



(51) Classification internationale des brevets :
G06K 9/00 (2006.01)

(21) Numéro de la demande internationale :
PCT/EP2016/066381

(22) Date de dépôt international :
11 juillet 2016 (11.07.2016)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
1556520 9 juillet 2015 (09.07.2015) FR

(71) Déposant : VALEO SCHALTER UND SENSOREN
GMBH [DE/DE]; Laiernstr. 12, 74321 Bietigheim-Bissingen (DE).

(72) Inventeurs : REBUT, Julien; C/o Valeo Comfort and Driving Assistance, 76 rue Auguste Perret - ZI Europarc, 94046 Créteil Cedex (FR). CSILLAG, Dora; C/o Valeo Comfort and Driving Assistance, 76 rue Auguste Perret - ZI Europarc, 94046 Créteil Cedex (FR).

(74) Mandataire : DELPLANQUE, Arnaud; C/o Valeo Comfort and Driving Assistance, Propriété Intellectuelle, 76, rue Auguste Perret - ZI Europarc, 94046 Créteil Cedex (FR).

(81) États désignés (sauf indication contraire, pour tout titre de protection nationale disponible) : AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) États désignés (sauf indication contraire, pour tout titre de protection régionale disponible) : ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), eurasien (AM, AZ, BY, KG, KZ, RU, TJ, TM), européen (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Publiée :

— avec rapport de recherche internationale (Art. 21(3))

(54) Title : DETECTION OF OBJECTS BY IMAGE PROCESSING

(54) Titre : DÉTECTION D'OBJETS PAR TRAITEMENT D'IMAGES

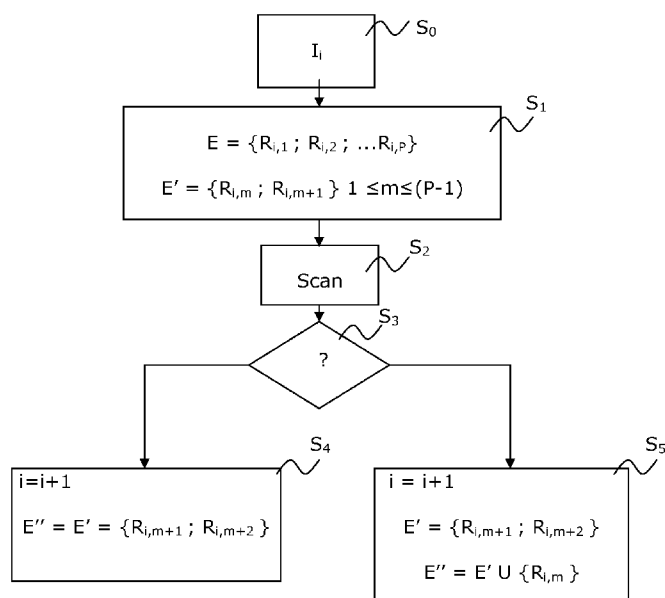


FIG. 3

(57) Abstract : The invention relates to a method for detecting objects by processing images captured by a camera, in which method an algorithm for shape recognition by learning is used, which is capable of detecting at least one object by performing, for each image of a series of images, a scan of a plurality of regions of interest by means of a sliding window having a predetermined fixed size, each region of interest corresponding to a different scale of each image. The invention involves selecting (S₁), from a predetermined set (E) of P regions of interest which are classified in a predefined order of consecutive scales, a first subset (E') of K consecutive regions of interest on which the sliding window performs the scan (S₂) for at least one first image.

(57) Abrégé : L'invention concerne un procédé de détection d'objets par traitement d'images captées par une caméra, dans lequel on utilise un algorithme de reconnaissance de formes par apprentissage apte à détecter au moins un objet en effectuant, pour chaque image d'une succession d'images, le balayage d'une pluralité de régions d'intérêt par une fenêtre glissante de taille fixe prédéterminée, chaque région d'intérêt correspondant à une échelle différente de chaque image. Conformément à l'invention, on sélectionne (S₁) dans un ensemble (E) prédéterminé de P régions d'intérêt, classées dans un ordre défini d'échelles successives, un premier sous-ensemble (E') de K régions d'intérêt successives

sur lesquelles le balayage par la fenêtre glissante est effectué (S₂) pour au moins une première image.

DETECTION D'OBJETS PAR TRAITEMENT D'IMAGES

5 La présente invention concerne de manière générale la détection d'objets par les techniques de reconnaissance de formes dans des images.

Un domaine plus particulièrement, bien que non exclusivement, concerné est celui des véhicules automobiles.

10 Il est connu d'équiper certains véhicules automobiles avec des systèmes d'aide à la conduite utilisant des images captées par une ou plusieurs caméras embarquées sur le véhicule pour détecter notamment les lignes de marquage routier, les obstacles, les autres véhicules, ou encore les panneaux de signalisation tels que les panneaux de prescription de vitesse.

15 La détection de piétons est une autre application importante dans les systèmes d'aide à la conduite automobile puisqu'il s'agit de concevoir des systèmes intelligents capables d'avertir ou de prévenir les accidents, en surveillant les alentours du véhicule par l'intermédiaire d'une ou plusieurs caméras. Un tel système de détection peut ainsi soit prévenir le conducteur de la présence d'un piéton à une distance jugée dangereuse, soit directement agir
20 sur le système de freinage du véhicule automobile. Les systèmes de détection de piétons sont en outre cruciaux pour les systèmes de conduite automatisés.

25 Que l'objet à détecter soit fixe ou mobile, on comprend ainsi qu'un tel système de détection doit pouvoir fournir une information suffisamment fiable et précise pour pouvoir dériver, en temps réel, une estimation de la distance à laquelle se situe l'objet détecté par rapport au véhicule.

On rappelle que la détection de piétons consiste à déterminer le plus précisément possible, la présence et la localisation de toutes les personnes susceptibles d'être présentes dans une image ou une succession d'images. Elle est généralement basée sur des techniques de reconnaissance de formes dont
30 le but est d'apprendre puis de retrouver dans l'image, l'allure générale d'une personne.

Les algorithmes mis en œuvre dans une détection de piétons par reconnaissance de formes sont généralement très complexes compte tenu

notamment de la très grande diversité entre les personnes (taille, poids, vêtements, postures...). Ces algorithmes de reconnaissance de formes sont souvent basés sur l'utilisation combinée de descripteurs et d'une méthode de classification binaire permettant de déterminer si une zone d'une image captée

5 par une caméra correspond à un piéton ou à du fond. Une méthode classique, connue sous le nom de la méthode de Viola et Jones, consiste à tester la présence d'un piéton dans une fenêtre ou imagerie de taille fixe, par exemple de 24 sur 48 pixels, à toutes les positions possibles dans l'image et pour plusieurs échelles de l'image. Le test consiste à générer un vecteur de

10 descripteurs de l'imagerie et à les comparer avec les descripteurs appris sur une base d'apprentissage. Plus précisément, pour chaque image captée par une caméra, on définit une pluralité de régions d'intérêt à balayer par la fenêtre, chaque région d'intérêt correspondant à une échelle donnée de l'image. A titre d'exemple, une vingtaine de régions d'intérêt, encore appelées

15 zones de balayage relatives à chaque échelle (« scan zone » en terminologie anglo-saxonne), doivent être traitées correspondant à des échelles de l'image variant de 0,025 (zoom avant) à 1 (taille réelle) par pas de 1,22.

La méthode consiste typiquement à balayer chaque région d'intérêt, de même taille que l'image de base, en déplaçant la fenêtre de taille fixe d'un

20 certain nombre de pixels à la fois. Un classifieur combiné à une base d'apprentissage comportant des échantillons de piétons et des échantillons de fond va traiter les différents vecteurs de descripteurs obtenus par la fenêtre glissante pour identifier les zones correspondant effectivement à un piéton par rapport à un modèle appris. Le résultat est délivré par le détecteur de piéton

25 sous forme d'une boîte englobante, généralement de forme rectangulaire, par piéton détecté. Les dimensions et/ou la position verticale de la boîte englobante relative à un piéton détecté sont alors utilisées pour estimer la distance à laquelle se situe le piéton par rapport au véhicule. Une telle méthode est par exemple implémentée par des détecteurs rapides de type

30 Adaboost. Les résultats de la détection rapide sont ensuite généralement affinés et validés par des classifieur plus performants.

On conçoit aisément qu'une telle méthode est très consommatrice en calcul et en temps, et de ce fait, incompatible avec la nécessité d'un traitement en temps réel primordial pour les systèmes d'aide à la conduite.

5 La présente invention a pour but de pallier les inconvénients des solutions connues basées sur l'algorithme de reconnaissance de formes de type Adaboost en proposant une stratégie de traitement des images qui soit moins consommatrice en temps de calcul.

10 Pour ce faire, l'invention a pour objet un procédé de détection d'objets par traitement d'images captées par une caméra, dans lequel on utilise un algorithme de détection et/ou de reconnaissance de formes apte à détecter au moins un objet en effectuant, pour chaque image d'une succession d'images, le balayage d'une pluralité de régions d'intérêt par une fenêtre glissante, chaque région d'intérêt correspondant à une échelle différente de chaque
15 image, le procédé étant caractérisé en ce qu'il comprend une étape de sélection, dans un ensemble prédéterminé de P régions d'intérêt classées dans un ordre défini d'échelles successives, d'un premier sous-ensemble de K régions d'intérêt successives sur lesquelles le balayage par la fenêtre glissante est effectué pour au moins une première image.

20 Outre les caractéristiques principales qui viennent d'être mentionnées dans le paragraphe précédent, le procédé selon l'invention peut présenter une ou plusieurs caractéristiques complémentaires parmi les suivantes :

25 - on mémorise chaque région d'intérêt du premier sous-ensemble ayant donné lieu à une détection d'un objet ; et on réitère le balayage pour les images suivantes sur un deuxième sous-ensemble de régions d'intérêt successives limité d'une part, aux régions d'intérêt mémorisées et aux régions contiguës aux régions d'intérêt mémorisées, et d'autre part, aux K régions d'intérêt du premier sous-ensemble, en déplaçant, à chaque itération, ledit premier sous-ensemble de K régions d'intérêt successives d'au moins une
30 région d'intérêt dans l'ordre défini d'échelles successives ;

 - ledit premier sous-ensemble de K régions d'intérêt successives est déplacé par exemple au plus de K régions d'intérêt ;

- le balayage peut être réitéré jusqu'à ce que le premier sous-ensemble de K régions d'intérêt successives inclue la région d'intérêt dudit ensemble prédéterminé correspondant à la plus grande échelle ;
- la plus grande échelle correspond par exemple à la taille réelle de
5 l'image captée par la caméra ;
- ledit premier sous-ensemble de K régions d'intérêt successives inclut dans un mode de réalisation possible la région d'intérêt dudit ensemble prédéterminé correspondant à la plus petite échelle ;
- ledit premier sous-ensemble comprend de préférence deux régions
10 d'intérêts successives ;
- chaque région d'intérêt peut être en outre avantageusement segmentée dans le sens de la largeur en une pluralité de zones d'intérêt, et le procédé comporte alors une étape de recadrage des régions d'intérêt mémorisées, voire des régions contigües aux régions d'intérêt, le recadrage
15 s'effectuant dans la zone d'intérêt ayant donné lieu à une détection d'objet.

L'invention a également pour objet un système de détection d'objets par traitement d'images captées par une caméra, le système comprenant un détecteur utilisant un algorithme de détection et/ou de reconnaissance de
20 formes apte à détecter au moins un objet en effectuant, pour chaque image d'une succession d'images, le balayage d'une pluralité de régions d'intérêt par une fenêtre glissante, chaque région d'intérêt correspondant à une échelle différente de chaque image, caractérisé en ce que ledit détecteur est programmé pour sélectionner dans un ensemble prédéterminé de P régions
25 d'intérêt, classées dans un ordre défini d'échelles successives, un premier sous-ensemble de K régions d'intérêt successives sur lesquelles le balayage par la fenêtre glissante est effectué pour une première image.

L'invention sera mieux comprise à la lecture de la description qui suit,
30 faite en référence aux figures annexées, dans lesquelles :

- la figure 1 représente schématiquement une situation rencontrée sur une route par un véhicule automobile équipé d'un système selon l'invention ;
- la figure 2 donne un synoptique simplifié illustrant le fonctionnement

d'un détecteur de piéton de type Adaboost ;

- la figure 3 donne un synoptique simplifié illustrant certaines étapes susceptibles d'être mise en œuvre dans un procédé conforme à l'invention ;
- la figure 4 illustre un principe de segmentation des régions d'intérêt.

5

Dans la suite et en référence à la figure 1, on suppose à titre d'exemple qu'un véhicule automobile 1 équipé d'une caméra , de paramètres de calibration connus et apte à capter des images successives de paysages, se
10 déplace sur une route 3. La caméra est, dans cet exemple, localisée à un endroit du véhicule qui correspond le mieux à ce que le conducteur voit, par exemple centrée au niveau du pare-brise à l'intérieur de l'habitacle. D'autres positionnements peuvent néanmoins être envisagés sans départir du cadre de l'invention, par exemple sur le pare-choc arrière pour une caméra de recul, ou
15 plus généralement à n'importe quel endroit du véhicule selon la zone que l'on souhaite observer. Le véhicule automobile 1 comporte en outre un module de traitement (non représenté) formant, avec la caméra 2, un système permettant de détecter la présence d'un piéton 4, et d'agir sur un système d'aide à la conduite en fonction d'une estimation de la distance séparant le
20 piéton 4 du véhicule automobile 1. La détection de piétons est prise ici à titre d'exemple non limitatif, le système pouvant être prévu pour la détection d'autres objets. Ainsi, la caméra 2 capture les images de la scène routière située à l'avant du véhicule 1 et fournit ces images au module de traitement d'images du système.

25 Comme représenté très schématiquement sur la figure 2, chaque image I_i captée par la caméra est délivrée à un détecteur 5, dans cet exemple non limitatif, un détecteur de piétons, comportant une base d'apprentissage 50 avec des exemples de piétons d'une part, et de fonds d'autre part, et son classifieur associé 51. Le classifieur 51 est apte à implémenter l'algorithme de
30 type Viola et Jones décrit ci-avant, avec les modifications liées à l'invention. On rappelle que cet algorithme consiste en général à balayer chaque région d'intérêt $R_{i,p}$ d'un ensemble E de P régions d'intérêt correspondant à P échelles

possibles de l'image d'origine I_i , par une imagerie de taille prédéterminée. Lorsqu'un piéton est détecté par ce traitement, le détecteur 5 de piéton délivre en sortie une boîte englobante 52 qui délimite une zone d'image représentative du piéton 4 détecté. Le traitement ci-dessus est réitéré pour
5 chaque image reçue.

Afin de limiter le temps de traitement précédent, le procédé conforme à l'invention propose de ne balayer, pour chaque image reçue I_i , qu'un sous-ensemble de régions d'intérêt comprenant au maximum un nombre limité de K régions d'intérêt successives (entier K inférieur à P), ainsi que des régions
10 d'intérêt ayant donné lieu à une détection pour l'image précédente I_{i-1} (appelées dans la suite régions d'intérêt validées ou mémorisées) ou contiguës à ces régions d'intérêt validées. A chaque itération, c'est-à-dire lors du traitement d'une nouvelle image captée, on change de préférence en outre les K régions d'intérêt successives qui vont être balayées.

15 Un exemple d'étapes susceptibles d'être effectuées conformément au procédé selon l'invention va être décrit en référence à la figure 3.

Sur cette figure 3, I_i est représentative d'une image captée à l'instant i par la caméra (étape S_0). L'ensemble E de régions d'intérêts qui devraient être balayées si l'on devait appliquer la méthode classique de Viola et Jones
20 par un détecteur à reconnaissance de formes de type Adaboost comprend normalement P régions d'intérêts notées par commodité $R_{i,p}$, correspondant à P échelles différentes de l'image I_i , classées par ordre croissant d'échelles successives. Ainsi, pour chaque image I_i :

$$E = \{ R_{i,1} ; R_{i,2} ; \dots R_{i,P} \}$$

25 avec $R_{i,P}$ la région d'intérêt correspondant à l'image I_i de la taille réelle (en d'autres termes, $R_{i,P} = I_i$), et $R_{i,1}$ la région d'intérêt correspondant à cette même image à la plus petite échelle, par exemple une échelle de 0,025 correspondant à un sous échantillonnage, c'est-à-dire une diminution de la résolution de l'image I_i .

30 On supposera dans la suite, pour simplifier, que la valeur P est égale à 8. En pratique, la valeur de P dépend de nombreux critères, tels que l'orientation de la caméra, la portée de détection que l'on cherche.

Ici, conformément à l'invention, on prévoit de sélectionner dans un premier temps (étape S_1) un premier sous-ensemble E' comportant seulement K régions d'intérêt successives parmi les P possibles. A titre d'exemple non limitatif, l'entier K est fixé égal à 2. Le premier sous-ensemble E' peut donc
 5 s'écrire sous la forme mathématique suivante :

$$E' = \{ R_{i,m} ; R_{i,m+1} \}$$

avec m un entier quelconque compris entre 1 et $P-1$.

On effectue alors, sur chacune des régions d'intérêt successives du premier sous-ensemble E' , ici au nombre de 2, un balayage classique par la
 10 fenêtre glissante, de manière à détecter la présence éventuelle d'un ou plusieurs piétons dans ces régions d'intérêt (étape S_2). Chaque région d'intérêt ayant effectivement donné lieu à une détection de piéton est alors mémorisée.

Pour l'image suivante ($i = i+1$), on réitère le balayage sur un
 15 deuxième sous-ensemble E'' dont le nombre de régions d'intérêt dépend des détections réalisées à l'étape précédente.

Plus précisément, dans le cas où le balayage précédent n'a permis aucune détection (branche de gauche du test référencé S_3 sur la figure 3), le deuxième sous-ensemble E'' correspond au contenu de premier sous-ensemble
 20 E' réactualisé en déplaçant ledit premier sous-ensemble d'au moins une région d'intérêt dans l'ordre croissant d'échelles successives. Dans l'exemple représenté sur la figure 3, ceci correspond à l'étape S_4 , dans laquelle le déplacement correspond à exactement une région d'intérêt. Il convient de noter, de façon plus générale, que le déplacement ne peut être supérieur à
 25 l'entier K . Ainsi, le deuxième sous-ensemble E'' peut s'écrire dans ce cas sous la forme mathématique suivante :

$$E'' = E' = \{ R_{i,m+1} ; R_{i,m+2} \}$$

30 ce qui signifie que, pour l'image suivante, le balayage par l'imagette ne s'effectuera là encore que sur deux régions d'intérêt successives.

Dans le cas contraire (branche de droite du test S_3 sur la figure 3), le deuxième sous-ensemble E'' de régions d'intérêt successives comprend d'une

part, les régions d'intérêt mémorisées, voire celles contiguës aux régions d'intérêt mémorisées, et d'autre part, le contenu du premier ensemble E' réactualisé comme expliqué précédemment, par déplacement d'au moins une région d'intérêt. Dans l'exemple représenté sur la figure 3, ceci correspond à l'étape S_5 , à l'issue de laquelle le deuxième sous-ensemble E'' peut s'écrire sous la forme mathématique suivante, en supposant que l'étape S_2 a permis la détection d'un objet dans la région d'intérêt $R_{i,m}$:

$$E'' = R_{i,m} \cup E' = R_{i,m} \cup \{ R_{i,m+1} ; R_{i,m+2} \}$$

ce qui signifie que, pour l'image suivante, le balayage par l'imagette ne s'effectuera que sur trois régions d'intérêt successives.

Le balayage est réitéré pour chaque image reçue selon le même principe, et ce, jusqu'à ce que le contenu réactualisé du premier sous-ensemble E' inclue la région d'intérêt $R_{i,P}$ correspondant à la plus grande échelle.

Pour fixer les idées, le tableau 1 ci-après montre les régions d'intérêt qui sont examinées à chaque itération sur 7 images successives, en fonction des détections réalisées. Dans cet exemple :

$P=8$, $K=2$ et $m=1$, de sorte que le premier sous-ensemble utilisé pour la première image reçue est

$$E' = \{ R_{1,1} ; R_{1,2} \}$$

et inclut donc ici la région d'intérêt de l'ensemble E prédéterminé correspondant à la plus petite échelle.

Image I_i	Régions d'intérêt balayées	Détections
I_1	$R_{1,1} ; R_{1,2}$	aucune
I_2	$R_{2,2} ; R_{2,3}$	aucune
I_3	$R_{3,3} ; R_{3,4}$	Piéton dans Région $R_{3,4}$
I_4	$R_{4,3} ; R_{4,4} ; R_{4,5}$	Piéton dans Région $R_{4,4}$ et piéton dans Région $R_{4,5}$
I_5	$R_{5,3} ; R_{5,4} ; R_{5,5} ; R_{5,6}$	aucune
I_6	$R_{6,6} ; R_{6,7}$	aucune
I_7	$R_{7,7} ; R_{7,8}$	aucune
I_8	$R_{8,1} ; R_{8,2}$	...

tableau 1

Dans le tableau précédent, on remarque que, tant qu'il n'y a pas de détection (deux première lignes du tableau), le premier ensemble E' est déplacé d'une région d'intérêt, et le balayage par l'imagette ne s'effectue que sur les deux régions d'intérêt de cet ensemble E' . A la première détection effectuée (ligne 3 du tableau), le premier ensemble est réactualisé par déplacement d'une région d'intérêt, mais on conserve également pour l'examen de l'image d'après la région d'intérêt $R_{4,3}$, laquelle correspond à une région contigüe et immédiatement inférieure (dans le sens des échelles croissantes) de celle ayant donné lieu à une détection sur la troisième image. L'avantage est ici d'éviter de perdre le suivi d'un objet qui se déplacerait en direction de la voiture. Certaines régions d'intérêt peuvent être abandonnées à l'itération suivante si elles n'ont pas donné lieu à des détections. Enfin, la totalité des huit régions d'intérêt ayant été balayées entre la première image et la septième image de balayage, la dernière ligne montre que l'on recommence par l'examen des deux premières régions d'intérêt pour la huitième image.

Il convient d'ailleurs de noter que la plage de balayage pour l'ensemble E' ne correspond pas forcément à la totalité de l'ensemble E d'origine, comprenant ici huit régions d'intérêt. On pourrait par exemple se limiter à un déplacement de l'ensemble E' sur les régions d'intérêt correspondant aux plus
5 grandes échelles de l'ensemble E .

L'exemple précédent illustre le gain de traitement que l'on peut atteindre par l'examen d'un nombre limité de régions d'intérêts.

Ce gain de traitement peut encore être amélioré en combinant le procédé de traitement précédent à un traitement de recadrage (« cropping »
10 en terminologie anglo-saxonne) des régions d'intérêt ayant donné lieu à une détection. Pour ce faire, on considère que chaque région d'intérêt $R_{i,m}$ peut être segmentée, dans le sens de la largeur, en une pluralité de zones, par exemple trois zones comme cela est illustré sur la figure 4 par des hachures d'orientation différentes. On note $R_{i,m}^L$ la zone de gauche, $R_{i,m}^C$, la zone
15 centrale, et $R_{i,m}^R$ la zone de droite. Lorsque le traitement d'une image I_i a donné le lieu à une détection dans une zone spécifique d'une région d'intérêt donnée, on recadre, à l'itération d'après, cette région d'intérêt uniquement sur cette zone spécifique. Les zones spécifiques se recouvrent de préférence dans une large mesure de manière à éviter de perdre le suivi d'un objet qui se
20 déplacerait rapidement en latéral dans les images. Le tableau 2 suivant montre en gras les modifications apportées au tableau 1 lorsque l'on fait intervenir le recadrage.

Image I_i	Régions d'intérêt balayées	Détections
I_1	$R_{1,1} ; R_{1,2}$	aucune
I_2	$R_{2,2} ; R_{2,3}$	aucune
I_3	$R_{3,3} ; R_{3,4}$	Piéton à droite dans Région $R_{3,4}$
I_4	$R_{4,3}^R ; R_{4,4}^R ; R_{4,5}$	Piéton dans Région $R_{4,4}^R$ et piéton à droite dans Région $R_{4,5}$
I_5	$R_{5,3}^R ; R_{5,4}^R ; R_{5,5}^R ;$ $R_{5,6}$	aucune
I_6	$R_{6,6} ; R_{6,7}$	aucune
I_7	$R_{7,7} ; R_{7,8}$	aucune
I_8	$R_{8,1} ; R_{8,2}$	...

tableau 2

Les simulations effectuées par la Demanderesse ont pu montrer que la solution décrite ci-dessus peut réduire d'un facteur 10 le temps de traitement nécessaire à l'examen d'une scène (succession d'images) lorsqu'aucun objet n'a été détecté. Bien sûr, ce facteur est moins élevé en cas de détection d'objets, mais le temps de traitement reste tout de même globalement amélioré par rapport à un algorithme classique de Viola et Jones.

Par ailleurs, le procédé de l'invention permet un niveau de précision de détection variable en fonction de la distance à laquelle se situent les objets, avec une précision plus élevée pour les objets proches, ce qui est tout à fait souhaitable lorsque l'on doit appliquer rapidement une stratégie d'évitement.

REVENDICATIONS

1. Procédé de détection d'objets par traitement d'images captées par une caméra (2), dans lequel on utilise un algorithme de détection et/ou de reconnaissance de formes apte à détecter au moins un objet en effectuant, pour chaque image d'une succession d'images, le balayage d'une pluralité de régions d'intérêt par une fenêtre glissante, chaque région d'intérêt correspondant à une échelle différente de chaque image, le procédé étant caractérisé en ce qu'il comprend une étape de sélection (S_1), dans un ensemble (E) prédéterminé de P régions d'intérêt classées dans un ordre défini d'échelles successives, d'un premier sous-ensemble (E') de K régions d'intérêt successives sur lesquelles le balayage par la fenêtre glissante est effectué (S_2) pour au moins une première image.
 2. Procédé selon la revendication 1, caractérisé en ce qu'il comprend en outre les étapes suivantes :
 - on mémorise chaque région d'intérêt du premier sous-ensemble (E') ayant donné lieu à une détection d'un objet ;
 - on réitère (S_5) le balayage pour les images suivantes sur un deuxième sous-ensemble (E'') de régions d'intérêt successives limité d'une part, aux régions d'intérêt mémorisées et aux régions contiguës aux régions d'intérêt mémorisées, et d'autre part, aux K régions d'intérêt du premier sous-ensemble, en déplaçant, à chaque itération, ledit premier sous-ensemble de K régions d'intérêt successives d'au moins une région d'intérêt dans l'ordre défini d'échelles successives.
 3. Procédé selon la revendication 2, caractérisé en ce que ledit premier sous-ensemble (E') de K régions d'intérêt successives est déplacé au plus de K régions d'intérêt.
 4. Procédé selon l'une quelconque des revendications 2 ou 3, caractérisé en ce que le balayage est réitéré jusqu'à ce que le premier sous-ensemble (E') de K régions d'intérêt successives inclue la région

d'intérêt dudit ensemble prédéterminé correspondant à la plus grande échelle.

5. Procédé selon la revendication 4, caractérisé en ce que la plus grande échelle correspond à la taille réelle de l'image captée par la caméra.

5 6. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que ledit premier sous-ensemble (E') de K régions d'intérêt successives inclut la région d'intérêt dudit ensemble prédéterminé correspondant à la plus petite échelle.

10 7. Procédé selon l'une quelconque des revendications précédentes, caractérisé en ce que ledit premier sous-ensemble (E') comprend deux régions d'intérêts successives.

15 8. Procédé selon l'une quelconque des revendications 2 à 5, caractérisé en ce que chaque région d'intérêt est en outre segmentée dans le sens de la largeur en une pluralité de zones d'intérêt, et le procédé comporte en outre une étape de recadrage des régions d'intérêt mémorisées, voire des régions contigües aux régions d'intérêt, le recadrage s'effectuant dans la zone d'intérêt ayant donné lieu à une détection d'objet.

20 9. Système de détection d'objets par traitement d'images captées par une caméra (2), le système comprenant un détecteur (5) utilisant un algorithme de détection et/ou de reconnaissance de formes apte à détecter au moins un objet en effectuant, pour chaque image d'une succession d'images, le balayage d'une pluralité de régions d'intérêt par une fenêtre glissante, chaque région d'intérêt correspondant à une échelle différente de chaque image, caractérisé en ce que ledit
25 détecteur est programmé pour sélectionner (S_1) dans un ensemble (E) prédéterminé de P régions d'intérêt, classées dans un ordre défini d'échelles successives, un premier sous-ensemble (E') de K régions d'intérêt successives sur lesquelles le balayage par la fenêtre glissante est effectué (S_2) pour une première image.

10. Système selon la revendication 9, caractérisé en ce que le détecteur est programmé en outre pour :

- mémoriser chaque région d'intérêt du premier sous-ensemble (E') ayant donné lieu à une détection d'un objet ;

5 - réitérer (S_5) le balayage pour les images suivantes sur un deuxième sous-ensemble (E'') de régions d'intérêt successives limité d'une part, aux régions d'intérêt mémorisées et aux régions contiguës aux régions d'intérêt mémorisées, et d'autre part, aux K régions d'intérêt du premier sous-ensemble, en déplaçant, à chaque itération, ledit premier sous-
10 ensemble de K régions d'intérêt successives d'au moins une région d'intérêt dans l'ordre défini d'échelles successives.

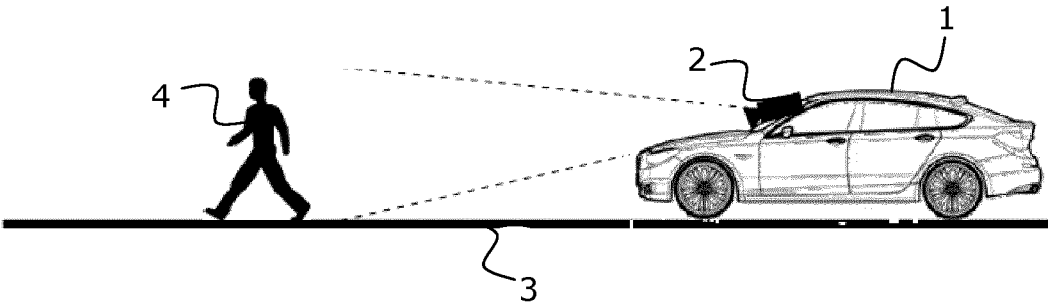


FIG. 1

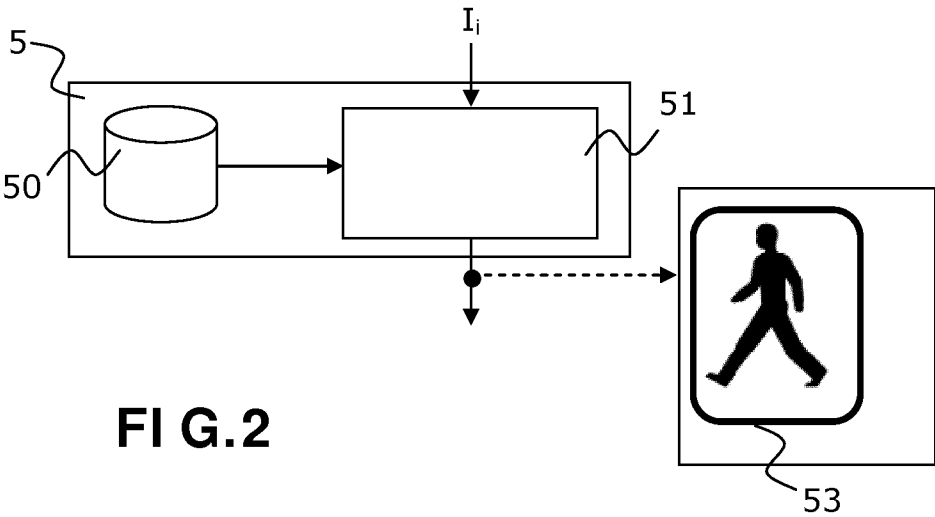


FIG. 2

2 / 2

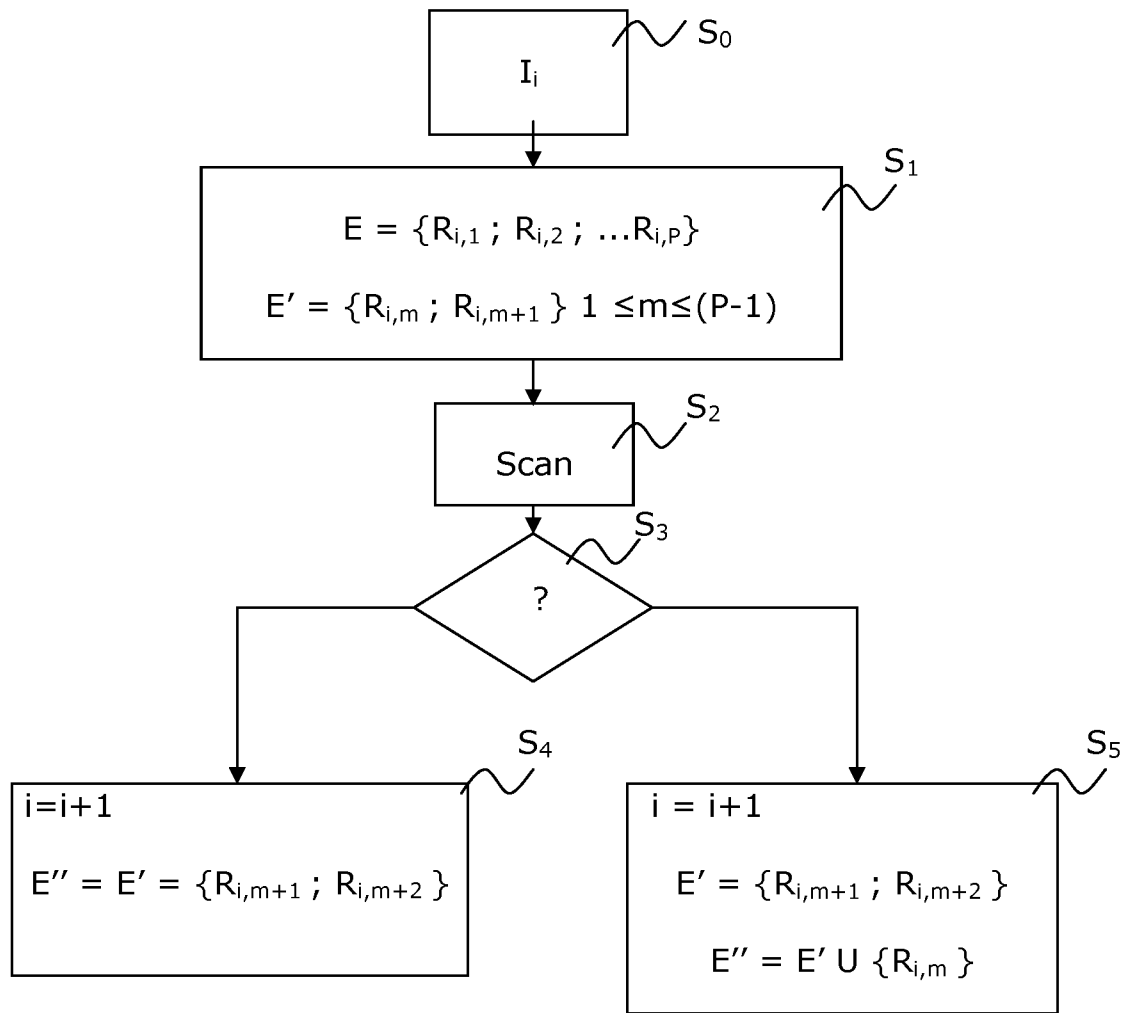


FIG. 3

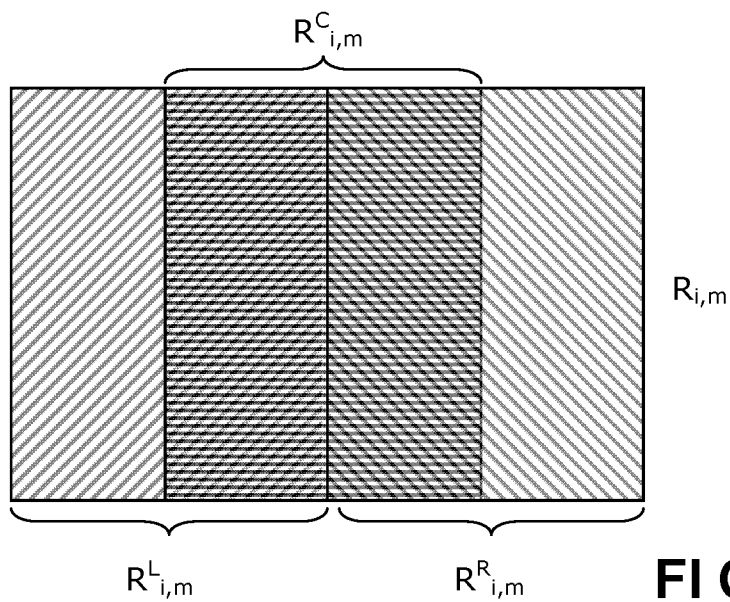


FIG. 4

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2016/066381

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06K9/00
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Shaoqing Ren ET AL: "Object Detection Networks on Convolutional Feature Maps", CoRR, 23 April 2015 (2015-04-23), XP055255038, Retrieved from the Internet: URL:http://arxiv.org/pdf/1504.06066v1.pdf [retrieved on 2016-03-03]	1,6,7,9
A	abstract section 5.3 section 7 ----- -/--	2-5,8,10



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

23 September 2016

Date of mailing of the international search report

30/09/2016

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Martinez, Francis

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2016/066381

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FEIDIE LIANG ET AL: "Fast Pedestrian Detection Based on Sliding Window Filtering", 4 December 2012 (2012-12-04), ADVANCES IN MULTIMEDIA INFORMATION PROCESSING PCM 2012, SPRINGER BERLIN HEIDELBERG, BERLIN, HEIDELBERG, PAGE(S) 811 - 822, XP047014270, ISBN: 978-3-642-34777-1	1,6,7,9
A	abstract section 1 -----	2-5,8
A	Alejandro González ET AL: "3D-Guided Multiscale Sliding Window for Pedestrian Detection" In: "Correct System Design", 9 June 2015 (2015-06-09), Springer International Publishing, Cham 032548, XP055255297, ISSN: 0302-9743 ISBN: 978-3-642-28871-5 vol. 9117, pages 560-568, DOI: 10.1007/978-3-319-19390-8_63, section 3 -----	1-10

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°

PCT/EP2016/066381

A. CLASSEMENT DE L'OBJET DE LA DEMANDE INV. G06K9/00 ADD.		
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB		
B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE Documentation minimale consultée (système de classification suivi des symboles de classement) G06K		
Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche		
Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés) EPO-Internal, WPI Data		
C. DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	Shaoqing Ren ET AL: "Object Detection Networks on Convolutional Feature Maps", CoRR, 23 avril 2015 (2015-04-23), XP055255038, Extrait de l'Internet: URL:http://arxiv.org/pdf/1504.06066v1.pdf [extrait le 2016-03-03]	1,6,7,9
A	abrégé section 5.3 section 7 ----- -/--	2-5,8,10
☒ Voir la suite du cadre C pour la fin de la liste des documents ☐ Les documents de familles de brevets sont indiqués en annexe		
* Catégories spéciales de documents cités:		
"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent "E" document antérieur, mais publié à la date de dépôt international ou après cette date "L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée) "O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens "P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée "T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention "X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément "Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier "&" document qui fait partie de la même famille de brevets		
Date à laquelle la recherche internationale a été effectivement achevée 23 septembre 2016		Date d'expédition du présent rapport de recherche internationale 30/09/2016
Nom et adresse postale de l'administration chargée de la recherche internationale Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Fonctionnaire autorisé Martinez, Francis

C(suite). DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	FEIDIE LIANG ET AL: "Fast Pedestrian Detection Based on Sliding Window Filtering", 4 décembre 2012 (2012-12-04), ADVANCES IN MULTIMEDIA INFORMATION PROCESSING PCM 2012, SPRINGER BERLIN HEIDELBERG, BERLIN, HEIDELBERG, PAGE(S) 811 - 822, XP047014270, ISBN: 978-3-642-34777-1	1,6,7,9
A	abrégé section 1	2-5,8
A	----- Alejandro González ET AL: "3D-Guided Multiscale Sliding Window for Pedestrian Detection" In: "Correct System Design", 9 juin 2015 (2015-06-09), Springer International Publishing, Cham 032548, XP055255297, ISSN: 0302-9743 ISBN: 978-3-642-28871-5 vol. 9117, pages 560-568, DOI: 10.1007/978-3-319-19390-8_63, section 3 -----	1-10