

(19) RÉPUBLIQUE FRANÇAISE  
—  
INSTITUT NATIONAL  
DE LA PROPRIÉTÉ INDUSTRIELLE  
—  
COURBEVOIE

(11) Nº de publication :  
(à n'utiliser que pour les  
commandes de reproduction)

**3 083 352**

(21) Nº d'enregistrement national :

**18 55947**

(51) Int Cl<sup>8</sup> : **G 06 K 9/00** (2018.01), B 60 W 30/00, G 06 T 7/593

(12)

## BREVET D'INVENTION

**B1**

(54) PROCÉDÉ ET DISPOSITIF DE DETECTION RAPIDE DE STRUCTURES REPETITIVES  
DANS L'IMAGE D'UNE SCÈNE ROUTIERE.

(22) Date de dépôt : 29.06.18.

(60) Références à d'autres documents nationaux  
apparentés :

(30) Priorité :

Demande(s) d'extension :

(43) Date de mise à la disposition du public  
de la demande : 03.01.20 Bulletin 20/01.

(71) Demandeur(s) : *CONTINENTAL AUTOMOTIVE*  
*FRANCE Société par actions simplifiée* —FR et  
*CONTINENTAL AUTOMOTIVE GMBH* — DE.

(45) Date de la mise à disposition du public du  
brevet d'invention : 05.06.20 Bulletin 20/23.

(72) Inventeur(s) : LUGEZ BORIS, GARCIA LUCIEN et  
FAURE ARNAUD.

(56) Liste des documents cités dans le rapport de  
recherche :

*Se reporter à la fin du présent fascicule*

(73) Titulaire(s) : *CONTINENTAL AUTOMOTIVE*  
*FRANCE Société par actions simplifiée*,  
*CONTINENTAL AUTOMOTIVE GMBH*.

(74) Mandataire(s) : *CONTINENTAL AUTOMOTIVE*  
*FRANCE Société par actions simplifiée*.

**FR 3 083 352 - B1**



L'invention se rapporte au domaine de l'assistance à la conduite d'un véhicule automobile et concerne plus particulièrement un procédé et un dispositif de traitement d'images acquises par une caméra d'un véhicule automobile ainsi qu'un véhicule automobile comprenant un tel dispositif.

5 De nos jours, il est connu d'équiper un véhicule automobile d'un système d'assistance à la conduite appelé communément ADAS (Advanced Driver Assistance System en langue anglaise). Un tel système comprend de manière connue une caméra vidéo, par exemple montée sur le pare-brise avant du véhicule, qui permet de générer un flux d'images représentant l'environnement dudit véhicule. Ces images sont exploitées  
10 par un dispositif de traitement d'images dans le but d'assister le conducteur, par exemple en détectant un obstacle ou bien le franchissement d'une ligne blanche. Les informations fournies par les images acquises par la caméra doivent donc être suffisamment précises et pertinentes pour permettre au système d'assister le conducteur de manière fiable.

15 En particulier, pour exploiter ces images, il est nécessaire de suivre le flot optique (« optical flow » en langue anglaise) d'un certain nombre de points dans la série d'images. Plus précisément, il s'agit de suivre la trajectoire de ces points dans une séquence d'images, c'est-à-dire dans une pluralité d'images successives, et de déterminer leur localisation dans l'espace.

20 Afin de suivre un point représenté par un pixel d'une première image dans une deuxième image consécutive, une solution connue consiste à analyser une zone de pixels autour dudit pixel dans la deuxième image afin de déterminer dans quelle direction le point se déplace entre la première image et la deuxième image.

25 Cependant, ce procédé de traitement d'images présente un inconvénient majeur. En effet, pour un point donné d'un motif identique qui se répète plusieurs fois sur chaque image d'une séquence d'images, le procédé détecte une pluralité de signatures similaires sur un même segment dit « épipolaire » selon lequel sont disposés les objets correspondant à ces motifs, par exemple une pluralité de poteaux d'une glissière de sécurité installée le long d'une route. Autrement dit, ce procédé de traitement d'images propose plusieurs solutions possibles d'emplacement d'un même point d'un motif répétitif  
30 dans chaque image de la séquence d'images. Ce faisant, les informations données au conducteur concernant la distance entre l'objet et le véhicule peuvent s'avérer fausses, ce qui peut présenter un risque pour la sécurité des occupants du véhicule.

Il existe donc le besoin d'une solution permettant de remédier au moins en partie à ces inconvénients en résolvant ce type d'ambiguïté.

A cette fin, l'invention a pour objet un procédé de traitement d'images permettant l'identification d'au moins un même point apparaissant sur chaque image d'une séquence d'images, ladite séquence d'images comprenant une première image, une deuxième image et une troisième image générées respectivement à un premier instant, à un deuxième instant, postérieur au premier instant, et à un troisième instant, postérieur au deuxième instant, par une caméra montée dans un véhicule automobile, ledit procédé étant remarquable en ce qu'il comprend les étapes de :

• détection d'une pluralité de points identiques dans chaque image de la séquence d'images relative à une pluralité de motifs répétitifs dans chaque image de la séquence d'images,

• pour chaque point identique détecté dans la deuxième image de la séquence d'images, calcul de la profondeur au premier instant du point de l'espace correspondant au dit point identique de la deuxième image,

• pour chaque point identique détecté dans la troisième image de la séquence d'images, calcul de la profondeur au premier instant du point de l'espace correspondant au dit point identique de la troisième image,

• identification, dans la deuxième image et dans la troisième image, des points identiques correspondants aux points de l'espace pour lesquels la profondeur calculée est identique, ces points de l'espace représentant respectivement dans la deuxième image et dans la troisième image un même et unique point réel.

Le procédé selon l'invention permet avantageusement de distinguer efficacement et de manière rapide des motifs répétitifs apparaissant dans chaque image d'une séquence d'images afin notamment de permettre une assistance fiable au conducteur du véhicule.

De préférence, la profondeur au premier instant du point de l'espace correspondant au dit point identique de la deuxième image est calculée à partir de la matrice de rotation et du vecteur de translation entre la première image et la deuxième image.

De préférence encore, la profondeur au premier instant du point de l'espace correspondant au dit point identique de la troisième image est calculée à partir de la matrice de rotation et du vecteur de translation entre la première image et la troisième image.

Avantageusement, le procédé comprend une étape de mémorisation, après l'étape de calcul, de la profondeur pour chaque point identique détecté dans la deuxième image de la séquence d'images.

Avantageusement encore, le procédé comprend une étape de mémorisation, après l'étape de calcul, de la profondeur pour chaque point identique détecté dans la troisième image de la séquence d'images.

De telles mémorisations permettent de sauvegarder les résultats obtenus par 5 les calculs de profondeur.

L'invention concerne également un module de traitement d'images permettant l'identification d'au moins un même point apparaissant sur chaque image d'une séquence d'images, ladite séquence d'images comprenant une première image, une deuxième image et une troisième image générées respectivement à un premier instant, à un 10 deuxième instant, postérieur au premier instant, et à un troisième instant, postérieur au deuxième instant, par une caméra montée dans un véhicule automobile, ledit module de traitement d'images étant caractérisé en ce qu'il est configuré pour:

- détecter une pluralité de points identiques dans chaque image de la séquence d'images relative à une pluralité de motifs répétitifs dans chaque image de la 15 séquence d'images,
- calculer de la profondeur au premier instant du point de l'espace correspondant au dit point identique de la deuxième image pour chaque point identique détecté dans la deuxième image de la séquence d'images,
- calculer la profondeur au premier instant du point de l'espace correspondant au 20 dit point identique de la troisième image, pour chaque point identique détecté dans la troisième image de la séquence d'images,
- identifier dans la deuxième image et dans la troisième image, des points identiques correspondants aux points de l'espace pour lesquels la distance calculée est identique, ces points de l'espace représentant respectivement dans 25 la deuxième image et dans la troisième image un même et unique point réel.

De préférence, le module de traitement d'images est configuré pour calculer la profondeur au premier instant du point de l'espace correspondant au dit point identique de la deuxième image à partir de la matrice de rotation et du vecteur de translation entre la première image et la deuxième image.

30 De préférence encore, le module de traitement d'images est configuré pour calculer la profondeur au premier instant du point de l'espace correspondant au dit point identique de la troisième image à partir de la matrice de rotation et du vecteur de translation entre la première image et la troisième image.

Avantageusement, le module de traitement d'images est configuré pour 35 mémoriser la profondeur pour chaque point identique détecté dans la deuxième image de la séquence d'images et la profondeur pour chaque point identique détecté dans la troisième image de la séquence d'images.

L'invention concerne aussi un véhicule automobile comprenant un module de traitement d'images tel que présenté précédemment.

- D'autres caractéristiques et avantages de l'invention apparaîtront lors de la description qui suit faite en regard des figures annexées données à titre d'exemples non limitatifs et dans lesquelles des références identiques sont données à des objets semblables.
- La **figure 1** illustre schématiquement un véhicule selon l'invention.
  - La **figure 2** illustre schématiquement une première image et une deuxième image, consécutive à la première image, acquises par la caméra.
  - La **figure 3** illustre schématiquement deux projections d'une zone de pixels d'une première image dans une deuxième image par homographie sur un plan de projection situé respectivement à une distance minimum et à une distance maximum prédéterminées.
  - La **figure 4** illustre schématiquement une projection par homographie, sur un plan de projection situé à une distance prédéterminée, d'un ensemble de points d'une première image dans une deuxième image, la correspondance entre les deux images étant réalisée via une matrice de rotation et un vecteur de translation de la caméra.
  - La **figure 5** illustre schématiquement un segment épipolaire obtenu dans une deuxième image par la projection d'un ensemble de pixels d'une première image sur un plan situé consécutivement à une pluralité de distances prédéterminées comprises entre une distance minimum prédéterminée et à une distance maximum prédéterminée.
  - La **figure 6** illustre schématiquement le déplacement d'un véhicule par rapport à une pluralité de structures répétitives.
  - La **figure 7** représente le modèle sténopé, utilisé ici pour décrire le processus de formation des images par une caméra.
  - La **figure 8** illustre un mode de réalisation de l'étape de détection du procédé de traitement d'images selon l'invention.
  - La **figure 9** illustre un mode de réalisation du procédé de traitement d'images selon l'invention.

Le dispositif de traitement d'images selon l'invention va maintenant être décrit en référence aux **figures 1 à 8**.

En référence à la **figure 1**, le dispositif 10 de traitement d'images est monté dans un véhicule 1 automobile. Par les termes « véhicule automobile », on entend de manière standard un véhicule routier mû par un moteur à explosion, à combustion interne,

électrique ou par turbine à gaz ou un moteur hybride tel que, par exemple, une voiture, une camionnette, un camion, etc.

Le dispositif 10 comprend une caméra 100 vidéo et un module de traitement d'images 200, par exemple mis en œuvre par un calculateur (non représenté) embarqué dans le véhicule 1. De manière préférée, la caméra 100 est montée au niveau de la partie supérieure centrale du pare-brise avant (non représenté) du véhicule 1. On notera toutefois que la caméra 100 pourrait être montée à tout autre endroit adapté du véhicule 1 (côté, arrière, etc.). La caméra 100 est configurée pour générer une pluralité d'images de l'environnement 2 du véhicule 1, par exemple la route 2A et ses accotements, afin que le module de traitement d'images 200 les exploite.

Le module de traitement d'images 200 traite les images générées par la caméra 100 afin de fournir une assistance à la conduite au conducteur du véhicule 1. A titre d'exemple, un tel traitement d'images peut consister à détecter des objets dans les images, tels que par exemple des panneaux, des bordures de route, des obstacles (piétons ou autre...), afin d'informer le conducteur respectivement d'une vitesse limite, d'un risque de sortie de route ou d'un risque de collision. Un tel traitement d'images peut aussi permettre de mesurer la hauteur d'un pont ou la position d'un obstacle afin d'en avertir le conducteur ou de procéder à un freinage d'urgence.

Dans le cas où des motifs répétitifs apparaissent sur les images successives générées par la caméra 100 (i.e. des pixels identiques apparaissant à plusieurs endroits des images), il peut être utile de les distinguer d'une image à la suivante pour éviter de provoquer des erreurs d'interprétation qui pourraient présenter un risque pour la conduite du véhicule 1. Ces motifs répétitifs peuvent par exemple correspondre à des poteaux d'une glissière d'autoroute défilant le long de la trajectoire du véhicule 1 de sorte que plusieurs d'entre eux apparaissent sur des images successives générées par la caméra 100.

Avant de pouvoir distinguer des groupes de pixels représentant des motifs répétitifs (donc identiques mais apparaissant à plusieurs endroits sur chaque image de la séquence d'images successives), il est tout d'abord nécessaire de détecter les points ou groupes de pixels correspondant à ces motifs répétitifs dans la séquence d'images.

Une forme de réalisation préférée permettant de détecter les points ou groupes de pixels correspondant à des motifs répétitifs dans la séquence d'images va maintenant être décrite. On notera toutefois que toute autre méthode de détection adaptée pourrait être utilisée sans que cela ne limite la portée de la présente invention.

Dans la forme de réalisation décrite à titre d'exemple, en référence à la **figure 2**, le module de traitement d'images 200 est configuré pour réaliser une première série de tâches. A titre d'exemple, cette première série de tâches est décrite ci-après en

référence à trois images successives d'une séquence d'images : une première image I1, une deuxième image I2, postérieure à la première image I1, et une troisième image I3 (non représentée), postérieure à la deuxième image I2, ces images ayant été acquises par la caméra 100 respectivement à un premier instant t1, un deuxième instant t2 et un 5 troisième instant t3, de préférence consécutifs.

Tout d'abord, le module de traitement d'images 200 est configuré pour déterminer, pour chaque pixel de la première image I1, pour chaque pixel de la deuxième image I2 et pour chaque pixel de la troisième image I3 (non représentée), une signature caractéristique dudit pixel.

10 Par ailleurs, en référence à la **figure 3**, le module de traitement d'images 200 est configuré pour déterminer une matrice de rotation R et un vecteur de translation T de la caméra 100 entre deux images consécutives, ici entre la première image I1 et la deuxième image I2, ainsi qu'entre la deuxième image I2 et la troisième image I3 (non représentée). La matrice de rotation R et le vecteur de translation T correspondent 15 respectivement à la rotation et à la translation de la caméra 100 dans un référentiel lié à l'objet.

La détermination de la matrice de rotation R et du vecteur de translation T peut, par exemple, être réalisée en utilisant un objet statique représenté sur les trois images, par exemple la route ou un pont, ou bien un objet dynamique se déplaçant par 20 rapport à la caméra 100, par exemple un autre véhicule.

Dans ce dernier cas, il faut savoir où se situe l'objet dans l'image (par exemple en utilisant une méthode de classification connue en soi) et déterminer alors la matrice de rotation R et le vecteur de translation T de la caméra 100 dans le référentiel lié 25 à l'objet, le vecteur de translation T étant dans ce cas pondérée d'un coefficient  $\alpha$  qui peut être déterminé de manière connue par traitement d'images en connaissant la taille de l'objet.

La détermination de la matrice de rotation R et du vecteur de translation T entre la première image I1 et la deuxième image I2 peut être réalisée par un algorithme de traitement d'images ou par l'utilisation d'un ou plusieurs capteurs inertIELS et/ou de 30 position géographique (par exemple GPS) de manière connue en soi.

Ensuite, en référence à la **figure 2**, le module de traitement d'images 200 est configuré pour sélectionner un pixel PA(1) dans la première image I1 pour lequel on cherche un pixel correspondant dans la deuxième image I2. Par les termes « pixel correspondant », on entend un pixel représentant une même zone spatiale de 35 l'environnement 2 du véhicule 1.

Toujours en référence à la **figure 2**, le module de traitement d'images 200 est également configuré pour déterminer dans la première image I1 un ensemble initial de

pixels comprenant le pixel sélectionné PA(1) et une pluralité de pixels PA(2), ..., PA(5), localisés au voisinage du pixel sélectionné PA(1).

Par le terme voisinage, on entend de manière connue une pluralité de pixels PA(2), ..., PA(5) adjacents au pixel sélectionné PA(1) dans la première image I1. Le 5 nombre de pixels du voisinage est de quatre dans cet exemple à des fins de clarté mais il va de soi que le voisinage du pixel sélectionné PA(1) peut comprendre plus ou moins de quatre pixels (plus généralement « n » pixels, « n » étant un entier naturel supérieur ou égal à 2).

Un tel ensemble initial de pixels PA(1), PA(2), ..., PA(5) peut être une fenêtre 10 ou une portion de la première image I1, par exemple une fenêtre de  $10 \times 10$  pixels. On notera ici que le nombre n de pixels et la forme de la fenêtre ou de la portion d'image peuvent varier à dessein. On notera aussi que l'on a représenté cinq pixels dans l'ensemble à titre d'exemple mais que le nombre de pixels de l'ensemble pourrait bien entendu être inférieur ou supérieur à cinq.

15 Le module de traitement d'images 200 est aussi configuré pour déterminer, dans la deuxième image I2, une pluralité d'ensembles de pixels PB(1), ..., PB(5) par projection, à partir de la matrice de rotation R et du vecteur de translation T déterminés, de l'ensemble initial de pixels PA(1), ..., PA(5) déterminé. Chaque ensemble de cette pluralité d'ensembles de pixels PB(1), ..., PB(5) comprend un premier pixel PB(1) 20 correspondant au pixel sélectionné PA(1) dans la première image I1.

En référence aux **figures 3** et **4**, cette projection est réalisée sur un plan W prédéterminé situé à différentes distances D prédéterminées du centre optique C1 de la caméra 100 pour la première image I1.

Plus précisément, en référence à la **figure 3**, le module de traitement 25 d'images 200 est tout d'abord configuré pour déterminer, dans la deuxième image I2, un premier ensemble de pixels PBmin(1), ..., PBmin(5) par projection, à partir de la matrice de rotation R et du vecteur de translation T déterminés, de l'ensemble initial de pixels PA(1), ..., PA(5) déterminé sur un plan W de projection prédéterminé situé à une distance minimale Dmin prédéterminée. Ce premier ensemble PBmin(1), ..., PBmin(5) comprend 30 un premier pixel PBmin(1) correspondant au pixel sélectionné PA(1) dans la première image I1.

Similairement, toujours en référence à la **figure 3**, le module de traitement d'images 200 est configuré pour déterminer, dans la deuxième image I2, un deuxième ensemble de pixels PBmax(1), ..., PBmax(5) par projection, à partir de la matrice de 35 rotation R et du vecteur de translation T déterminées, de l'ensemble initial déterminé PA(1), ..., PA(5) sur le plan W situé à une distance maximale prédéterminée Dmax. Ce

deuxième ensemble  $PB_{max}(1), \dots, PB_{max}(5)$  comprend un deuxième pixel  $PB_{max}(1)$  correspondant au pixel sélectionné  $PA(1)$  dans la première image  $I1$ .

En référence maintenant à la **figure 4**, la détermination de la projection d'un pixel de la première image  $I1$  dans la deuxième image  $I2$  via un plan  $W$  de projection 5 correspond à une homographie du pixel dans un premier repère, qui a pour origine le centre optique  $C1$  de la caméra 100 lors de l'acquisition de la première image  $I1$ , à un deuxième repère qui a pour origine le centre optique  $C2$  de la caméra 100 lors de l'acquisition de la deuxième image  $I2$ , via le plan  $W$  de projection situé à l'une des distances déterminées. Une telle projection par homographie est réalisée à partir de la 10 matrice de rotation  $R$  et du vecteur de translation  $T$ , d'un vecteur  $n_2$  normal au plan  $W$  de projection et de la distance  $D$  du plan  $W$  par rapport au centre optique  $C1$  de la manière suivante :  $H = R - \left(\frac{T \times n_2^T}{D}\right)$ .

Une telle homographie étant connue en soi, elle ne sera donc pas davantage détaillée ici.

15 En théorie, la distance minimale  $D_{min}$  est nulle et la distance maximale  $D_{max}$  correspond à l'infini. En pratique, on peut choisir une distance minimale  $D_{min}$  de quelques mètres, par exemple 5 mètres, et une distance maximale  $D_{max}$  de quelques dizaines de mètres, par exemple 50 ou 100 mètres, notamment lorsque le pixel sélectionné  $PA(1)$  dans la première image  $I1$  représente un objet identifié dont on connaît l'éloignement par 20 rapport à la caméra 100, par exemple la route défilant dans la partie inférieure des images acquises par la caméra 100.

Le plan  $W$  de projection peut être déterminé selon la nature du pixel sélectionné  $PA(1)$ . Par exemple, lorsque le pixel sélectionné  $PA(1)$  représente une zone de la route 2A sur laquelle circule le véhicule 1, un plan  $W$  correspondant au plan de la 25 route 2A peut être utilisé comme illustré schématiquement sur la **figure 2**. Ce plan  $W$  est un plan de projection virtuel qui va permettre d'estimer par homographie, c'est-à-dire de manière géométrique, la position correspondante d'un pixel de la première image  $I1$  dans la deuxième image  $I2$ .

Une fois le premier pixel  $PB_{min}(1)$  et le deuxième pixel  $PB_{max}(1)$  30 sélectionnés, le module de traitement d'images 200 est configuré pour déterminer un segment  $U$  reliant le premier pixel  $PB_{min}(1)$  et le deuxième pixel  $PB_{max}(1)$ , appelé segment épipolaire.

En référence à la **figure 5**, le module de traitement d'images 200 est ensuite configuré pour sélectionner une pluralité de points  $PB_2(1), \dots, PB_5(1)$  le long du segment 35  $U$  déterminé. De manière avantageuse, les points  $PB_2(1), \dots, PB_5(1)$  peuvent être répartis le long du segment  $U$  en partant du premier pixel  $PB_{min}(1)$  en étant espacés d'un pas « step » de largeur prédéterminée, par exemple tous les 0,1 pixels. Le nombre de points

$PB_2(1), \dots, PB_5(1)$  sélectionnés sur le segment U est de quatre dans cet exemple mais il va de soi qu'il pourrait être supérieur ou inférieur à quatre.

Le module de traitement d'images 200 est configuré pour associer chaque point  $PB_2(1), \dots, PB_5(1)$  sélectionné à un pixel de la deuxième image I2. En pratique, on 5 considère qu'un point est associé lorsque ses coordonnées dans l'image correspondent aux coordonnées d'un pixel, c'est-à-dire que le point sélectionné se trouve dans la zone de l'image couverte (ou représentée) par ledit pixel.

Le module de traitement d'images 200 est configuré pour comparer la signature de chaque pixel correspondant à un point  $PB_2(1), \dots, PB_5(1)$  sélectionné avec la 10 signature du pixel sélectionné dans la première image I1.

Le module de traitement d'images 200 est configuré pour détecter un motif répétitif entre la première image I1 et la deuxième image I2 lorsqu'une pluralité de signatures de pixels associés aux points  $PB_2(1), \dots, PB_5(1)$  sélectionnés sur le segment U correspondent à la signature du pixel sélectionné dans la première image I1.

15 Le module de traitement d'images 200 est également configuré pour réaliser les tâches décrites ci-avant entre la deuxième image I2 et la troisième image I3 (non représentée) et plus généralement entre deux images successives d'une séquence d'images de manière à détecter des motifs répétitifs dans sur chaque image de la séquence d'images I1, I2, I3.

20 Ensuite, afin d'identifier un point ou un groupe de pixel correspondant à un même motif sur deux images successives, parmi une pluralité de motifs successifs, le module de traitement d'images 200 est configuré pour réaliser une deuxième série de tâches.

Tout d'abord, le module de traitement d'images 200 est configuré pour 25 détecter une pluralité de points identiques dans chaque image de la séquence d'images I1, I2, I3 relative à une pluralité de motifs répétitifs dans chaque image de la séquence d'images I1, I2, I3.

Le module de traitement d'images 200 est configuré, pour chaque point identique détecté dans la deuxième image I2 de la séquence d'images I1, I2, I3, pour 30 calculer la profondeur au premier instant  $t_1$  du point de l'espace correspondant au dit point identique de la deuxième image I2.

Le module de traitement d'images 200 est configuré, pour chaque point identique détecté dans la troisième image I3 de la séquence d'images I1, I2, I3, pour calculer la profondeur au premier instant  $t_1$  du point de l'espace correspondant au dit 35 point identique de la troisième image I3 (non représentée).

Le module de traitement d'images 200 est configuré pour identifier, dans la deuxième image I2 et dans la troisième image I3 (non représentée), des points identiques

correspondants aux points de l'espace pour lesquels la profondeur calculée est identique, ces points de l'espace représentant respectivement dans la deuxième image I2 et dans la troisième image I3 (non représentée) un même et unique point réel.

L'invention va maintenant être décrite dans sa mise en œuvre en référence  
5 aux **figures 8 et 9**.

Tout d'abord, dans une étape E0, en référence notamment aux **figures 2 et 8**,  
la caméra 100 acquiert une première image I1 et une deuxième image I2, de préférence  
consécutives. On notera que le procédé peut être mis en œuvre en prenant comme  
première image I1 l'image dont l'acquisition est temporellement la première (image  
10 antérieure ou précédente dans le flot d'images acquises par la caméra) et comme  
deuxième image I2, l'image successive dans le flot d'images. A contrario, le procédé peut  
être mis en œuvre en prenant comme première image I1 l'image dont l'acquisition est  
temporellement postérieure (image suivante dans le flot d'images acquises par la caméra)  
à la deuxième image I2 et comme deuxième image I2, l'image précédente dans le flot  
15 d'images. Dans l'exemple qui suit, on choisit l'image antérieure comme première image I1  
et l'image postérieure (i.e. successive dans le flot d'images) comme deuxième image I2.

Dans une étape E1, le module de traitement d'images 200 détermine, pour  
chaque pixel de la première image I1 et pour chaque pixel de la deuxième image I2, une  
signature caractéristique dudit pixel.

20 Une telle signature peut être déterminée à partir de l'analyse locale d'une  
image et caractérise ainsi le contenu de cette image. Il existe plusieurs façons de créer  
une signature basée sur des descripteurs. Ces descripteurs peuvent avoir certaines  
propriétés comme l'invariance par rotation, par grossissement ou changement de  
luminosité. Le cas d'utilisation définit en général les propriétés nécessaires aux  
25 descripteurs et permet de déterminer la méthode la plus avantageuse. Par exemple, la  
transformation de Census donne une signature basée sur les différences relatives  
d'intensité (Zabih, R., Woodfill, J.I.: Non-parametric Local Transforms for Computing  
Visual Correspondence. In: Eklundh, J.-O. (ed.) ECCV 1994. LNCS, vol. 801, pp. 151–  
158. Springer, Heidelberg (1994)). Un autre exemple est la transformation de  
30 caractéristiques visuelles invariante à l'échelle (SIFT : Scale-Invariant Feature Transform  
en langue anglaise).

En vision par ordinateur, l'extraction de caractéristiques visuelles (ou visual  
features extraction en anglais) consiste en des transformations mathématiques calculées  
sur les pixels d'une image numérique. Les caractéristiques visuelles permettent  
35 généralement de mieux rendre compte de certaines propriétés visuelles de l'image,  
utilisées pour des traitements ultérieurs entrant dans le cadre d'applications telles que la  
détection d'objets ou la recherche d'images par le contenu.

La caractérisation d'une image peut être calculée en un nombre restreint de pixel. Il faut pour cela d'abord détecter les zones d'intérêt de l'image puis calculer en chacune de ces zones un vecteur caractéristique. Ces zones d'intérêt sont par exemple les arêtes ou les points saillants de l'image (zones de fort contraste). Il peut aussi s'agir 5 de points pris aléatoirement ou régulièrement dans l'image (échantillonnage dit dense).

Le vecteur caractéristique contient parfois des données provenant de la détection, telles que l'orientation de l'arête ou la magnitude du gradient dans la zone d'intérêt. Généralement, le vecteur caractéristique en un pixel est calculé sur un voisinage de ce pixel, c'est-à-dire à partir d'une imagette centrée sur ce pixel. Il peut être calculé à 10 différentes échelles de manière à s'affranchir du facteur de zoom. Parmi les caractéristiques locales couramment calculées, on retrouve des motifs préalablement utilisés globalement, tels que des histogrammes de couleur ou des vecteurs rendant compte de l'orientation des gradients des niveaux de gris. Trois exemples de voisinages utilisés pour définir une texture et calculer un motif binaire local (LBP). Certaines 15 méthodes telles SIFT ou SURF incluent à la fois la détection de zone d'intérêt et le calcul d'un vecteur caractéristique en chacune de ces zones. Concernant le vecteur caractéristique, les SIFT sont grossièrement un histogramme des orientations du gradient et les SURF consistent en le calcul d'approximation d'ondelettes de Haar. Dans une veine similaire, les motifs binaires locaux sont un calcul de cooccurrence des niveaux de gris 20 locaux d'une image et les HOG des histogrammes de gradients orientés, assez similaires aux SIFT, calculés selon un échantillonnage dense.

En référence à la **figure 3**, le module de traitement d'images 200 détermine ensuite, dans une étape E2, la matrice de rotation R et le vecteur de translation T entre la première image I1 et la deuxième image I2.

25 Le module de traitement d'images 200 sélectionne, dans une étape E3, un pixel sélectionné PA(1) dans la première image I1 pour lequel un pixel correspondant dans la deuxième image I2 doit être déterminé.

De manière avantageuse, le module de traitement d'images 200 peut tout d'abord sélectionner un pixel, qui n'a pas de flot optique dans la deuxième image I2, dans 30 une zone d'intérêt de la première image I1, par exemple un côté de l'image représentant le bord de la route lorsqu'on souhaite prévenir le conducteur d'un risque de sortie de route, le dessus de l'image représentant le ciel lorsqu'on souhaite détecter un pont pour en évaluer la hauteur etc.

Ensuite, en référence à la **figure 2**, pour le pixel sélectionné PA(1), le module 35 de traitement d'images 200 détermine dans la première image I1, dans une étape E4, un ensemble initial de pixels PA(1), ..., PA(5) comprenant le pixel sélectionné PA(1) et une pluralité de pixels PA(2), ..., PA(5) localisés au voisinage du pixel sélectionné PA(1).

Dans l'exemple de la **figure 2**, par souci de clarté, seuls quatre pixels PA(2), PA(3), PA(4) et PA(5) du voisinage du pixel sélectionné PA(1) ont été sélectionnés.

En référence à la **figure 3**, le module de traitement d'images 200 détermine ou sélectionne (par exemple dans une liste stockée dans une zone mémoire (non représentée) du dispositif 10) dans une étape E5 une distance minimale Dmin et une distance maximale Dmax entre lesquelles se situe le point spatial représenté par le pixel sélectionné PA(1). Ces valeurs Dmin et Dmax peuvent être déterminées comme décrit ci-avant ou bien en utilisant l'historique des images précédentes ou une connaissance géométrique de la scène, comme par exemple une route plane.

10 Dans une étape E6, le module de traitement d'images 200 détermine un plan W de projection, du pixel sélectionné PA(1), dans la deuxième image I2, par exemple un plan vertical situé face à l'avant du véhicule 1. On notera que les étapes E5 et E6 peuvent être inversées.

15 En référence aux **figures 2** et **3**, dans une étape E7, le module de traitement d'images 200 détermine, dans la deuxième image I2, un premier ensemble de pixels PBmin(1), ..., PBmin(5) par projection homographique de l'ensemble initial de pixels PA(1), ..., PA(5) déterminé sur le plan W, en considérant que ce dernier est placé à la distance minimale déterminée Dmin. Ce premier ensemble PBmin(1), ..., PBmin(5) comprend un premier pixel PBmin(1) correspondant, par projection homographique sur le 20 plan W à la distance minimale Dmin, au pixel sélectionné PA(1) dans la première image I1.

25 Similairement, dans une étape E8, le module de traitement d'images 200 détermine, dans la deuxième image I2, un deuxième ensemble de pixels PBmax(1), ..., PBmax(5) par projection homographique de l'ensemble initial PA(1), ..., PA(5) déterminé sur le plan W pré-déterminé situé à la distance maximale Dmax pré-déterminée. Ce deuxième ensemble PBmax(1), ..., PBmax(5) comprend un deuxième pixel PBmax(1) correspondant, par projection homographique du plan W à la distance maximale Dmax, au pixel sélectionné PA(1) dans la première image I1. Il va de soi que les étapes E7 et E8 peuvent être inversées.

30 En référence aux **figures 3** à **5**, le module de traitement d'images 200 détermine ensuite, dans une étape E9, le segment U reliant le premier pixel PBmin(1) et le deuxième pixel PBmax(1) dans la deuxième image I2 puis sélectionne, dans une étape E10, une pluralité de points PB<sub>2</sub>(1), ..., PB<sub>5</sub>(1) sur le segment U ainsi déterminé, par exemple tous les 0,1 pixels.

35 Au lieu d'un unique segment, on pourrait également considérer un cylindre épipolaire qui se définit comme étant un segment épipolaire d'une épaisseur supérieure à

un pixel (de préférence une épaisseur d'un nombre impair de pixels, ainsi la dilatation du segment se fait de part et d'autre du segment d'épaisseur 1).

Le module de traitement d'images 200 associe ensuite dans une étape E11 chaque point  $PB_2(1), \dots, PB_5(1)$  sélectionné sur le segment U à un pixel de la deuxième image I2.

Le module de traitement d'images 200 compare alors dans une étape E12 la signature de chaque pixel associé à un point  $PB_2(1), \dots, PB_5(1)$  sélectionné avec la signature du pixel sélectionné dans la première image I1.

Enfin, dans une étape E13, le module de traitement d'images 200 détecte un point d'un motif répétitif entre la première image I1 et la deuxième image I2 lorsqu'une pluralité de signatures de pixels associés à la pluralité de points  $PB_2(1), \dots, PB_5(1)$  sélectionnés sur le segment U correspondent à la signature du pixel sélectionné dans la première image I1.

Dans l'exemple non limitatif qui a été décrit ci-avant, un unique segment épipolaire était utilisé mais on notera qu'en variante, une pluralité de segments épipolaires, par exemple parallèles, pourrait être utilisée, chaque segment étant défini par une pluralité de points associés à des pixels de la deuxième image I2 de sorte que la signature de chaque point soit comparée à la signature du pixel associé dans la première image I1 afin de détecter de manière encore plus précise des motifs répétitifs dans les images. De même, on pourrait définir, plutôt qu'un segment ou une pluralité de segments, un cylindre épipolaire comprenant le segment U et une pluralité de points associés à des pixels de la deuxième image I2 et dont on pourrait comparer les signatures avec les pixels correspondant de la première image I1 afin de détecter de manière encore plus précise des motifs répétitifs dans les images.

On notera par ailleurs que pour les objets dynamiques, la rotation et translation entre les images sont de préférence exprimés par rapport à un repère monde qui est lié à l'objet dynamique.

Ensuite, le module de traitement d'images 200 va identifier les points ou groupes de pixels correspondants à des motifs répétitifs afin de les distinguer entre eux sur une même image. Autrement dit, le module de traitement d'images 200 va calculer la profondeur des points de l'espace représentés dans les motifs répétitifs sur chaque image de la séquence d'images I1, I2 et I3 consécutives.

On notera tout d'abord qu'à ce stade, la matrice de rotation R et le vecteur de translation T subies par la caméra 100 entre le premier instant  $t_1$  et le deuxième instant  $t_2$  ont été déterminés lors de l'étape E1 avec la méthode décrite ci-avant et sont connus. De même, la matrice de rotation R et le vecteur de translation T subies par la caméra 100

entre le deuxième instant  $t_2$  et le troisième instant  $t_3$  ont été déterminés lors de la détection des motifs répétitifs décrite ci-avant et sont connus.

En référence à la **figure 6**, il est représenté l'exemple d'un motif répétitif généré par deux points de l'espace A et B qui ont la même signature dans chaque image de la séquence d'images  $I_1$ ,  $I_2$  et  $I_3$  consécutives, cet exemple pouvant être généralisé à autant de points que le motif répétitif en comprend. Par ailleurs, il est aussi représenté le mouvement de la caméra 100 en fonction de ces points de l'espace A et B. En effet, les trois flèches verticales représentent l'emplacement de la caméra 100, et donc du véhicule, ainsi que le sens et la direction de déplacement dudit véhicule, aux trois instants consécutifs  $t_1$ ,  $t_2$ ,  $t_3$ .

Tout d'abord, en référence à la **figure 7**, il est représenté le modèle sténopé, utilisé ici pour décrire le processus de formation des images par la caméra 100. Autrement dit, le modèle sténopé permet de décrire comment projeter au moins un objet 3D (en trois dimensions) sur une image 2D (en deux dimensions).

15 La caméra 100 possède les paramètres intrinsèques suivants :

- un centre optique  $O$ ,
- une distance focale  $f$ ,
- un point principal  $F$ .

Le modèle dit « du sténopé », connu en soi, est ici représenté dans un repère orthonormé  $(O, \vec{i}, \vec{j}, \vec{k})$  en trois dimensions, où le centre optique  $O$  est l'origine du repère. 20 Dans ce repère orthonormé  $(O, \vec{i}, \vec{j}, \vec{k})$ , le point principal  $F$  possède les coordonnées  $(0, 0, f)$ .

Chaque image formée par la caméra 100 se situe dans un plan appelée « plan de l'image »  $xOy$  d'origine  $O$ . Le point de l'espace A est un point 3D ayant pour 25 coordonnées cartésiennes  $(X_A(t_2), Y_A(t_2), Z_A(t_2))$  dans le repère orthonormé  $(O, \vec{i}, \vec{j}, \vec{k})$  au deuxième instant  $t_2$ . Le point  $P_2$  en deux dimensions appartient à la deuxième image  $I_2$  et représente la projection du point de l'espace A de l'objet réel dans le plan de l'image  $xOy$  dans lequel se situe ladite deuxième image  $I_2$ . Ledit point  $P_2$  possède pour 30 coordonnées  $(x_{b_2}(t_2), y_{b_2}(t_2), f)$  au deuxième instant  $t_2$  dans le repère orthonormé  $(O, \vec{i}, \vec{j}, \vec{k})$  avec :

$$x_{b_2}(t_2) = f \times \frac{X_A(t_2)}{Z_A(t_2)} \quad [1]$$

$$y_{b_2}(t_2) = f \times \frac{Y_A(t_2)}{Z_A(t_2)} \quad [2]$$

Le modèle du sténopé permet ensuite de déterminer la profondeur des points de l'espace A et B puisqu'on connaît la position de la projection dans chaque image de la 35 séquence d'image  $I_1$ ,  $I_2$ ,  $I_3$  des points de l'espace A et B des motifs répétitifs.

En référence à la **figure 9**, lors d'un premier calcul, appelé « Calcul 1 » (étape E14), le module de traitement d'images 200 calcule la profondeur du point de l'espace A dans le repère caméra au premier instant t1 à partir de la deuxième image I2.

Tout d'abord, le module de traitement d'images 200 associe la projection dans 5 la première image I1 du point de l'espace A et la projection dans la deuxième image I2 du point de l'espace A.

Un premier calcul entre un premier instant t1 et un deuxième instant t2 permet de déterminer la profondeur  $Z_A(t1)$  dans le repère caméra, du point de l'espace A (point en trois dimensions) au premier instant t1 à partir des coordonnées 2D du point 10 P1 ( $x_{b1}(t1); y_{b1}(t1); f$ ) qui correspond à la projection du point de l'espace A dans l'image au premier instant t1, des coordonnées 2D du point P2 de la projection du point A au deuxième instant t2 et de la connaissance de la matrice de rotation R et du vecteur de translation T de la caméra entre le premier instant t1 et le deuxième instant t2.

Précédemment, la matrice de rotation R et le vecteur de translation T ont été 15 déterminés et permettent ainsi de calculer pour tout point de l'espace A au premier instant t1, noté A(t1), le point de l'espace A correspondant au deuxième instant t2, noté A(t2), selon la formule suivante :

$$A(t2) = R \times A(t1) + T \quad [3]$$

Sachant que :

$$20 \quad T = \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix} \text{ et } R = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}$$

L'étape E13 a permis la détection d'une pluralité de points identiques, définie comme une pluralité de points représentant le même point de l'espace dans chaque image de la séquence d'images I1, I2, I3, relative à une pluralité de motifs répétitifs dans chaque image de la séquence d'images I1, I2, I3. Ensuite, on sélectionne un point P2 au 25 deuxième instant t2, de coordonnées ( $x_{b2}(t2); y_{b2}(t2); f$ ) dans la deuxième image I2. Ledit point P2 correspond à un des points identiques détectés dans la deuxième image I2 de la séquence d'images I1, I2, I3. Il s'en suit le calcul de la profondeur entre la caméra 100 et le point de l'espace correspondant au dit point P2 identique de la deuxième image I2 au premier instant t1.

30 Grâce à la matrice de rotation R et au vecteur de translation T on peut déterminer la position dudit point de l'espace A dans la première image I1 au premier instant t1, c'est-à-dire à l'instant précédent le deuxième instant t2. Ainsi, pour déterminer la position dudit point de l'espace A, il faut calculer  $Z_A(t1)$ , profondeur du point de l'espace A, pour en déduire la distance entre le centre optique O de la caméra 100 et le point de 35 l'espace A.

Pour cela, l'équation [3] est développée et on isole les termes  $X_A(t2)$ ,  $Y_A(t2)$ ,  $Z_A(t2)$ , ce qui permet d'obtenir les expressions de  $X_A(t2)$ ,  $Y_A(t2)$  et  $Z_A(t2)$  en fonction de  $X_A(t1)$ ,  $Y_A(t1)$ ,  $Z_A(t1)$  :

$$5 \quad \begin{aligned} X_A(t2) &= R_{11} \times X_A(t1) + R_{12} \times Y_A(t1) + R_{13} \times Z_A(t1) + T_x \\ Y_A(t2) &= R_{21} \times X_A(t1) + R_{22} \times Y_A(t1) + R_{23} \times Z_A(t1) + T_y \\ Z_A(t2) &= R_{31} \times X_A(t1) + R_{32} \times Y_A(t1) + R_{33} \times Z_A(t1) + T_z \end{aligned}$$

A partir du résultat précédent et des équations [1] et [2] et en factorisant  $Z_A(t1)$  dans le membre de droite, on obtient :

$$10 \quad \begin{aligned} \frac{x_{b2}(t2) \times Z_A(t2)}{f} &= \left( \frac{R_{11} \times x_{b1}(t1)}{f} + \frac{R_{12} \times y_{b1}(t1)}{f} + R_{13} \right) \times Z_A(t1) + T_x & [4] \\ \frac{y_{b2}(t2) \times Z_A(t2)}{f} &= \left( \frac{R_{21} \times x_{b1}(t1)}{f} + \frac{R_{22} \times y_{b1}(t1)}{f} + R_{23} \right) \times Z_A(t1) + T_y & [5] \\ Z_A(t2) &= \left( \frac{R_{31} \times x_{b1}(t1)}{f} + \frac{R_{32} \times y_{b1}(t1)}{f} + R_{33} \right) \times Z_A(t1) + T_z & [6] \end{aligned}$$

En injectant l'équation [6] dans les équations [4] et [5], on obtient :

$$\begin{aligned} \frac{x_{b2}(t2) \times [a_3 \times Z_A(t1) + T_z]}{f} &= a_1 \times Z_A(t1) + T_x & [4bis] \\ \frac{y_{b2}(t2) \times [a_3 \times Z_A(t1) + T_z]}{f} &= a_2 \times Z_A(t1) + T_y & [5bis] \end{aligned}$$

15 avec :

$$\begin{aligned} a_1 &= \frac{R_{11} \times x_{b1}(t1)}{f} + \frac{R_{12} \times y_{b1}(t1)}{f} + R_{13} \\ a_2 &= \frac{R_{21} \times x_{b1}(t1)}{f} + \frac{R_{22} \times y_{b1}(t1)}{f} + R_{23} \\ a_3 &= \frac{R_{31} \times x_{b1}(t1)}{f} + \frac{R_{32} \times y_{b1}(t1)}{f} + R_{33} \end{aligned}$$

On obtient deux solutions pour  $Z_A(t1)$  :

$$20 \quad \begin{aligned} Z_{A1}(t1) &= \frac{T_x - \frac{T_z \times x_{b2}(t2)}{f}}{\frac{x_{b2}(t2) \times a_3}{f} - a_1} \\ Z_{A1}(t1) &= \frac{T_y - \frac{T_z \times y_{b2}(t2)}{f}}{\frac{y_{b2}(t2) \times a_3}{f} - a_2} \end{aligned}$$

La profondeur  $Z_{A1}(t1)$  correspond à une première possibilité de profondeur pour le point de l'espace A au premier instant t1. Notons qu'en théorie, ces deux différentes formules de profondeur  $Z_{A1}(t1)$  donnent un résultat identique ou sensiblement identique. Ce résultat est mémorisé dans une étape E15.

Ensuite, le module de traitement d'images 200 associe la projection dans la première image I1 du point de l'espace A et la projection dans la deuxième image I2 du

point de l'espace B. Plus précisément, en référence à la **figure 6**, pour confirmer la profondeur du point de l'espace A, il faut faire un deuxième calcul de profondeur. En effet, on observe alors la projection du point de l'espace B au deuxième instant  $t_2$ . Par ce calcul, on obtiendra une profondeur  $Z_{A2}(t_1)$  étant une deuxième possibilité de profondeur 5 pour le point de l'espace A au premier instant  $t_1$ .

La méthode étant identique à celle utilisée pour la profondeur  $Z_{A1}(t_1)$  du point P1, les formules de profondeur  $Z_{A1}(t_1)$  et  $Z_{A2}(t_1)$  sont identiques aussi. Ainsi, on obtient les résultats suivants :

$$Z_{A2}(t_1) = \frac{T_x - \frac{T_z \times x_{b2}(t_2)}{f}}{\frac{x_{b2}(t_2) \times a_3}{f} - a_1}$$

$$10 \quad Z_{A2}(t_1) = \frac{T_y - \frac{T_z \times y_{b2}(t_2)}{f}}{\frac{y_{b2}(t_2) \times a_3}{f} - a_2}$$

Cependant,  $x_{b2}$  et  $y_{b2}$  sont différents dans les formules des profondeurs  $Z_{A1}(t_1)$  et  $Z_{A2}(t_1)$ . En effet, la projection de A2 dans l'image I2 n'est pas la même que celle de A1 dans l'image I2. Par ailleurs, lesdites deux valeurs de profondeur  $Z_{A1}(t_1)$  et  $Z_{A2}(t_1)$  sont sauvegardées. Ici, on se retrouve donc avec deux possibilités de profondeur, 15  $Z_{A1}(t_1)$  et  $Z_{A2}(t_1)$ , ce qui signifie qu'on a plusieurs solutions concernant la profondeur du point de l'espace A. A ce stade, il n'est pas encore possible de faire l'association entre le point de l'espace A et sa projection dans chaque image, c'est-à-dire qu'on ne connaît pas quelle projection correspond au point de l'espace A ou au point de l'espace B.

20 Lors d'une deuxième phase de calcul, appelé « Calcul 2 » (étape E16), le module de traitement d'images 200 calcule la profondeur du point de l'espace A dans le repère caméra au premier instant  $t_1$  à partir de la troisième image I3 (non représentée).

Le but de cette étape est de résoudre l'ambiguïté présente à l'étape précédente, c'est-à-dire pouvoir savoir quelle projection dans chaque image de la séquence d'images I1, I2, I3 correspond au point de l'espace A et quelle projection dans 25 chaque image de la séquence d'images I1, I2, I3 correspond au point de l'espace B.

On réalise les mêmes calculs que précédemment mais entre le premier instant  $t_1$  et le troisième instant  $t_3$ . De la même façon, on se retrouve donc avec deux possibilités de profondeur,  $Z_{A3}(t_1)$  et  $Z_{A4}(t_1)$ , ce qui signifie qu'on a plusieurs distances possibles entre la caméra 100 et l'objet réel.

30 Puisque le premier instant  $t_1$  et le troisième instant  $t_3$  ne sont pas des instants directement consécutifs, il faut réaliser une étape intermédiaire. En effet, on connaît la matrice de rotation R et la matrice de translation T entre deux instants consécutifs, c'est-à-dire entre le premier instant  $t_1$  et le deuxième instant  $t_2$  et entre le deuxième instant  $t_2$  et le troisième instant  $t_3$ . Ainsi, afin de connaître A( $t_3$ ), il faut cumuler les opérations de

rotation et de translation successives, avec  $R_{2/3}$  et  $T_{2/3}$  étant respectivement la matrice de rotation entre le deuxième instant  $t_2$  et le troisième instant  $t_3$ , la matrice de translation entre le deuxième instant  $t_2$  et le troisième instant  $t_3$ .

On a tout d'abord :

5 
$$A(t_2) = R \times A(t_1) + T \quad [7]$$

$$A(t_3) = R_{2/3} \times A(t_2) + T_{2/3} \quad [8]$$

Ainsi, si on remplace [7] dans [8], on obtient :

$$A(t_3) = R_{2/3} \times (R \times A(t_1) + T) + T_{2/3}$$

$$A(t_3) = R_{2/3} \times R \times A(t_1) + R_{2/3}T + T_{2/3}$$

10 Par ailleurs, lesdites deux valeurs de profondeur  $Z_{A3}(t_1)$  et  $Z_{A4}(t_1)$  sont mémorisées dans une étape E17.

Le module de traitement d'images 200 réalise la comparaison, dans en une étape E18, de tous les résultats de valeurs potentielles de profondeur obtenues lors du « calcul 1 » et du « calcul 2 », deux valeurs de profondeurs parmi  $Z_{A1}(t_1)$ ,  $Z_{A2}(t_1)$ , 15  $Z_{A3}(t_1)$  et  $Z_{A4}(t_1)$  étant identiques. Ces deux profondeurs de même valeur correspondent à un même et unique point réel, correspondant au point réel de l'espace A. Ainsi, le module de traitement d'images 200 identifie le point réel recherché (étape E19).

20 Dans l'exemple illustré en référence à la **figure 6**,  $Z_{A1}(t_1)$  et  $Z_{A4}(t_1)$  sont des valeurs de profondeurs identiques et correspondent donc à la profondeur du point réel de l'espace A.  $Z_{A2}(t_1)$  et  $Z_{A3}(t_1)$  correspondent respectivement à A2 et A3, des points fictifs de l'espace dont la valeur de profondeur est différente de la valeur de profondeur du point réel de l'espace A. Autrement dit : les points fictifs de l'espace A2 et A3 ne correspondent à aucun point réel (points de l'espace A ou B).

25 L'invention permet donc avantageusement de distinguer des motifs identiques se répétant dans les images successives d'une séquence d'images I1, I2, I3 afin d'assister efficacement le conducteur dans sa conduite.

## REVENDICATIONS

1. Procédé de traitement d'images permettant l'identification d'au moins un même point apparaissant sur chaque image d'une séquence d'images (I1, I2), ladite séquence d'images (I1, I2, I3) comprenant une première image (I1), une deuxième image (I2) et une troisième image (I3) générées respectivement à un premier instant (t1), à un deuxième instant (t2), postérieur au premier instant (t1), et à un troisième instant (t3), postérieur au deuxième instant (t2), par une caméra (100) montée dans un véhicule (1) automobile, ledit procédé étant **caractérisé en ce qu'il comprend** les étapes de :
- détection (E1-E13) d'une pluralité de points identiques dans chaque image de la séquence d'images (I1, I2, I3) relative à une pluralité de motifs répétitifs dans chaque image de la séquence d'images (I1, I2, I3),
  - pour chaque point identique détecté dans la deuxième image (I2) de la séquence d'images (I1, I2, I3), calcul (E14) de la profondeur ( $Z_{A1}(t1)$ ,  $Z_{A2}(t1)$ ) au premier instant (t1) du point de l'espace correspondant au dit point identique de la deuxième image (I2),
  - pour chaque point identique détecté dans la troisième image de la séquence d'images (I1, I2, I3), calcul (E16) de la profondeur ( $Z_{A3}(t1)$ ,  $Z_{A4}(t1)$ ) au premier instant (t1) du point de l'espace correspondant au dit point identique de la troisième image (I3),
  - identification (E19), dans la deuxième image (I2) et dans la troisième image (I3), des points identiques correspondants aux points de l'espace pour lesquels la profondeur calculée est identique, ces points de l'espace représentant respectivement dans la deuxième image (I2) et dans la troisième image (I3) un même et unique point réel.
- 25 2. Procédé selon la revendication 1, dans lequel la profondeur ( $Z_{A1}(t1)$ ,  $Z_{A2}(t1)$ ) au premier instant (t1) du point de l'espace correspondant au dit point identique de la deuxième image (I2) est calculée à partir de la matrice de rotation (R) et du vecteur de translation (T) entre la première image (I1) et la deuxième image (I2).
- 30 3. Procédé selon l'une des revendications précédentes, dans lequel la profondeur ( $Z_{A3}(t1)$ ,  $Z_{A4}(t1)$ ) au premier instant (t1) du point de l'espace correspondant au dit point identique de la troisième image (I3) est calculée à partir de la matrice de rotation et du vecteur de translation entre la première image (I1) et la troisième image (I2).

4. Procédé selon l'une des revendications précédentes, comprenant une étape (E15) de mémorisation, après l'étape de calcul, de la profondeur ( $Z_{A1}(t1)$ ,  $Z_{A2}(t1)$ ) pour chaque point identique détecté dans la deuxième image (I2) de la séquence d'images (I1, I2, I3).
5. Procédé selon l'une des revendications précédentes, comprenant une étape (E17) de mémorisation, après l'étape de calcul, de la profondeur ( $Z_{A3}(t1)$ ,  $Z_{A4}(t1)$ ) pour chaque point identique détecté dans la troisième image (I3) de la séquence d'images (I1, I2, I3).
6. Module de traitement d'images (200) permettant l'identification d'au moins un même point apparaissant sur chaque image d'une séquence d'images (I1, I2, I3), ladite séquence d'images (I1, I2, I3) comprenant une première image (I1), une deuxième image (I2) et une troisième image (I3) générées respectivement à un premier instant (t1), à un deuxième instant (t2), postérieur au premier instant (t1), et à un troisième instant (t3), postérieur au deuxième instant (t2), par une caméra (100) montée dans un véhicule (1) automobile, ledit module de traitement d'images (200) étant **caractérisé en ce qu'il est** configuré pour:
- détecter une pluralité de points identiques dans chaque image de la séquence d'images (I1, I2, I3) relative à une pluralité de motifs répétitifs dans chaque image de la séquence d'images (I1, I2, I3),
  - calculer de la profondeur ( $Z_{A1}(t1)$ ,  $Z_{A2}(t1)$ ) au premier instant (t1) du point de l'espace correspondant au dit point identique de la deuxième image (I2) pour chaque point identique détecté dans la deuxième image (I2) de la séquence d'images (I1, I2, I3),
  - calculer la profondeur ( $Z_{A3}(t1)$ ,  $Z_{A4}(t1)$ ) au premier instant (t1) du point de l'espace correspondant au dit point identique de la troisième image (I3), pour chaque point identique détecté dans la troisième image (I3) de la séquence d'images (I1, I2, I3),
  - identifier dans la deuxième image (I2) et dans la troisième image (I3), des points identiques correspondants aux points de l'espace pour lesquels la distance calculée est identique, ces points de l'espace représentant respectivement dans la deuxième image (I2) et dans la troisième image (I3) un même et unique point réel.
7. Module de traitement d'images (200) selon la revendication précédente, ledit module de traitement d'images (200) étant configuré pour calculer la profondeur ( $Z_{A1}(t1)$ ,  $Z_{A2}(t1)$ ) au premier instant (t1) du point de l'espace correspondant au dit point identique

de la deuxième image (I2) à partir de la matrice de rotation et du vecteur de translation entre la première image (I1) et la deuxième image (I2).

**8.** Module de traitement d'images (200) selon l'une des revendications 6 et 7, ledit module de traitement d'images (200) étant configuré pour calculer la 5 profondeur ( $Z_{A3}(t1)$ ,  $Z_{A4}(t1)$ ) au premier instant (t1) du point de l'espace correspondant au dit point identique de la troisième image (I3) à partir de la matrice de rotation (R) et du vecteur de translation (T) entre la première image (I1) et la troisième image (I2).

**9.** Module de traitement d'images (200) selon l'une des revendications 6 à 8, ledit module de traitement d'images (200) étant configuré pour mémoriser la 10 profondeur ( $Z_{A1}(t1)$ ,  $Z_{A2}(t1)$ ) pour chaque point identique détecté dans la deuxième image (I2) de la séquence d'images (I1, I2, I3) et la profondeur ( $Z_{A3}(t1)$ ,  $Z_{A4}(t1)$ ) pour chaque point identique détecté dans la troisième image (I3) de la séquence d'images (I1, I2, I3).

**10.** Véhicule automobile comprenant un module de traitement d'images (200) 15 selon l'une des revendications 6 à 9.

Fig 1

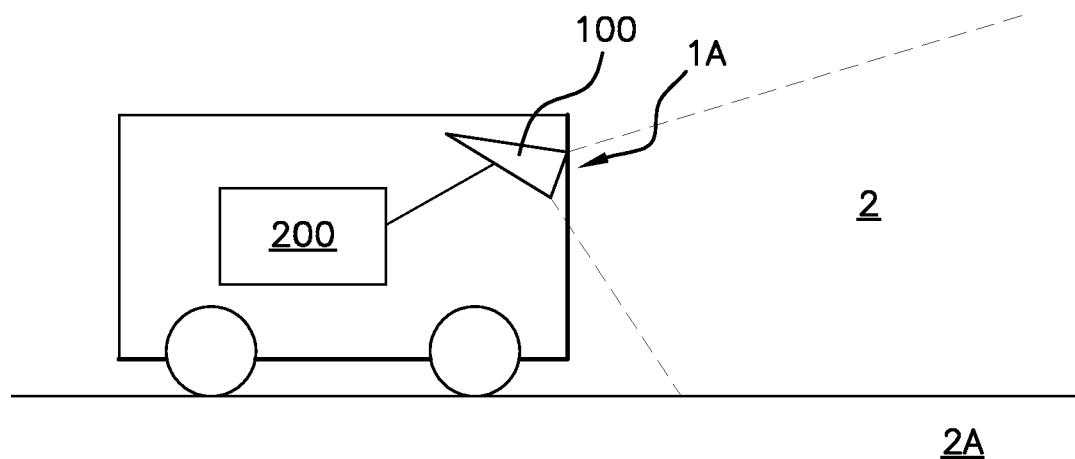


Fig 2

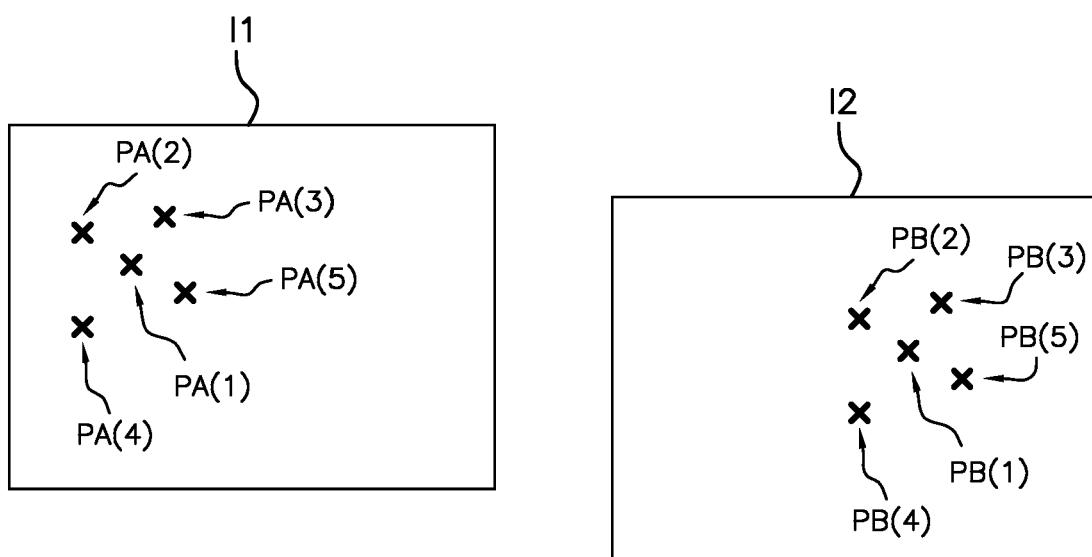


Fig 3

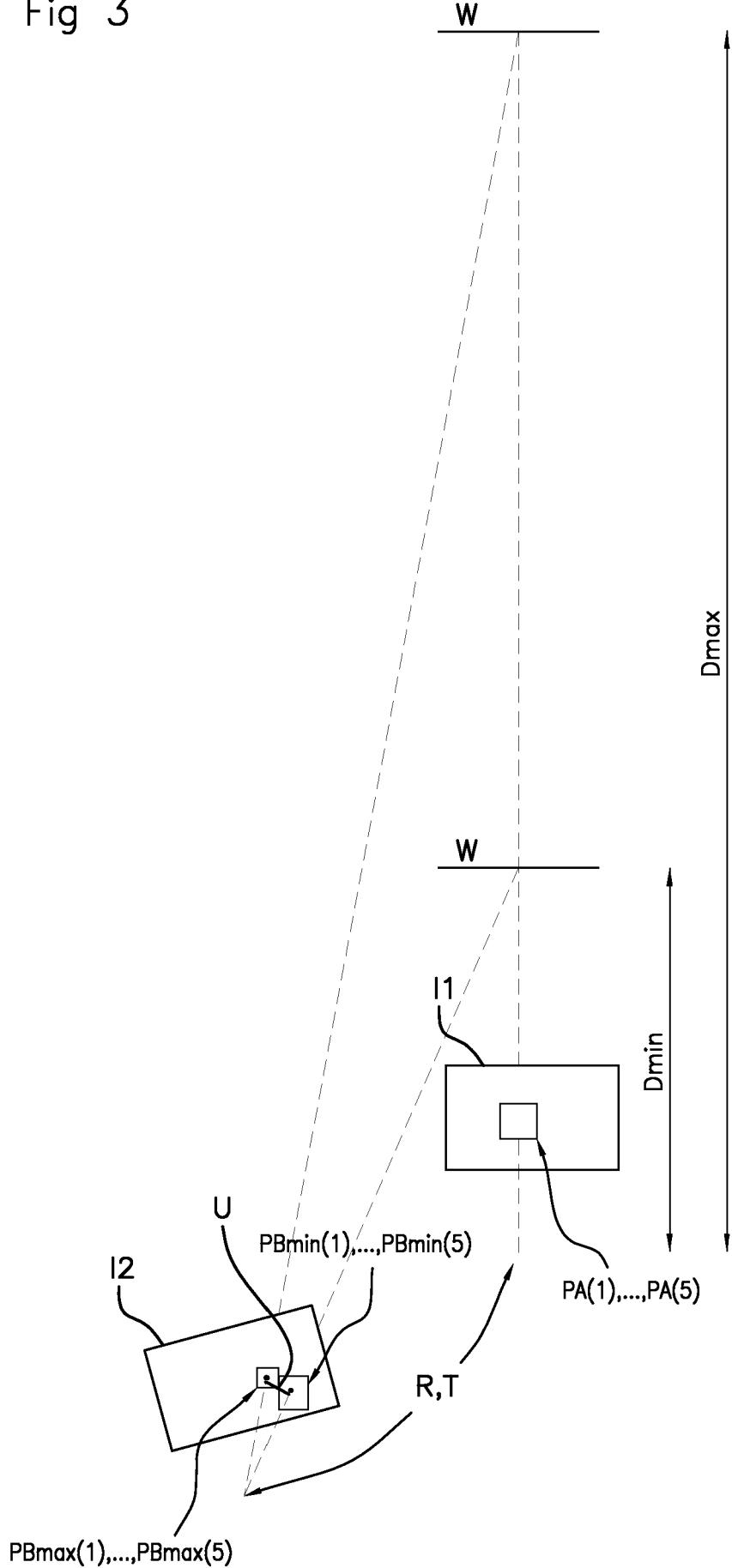
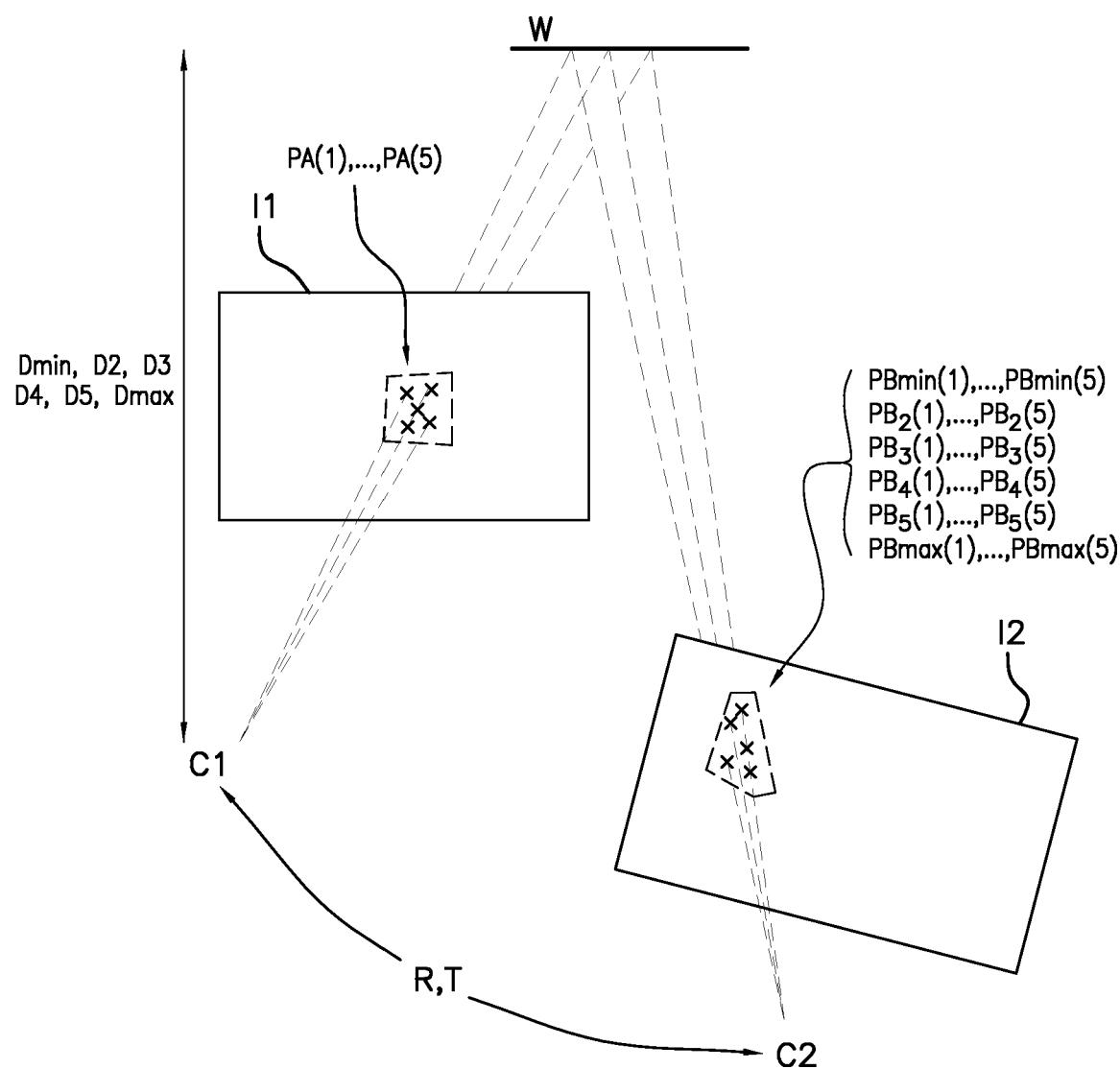


Fig 4



4/7

Fig 5

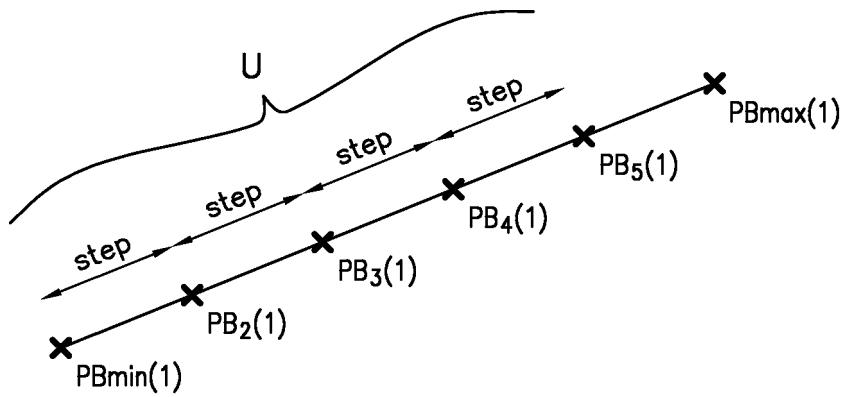


Fig 6

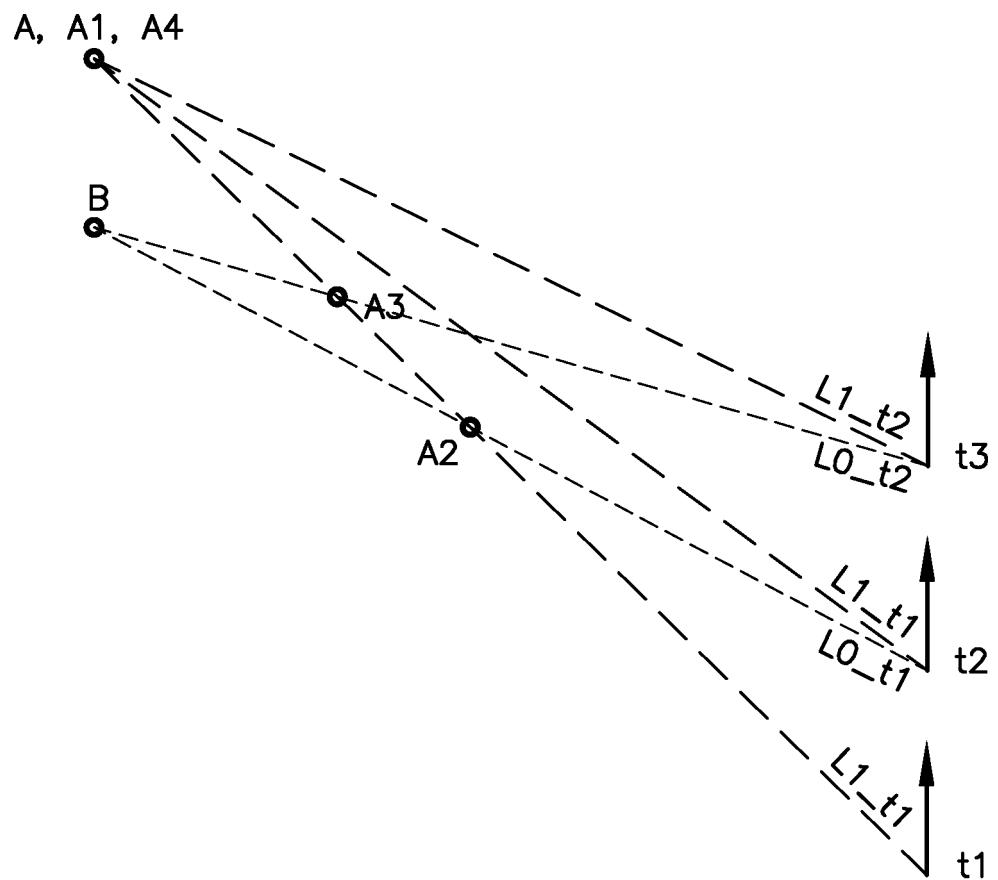


Fig 7

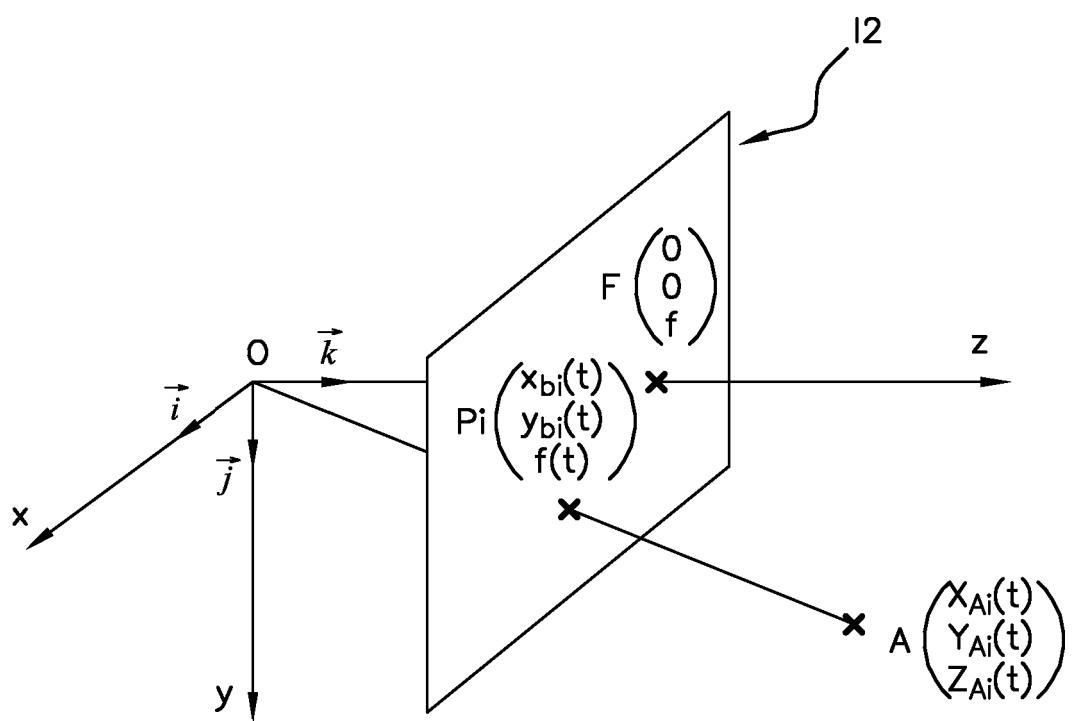


Fig 8

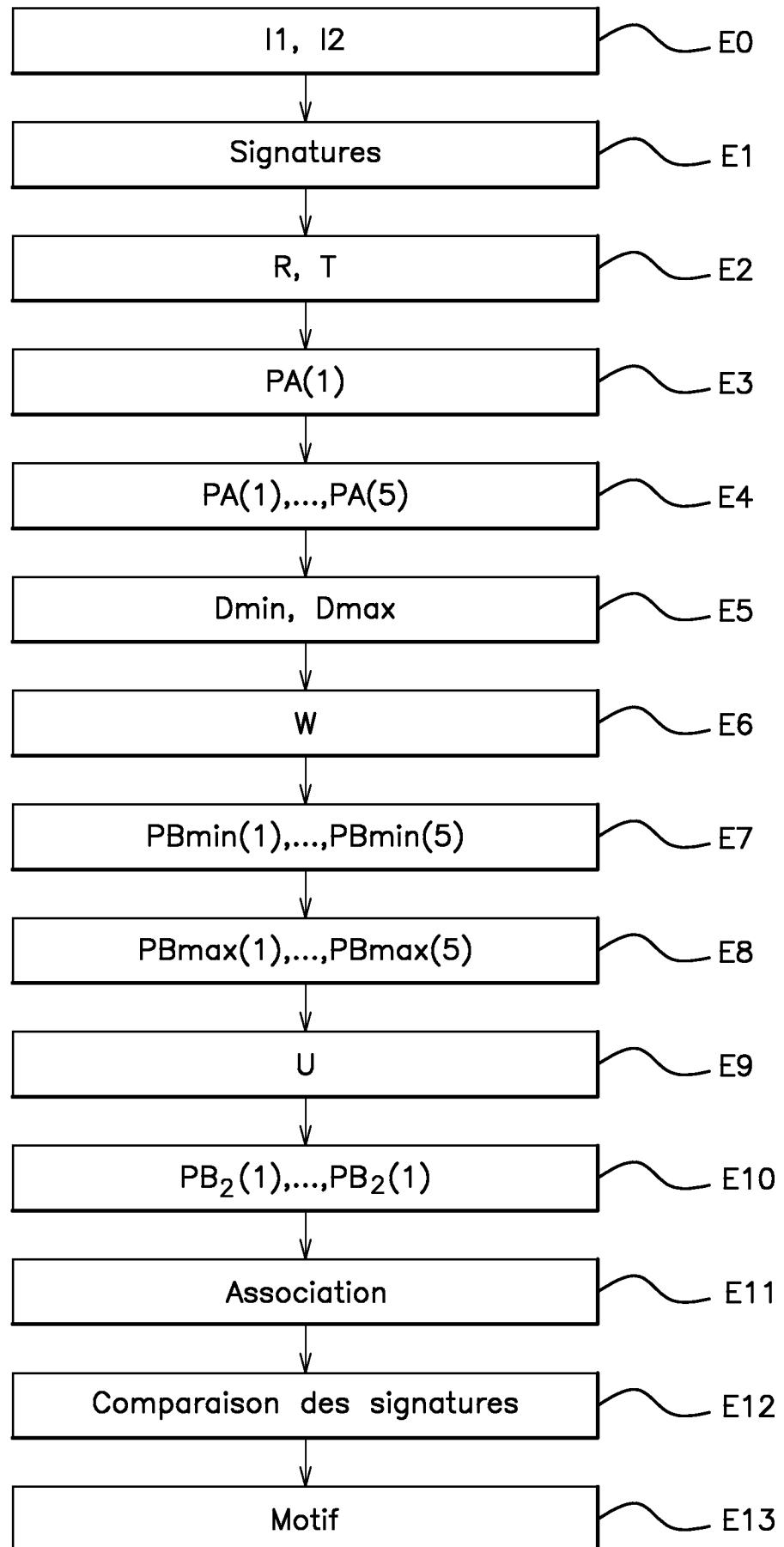
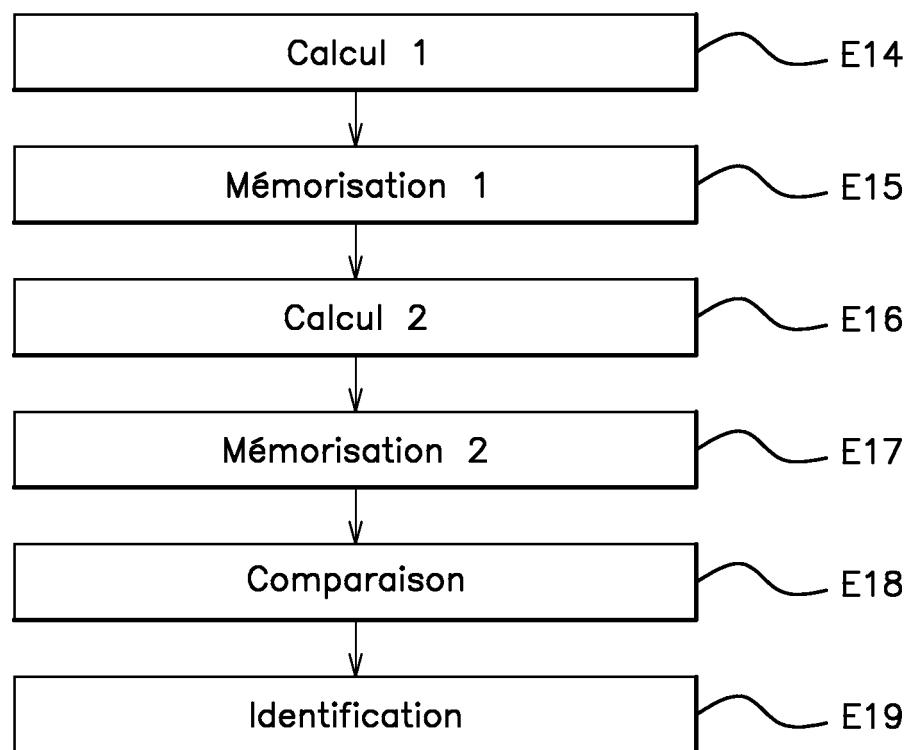


Fig 9



# RAPPORT DE RECHERCHE

articles L.612-14, L.612-53 à 69 du code de la propriété intellectuelle

## OBJET DU RAPPORT DE RECHERCHE

L'I.N.P.I. annexe à chaque brevet un "RAPPORT DE RECHERCHE" citant les éléments de l'état de la technique qui peuvent être pris en considération pour apprécier la brevetabilité de l'invention, au sens des articles L. 611-11 (nouveauté) et L. 611-14 (activité inventive) du code de la propriété intellectuelle. Ce rapport porte sur les revendications du brevet qui définissent l'objet de l'invention et délimitent l'étendue de la protection.

Après délivrance, l'I.N.P.I. peut, à la requête de toute personne intéressée, formuler un "AVIS DOCUMENTAIRE" sur la base des documents cités dans ce rapport de recherche et de tout autre document que le requérant souhaite voir prendre en considération.

## CONDITIONS D'ETABLISSEMENT DU PRESENT RAPPORT DE RECHERCHE

Le demandeur a présenté des observations en réponse au rapport de recherche préliminaire.

Le demandeur a maintenu les revendications.

Le demandeur a modifié les revendications.

Le demandeur a modifié la description pour en éliminer les éléments qui n'étaient plus en concordance avec les nouvelles revendications.

Les tiers ont présenté des observations après publication du rapport de recherche préliminaire.

Un rapport de recherche préliminaire complémentaire a été établi.

## DOCUMENTS CITES DANS LE PRESENT RAPPORT DE RECHERCHE

La répartition des documents entre les rubriques 1, 2 et 3 tient compte, le cas échéant, des revendications déposées en dernier lieu et/ou des observations présentées.

Les documents énumérés à la rubrique 1 ci-après sont susceptibles d'être pris en considération pour apprécier la brevetabilité de l'invention.

Les documents énumérés à la rubrique 2 ci-après illustrent l'arrière-plan technologique général.

Les documents énumérés à la rubrique 3 ci-après ont été cités en cours de procédure, mais leur pertinence dépend de la validité des priorités revendiquées.

Aucun document n'a été cité en cours de procédure.

**1. ELEMENTS DE L'ETAT DE LA TECHNIQUE SUSCEPTIBLES D'ETRE PRIS EN  
CONSIDERATION POUR APPRECIER LA BREVETABILITE DE L'INVENTION**

EP 2 431 917 A1 (MOBILEYE TECHNOLOGIES LTD  
[CY]) 21 mars 2012 (2012-03-21)

Frédéric Sur ET AL: "Image point  
correspondences and repeated patterns",  
HAL Id: inria-00609998,  
20 juillet 2011 (2011-07-20), XP055556303,  
Extrait de l'Internet:  
URL:<https://hal.inria.fr/inria-00609998/document>  
[extrait le 2019-02-13]

MARIA KUSHNIR ET AL: "Epipolar Geometry  
Estimation for Urban Scenes with  
Repetitive Structures",  
IEEE TRANSACTIONS ON PATTERN ANALYSIS AND  
MACHINE INTELLIGENCE,  
vol. 36, no. 12,  
16 juillet 2014 (2014-07-16), pages  
2381-2395, XP055556297,  
USA  
ISSN: 0162-8828, DOI:  
10.1109/TPAMI.2014.2339862

**2. ELEMENTS DE L'ETAT DE LA TECHNIQUE ILLUSTRANT L'ARRIERE-PLAN  
TECHNOLOGIQUE GENERAL**

NEANT

**3. ELEMENTS DE L'ETAT DE LA TECHNIQUE DONT LA PERTINENCE DEPEND  
DE LA VALIDITE DES PRIORITES**

NEANT