmable Gate Array*), ou encore sous forme d'un circuit intégré dédié, tel qu'un ASIC (de l'anglais *Applications Specific Integrated Circuit*).

25 La présente invention s'applique à tout algorithme itératif d'estimation de mouvement entre images partant d'une estimation initiale du mouvement : typiquement celles de type Newton ou Gauss-Newton ou encore de descente de gradient etc.

Dans d'autres modes de réalisation, le bloc de détermination de mouvement 3, n'est pas solidaire de la caméra 2 et/ou n'est pas une (mini-)centrale inertielle, mais par
 30 exemple un déplacement donné par un moteur pas à pas ou tout autre système capable de fournir une estimation physique du déplacement de la caméra.

REVENDICATIONS

1. Procédé d'estimation de mouvement entre deux images (I_1, I_2) mis en œuvre par un dispositif de traitement électronique (10), lesdites images étant capturées à l'aide d'un appareil de capture d'images (2), ledit procédé comprenant les étapes consistant à :

- recueillir les données définissant chacune des deux images ;
- déterminer un vecteur de déplacement d'initialisation caractérisant un déplacement entre lesdites deux images ;

- mettre en œuvre, à partir du vecteur de déplacement d'initialisation, un algorithme itératif délivrant à chaque itération un vecteur de déplacement actualisé en fonction d'une fonction d'erreur à minimiser entre l'une des deux images et la transformée de l'autre image par un déplacement défini par le vecteur de déplacement actualisé,

- ledit procédé étant caractérisé en ce qu'il comprend les étapes consistant à :
 - recueillir des informations définissant le mouvement de l'appareil de capture d'images entre les deux images ;

- recueillir la valeur de paramètre(s) de capture de l'appareil de capture d'images utilisé(s) pour la capture des deux images parmi une ouverture angulaire, un facteur de zoom et la résolution des images ;

- déterminer le vecteur de déplacement d'initialisation par projection, dans le plan image de l'appareil de capture d'image, du déplacement entre les deux images calculé en fonction des informations recueillies définissant le mouvement de l'appareil de capture d'images entre les deux images et en fonction de la valeur de chaque paramètre de capture.

2. Procédé d'estimation de mouvement entre deux images (I_1, I_2) selon la revendication 1, selon lequel les informations définissant le mouvement de l'appareil de capture d'images entre les deux images sont fournies par un dispositif de détermination de mouvement attaché à l'appareil de capture d'images.

3. Procédé d'estimation de mouvement entre deux images (I_1, I_2) selon la revendication 1, selon lequel le vecteur de déplacement d'initialisation est pris égal à $X_0^{opt} = (a_1^{opt}, a_2^{opt}, a_3^{opt}, a_4^{opt}, a_5^{opt}, a_6^{opt})^T$, où

$$\begin{cases}
 a_1^{opt} = \frac{\left(\frac{W-1}{2} \left(1 + \frac{\tan(\psi)}{\tan\left(\frac{\alpha_\psi}{2}\right)} - s \times \cos(\varphi) \right) + \frac{H-1}{2} s \times \sin(\varphi) \right)}{2^{L-1}} \\
 a_2^{opt} = s \times \cos(\varphi) - 1 \\
 a_3^{opt} = -s \times \sin(\varphi) \\
 a_4^{opt} = \frac{\left(\tan(\theta) \times \sqrt{\frac{(H-1)^2}{4 \tan\left(\frac{\alpha_\theta}{2}\right)^2} + \frac{(W-1)^2}{4} \frac{\tan(\psi)^2}{\tan\left(\frac{\alpha_\psi}{2}\right)^2}} - \frac{W-1}{2} s \times \sin(\varphi) + \frac{H-1}{2} (1 - s \times \cos(\varphi)) \right)}{2^{L-1}} \\
 a_5^{opt} = s \times \sin(\varphi) \\
 a_6^{opt} = s \times \cos(\varphi) - 1
 \end{cases}$$

où L est le nombre de niveaux d'une pyramide de réduction successive de résolution d'image dans le cas d'un procédé d'estimation de mouvement multi-résolution ou est égal à 1 sinon ; (ψ , θ , φ ,) sont les angles d'Euler indiqués par les informations recueillies, φ représentant en outre la rotation de la caméra autour de sa ligne de visée, les images ayant W pixels en largeur et H pixels en hauteur.

4. Programme d'ordinateur comportant des instructions logicielles qui, lorsqu'elles sont exécutées par un ordinateur, mettent en œuvre un procédé selon l'une quelconque des revendications précédentes.

5. Dispositif électronique d'estimation de mouvement entre deux images (I_1 , I_2) comprenant une entrée pour recueillir les données définissant chacune des deux images et étant adapté pour déterminer un vecteur de déplacement d'initialisation caractérisant un déplacement entre lesdites deux images et pour mettre en œuvre, à partir du vecteur de déplacement d'initialisation, un algorithme itératif délivrant à chaque itération un vecteur de déplacement actualisé en fonction d'une fonction d'erreur à minimiser entre l'une des deux images et la transformée de l'autre image par un déplacement défini par le vecteur de déplacement actualisé,

- ledit dispositif étant caractérisé en ce qu'il est adapté pour recueillir des informations définissant le mouvement de l'appareil de capture d'images entre les deux images, pour recueillir la valeur de paramètre(s) de capture d'un appareil de capture d'images utilisé(s) pour la capture des deux images parmi une ouverture angulaire, un facteur de zoom et la résolution des images et pour déterminer le vecteur de déplacement d'initialisation par projection, dans le plan image de l'appareil de capture d'image, du déplacement entre les deux images calculé en fonction des informations recueillies définissant le mouvement de l'appareil de

16

capture d'images entre les deux images et en fonction de la valeur de chaque paramètre de capture.

6. Dispositif électronique d'estimation de mouvement entre deux images (I_1 , I_2) selon la revendication 5, adapté pour prendre le vecteur de déplacement d'initialisation égal à $X_0^{opt} = (a_1^{opt}, a_2^{opt}, a_3^{opt}, a_4^{opt}, a_5^{opt}, a_6^{opt})^T$, où

5

$$\begin{cases} a_1^{opt} = \frac{\left(\frac{W-1}{2} \left(1 + \frac{\tan(\psi)}{\tan\left(\frac{\alpha_\psi}{2}\right)} - s \times \cos(\varphi) \right) + \frac{H-1}{2} s \times \sin(\varphi) \right)}{2^{L-1}} \\ a_2^{opt} = s \times \cos(\varphi) - 1 \\ a_3^{opt} = -s \times \sin(\varphi) \\ a_4^{opt} = \frac{\left(\tan(\theta) \times \sqrt{\frac{(H-1)^2}{4 \tan\left(\frac{\alpha_\theta}{2}\right)^2} + \frac{(W-1)^2}{4} \frac{\tan(\psi)^2}{\tan\left(\frac{\alpha_\psi}{2}\right)^2}} - \frac{W-1}{2} s \times \sin(\varphi) + \frac{H-1}{2} (1 - s \times \cos(\varphi)) \right)}{2^{L-1}} \\ a_5^{opt} = s \times \sin(\varphi) \\ a_6^{opt} = s \times \cos(\varphi) - 1 \end{cases}$$

où L est le nombre de niveaux d'une pyramide de réduction successive de résolution d'image dans le cas d'un procédé d'estimation de mouvement multi-résolution ou est égal à 1 sinon ; (ψ , θ , φ ,) sont les angles d'Euler indiqués par les informations recueillies, φ représentant en outre la rotation de la caméra autour de sa ligne de visée, les images ayant W pixels en largeur et H pixels en hauteur.

10