

12

DEMANDE DE BREVET D'INVENTION

A1

22 Date de dépôt : 28.02.19.

30 Priorité :

43 Date de mise à la disposition du public de la demande : 04.09.20 Bulletin 20/36.

56 Liste des documents cités dans le rapport de recherche préliminaire : *Se reporter à la fin du présent fascicule*

60 Références à d'autres documents nationaux apparentés :

Demanda(s) d'extension :

71 Demandeur(s) : PSA Automobiles SA Société anonyme — FR.

72 Inventeur(s) : LARGE FREDERIC et BOUTARD PATRICK.

73 Titulaire(s) : PSA Automobiles SA Société anonyme.

74 Mandataire(s) :

54 APPRENTISSAGE D'UN RÉSEAU DE NEURONES PAR COMPARAISONS DE VALEURS PRODUITES PAR UN SIMULATEUR DE CONDUITE D'UN VÉHICULE FONCTIONNANT EN PARALLÈLE.

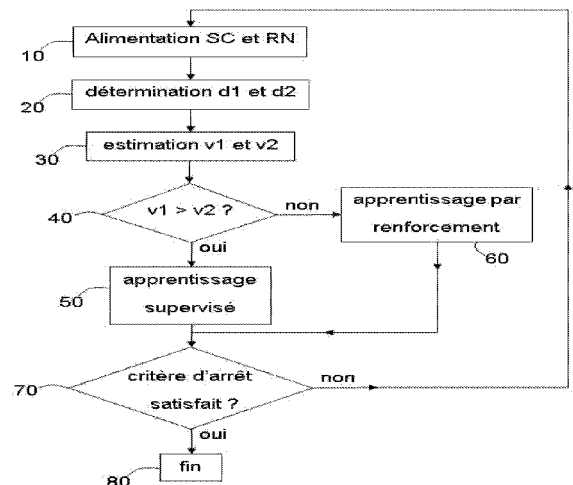
57 Un procédé forme un réseau de neurones déterminant des décisions de conduite pour un véhicule virtuel à conduite autonome. Ce procédé comprend :

- une première étape (10-20) où l'on fait fonctionner en parallèle, avec un ensemble de données d'environnement et un trajet ou une manœuvre, un simulateur de conduite pour qu'il détermine des premières décisions de conduite avec un ensemble de règles prédéfinies, et le réseau de neurones pour qu'il détermine des secondes décisions de conduite,

- une deuxième étape (30) où l'on estime des première et seconde valeurs représentatives de performances des premières et secondes décisions de conduite, et

- une troisième étape (40-80) où l'on modifie le réseau de neurones par apprentissage de type supervisé ou renforcement selon que la première valeur est supérieure ou inférieure à la seconde valeur, puis on réitère les première (10-20) à troisième (40-80) étapes tant qu'un critère d'arrêt n'est pas satisfait.

Figure à publier avec l'abrégié: Fig. 2



## **Description**

### **Titre de l'invention : APPRENTISSAGE D'UN RÉSEAU DE NEURONES PAR COMPARAISONS DE VALEURS PRODUITES PAR UN SIMULATEUR DE CONDUITE D'UN VÉHICULE FONCTIONNANT EN PARALÈLE**

#### **Domaine technique de l'invention**

[0001] L'invention concerne les véhicules terrestres à conduite au moins partiellement automatisée (ou autonome) et éventuellement de type automobile, et plus précisément l'apprentissage des réseaux de neurones qui sont chargés de déterminer des décisions de conduite pour de tels véhicules.

#### **Etat de la technique**

[0002] Dans ce qui suit on considère qu'un véhicule terrestre est à conduite au moins partiellement automatisée (ou autonome) lorsqu'il peut être conduit sur une route de façon automatisée (partielle ou totale (sans intervention de son conducteur)) pendant une phase de conduite automatisée, ou de façon manuelle (et donc avec intervention de son conducteur sur le volant et/ou les pédales) pendant une phase de conduite manuelle. Par exemple, la conduite automatisée (ou autonome) d'un véhicule (à conduite au moins partiellement automatisée) peut consister à diriger partiellement ou totalement ce véhicule ou à apporter tout type d'aide à une personne physique conduisant ce véhicule. Cela couvre donc toute conduite automatisée (ou autonome), du niveau 1 au niveau 5 dans le barème de l'OICA (Organisation International des Constructeurs Automobiles).

[0003] Par ailleurs, on entend ici par « véhicule terrestre » tout type de véhicule pouvant circuler sur une route, et notamment un véhicule automobile, un véhicule utilitaire, un cyclomoteur, une motocyclette, un minibus, un car, un robot de stockage dans un entrepôt, ou un engin de voirie.

[0004] De nombreux procédés d'assistance à la conduite de véhicules (à conduite au moins partiellement automatisée) ont pour but d'adapter la conduite de ces véhicules en fonction d'informations déterminées dans leur environnement à partir d'une analyse de données acquises dans ce dernier par leurs capteurs. Cette adaptation se fait en fonction de décisions de conduite qui peuvent être déterminées par un réseau de neurones ayant été préalablement formé (ou entraîné) à cet effet, et éventuellement capable de généralisation sur des cas non appris (voire d'optimisation sur un critère donné). Par exemple, lorsque le réseau de neurones a été formé (ou entraîné) à partir d'essais/erreurs, son apprentissage est dit « par renforcement ». Chaque essai/erreur est sanctionné par un

système de bonus/malus sans que la solution attendue ne soit explicitement fournie au système (car souvent non connue). L'exploitation de règles pour définir le bonus/malus peut être une option.

[0005] Un inconvénient de ce type de réseau de neurones réside dans le fait qu'il est formé sans que l'on connaisse son niveau minimum de performance, et donc la pertinence des décisions de conduite qu'il prend peut varier très notablement d'un environnement (correspondant à une situation particulière) à un autre environnement (correspondant à une autre situation particulière), voire parfois peut être non pertinente.

[0006] L'invention a donc notamment pour but d'améliorer la situation.

### **Présentation de l'invention**

[0007] Elle propose notamment à cet effet un procédé d'apprentissage destiné à former un réseau de neurones chargé de déterminer des décisions de conduite pour un véhicule virtuel à conduite au moins partiellement autonome.

[0008] Ce procédé d'apprentissage se caractérise par le fait qu'il comprend :

[0009] - une première étape dans laquelle on fait fonctionner en parallèle, avec un même ensemble de données d'environnement définissant un environnement du véhicule et une définition d'un trajet ou d'une manœuvre, d'une part, un simulateur de conduite pour qu'il détermine au moins une première décision de conduite pour le véhicule avec un ensemble de règles prédéfinies, et, d'autre part, le réseau de neurones pour qu'il détermine au moins une seconde décision de conduite pour le véhicule,

[0010] - une deuxième étape dans laquelle on estime des première et seconde valeurs représentatives respectivement de performances des première(s) et seconde(s) décisions de conduite, et

[0011] - une troisième étape dans laquelle on modifie de façon automatisée le réseau de neurones soit par apprentissage de type supervisé lorsque la première valeur est supérieure à la seconde valeur, soit par apprentissage de type renforcement lorsque la première valeur est inférieure à la seconde valeur, puis on réitère les première à troisième étapes, avec à chaque fois un nouvel ensemble de données d'environnement et une nouvelle définition de trajet ou de manœuvre, tant qu'un critère d'arrêt n'est pas satisfait.

[0012] Ainsi, lorsque le critère d'arrêt est satisfait (hormis en cas de collision) on considère que le réseau de neurones est devenu suffisamment performant. Lorsque le critère d'arrêt est une collision et qu'il est satisfait, on n'a pas besoin d'aller plus loin dans la simulation, car on sait que la suite de commandes choisie n'est pas la bonne.

[0013] Le procédé d'apprentissage selon l'invention peut comporter d'autres caractéristiques qui peuvent être prises séparément ou en combinaison, et notamment :

[0014] - le critère d'arrêt peut, par exemple, être choisi parmi un niveau de performance

global, fonction des secondes valeurs déterminées successivement lors des itérations, et dépassant un seuil prédéfini, un taux d'erreur devenant inférieur à un seuil prédéfini, une évolution des secondes valeurs, déterminées successivement lors des itérations, qui tend vers une valeur prédéfinie et dépasse un seuil prédéfini, et un nombre n, de fois où les secondes valeurs déterminées ont été supérieures aux premières valeurs correspondantes, qui atteint un seuil prédéfini ;

- [0015] - dans sa troisième étape la modification du réseau de neurones par apprentissage de type supervisé peut consister à fournir ou indiquer au réseau de neurones l'ensemble de règles devant être respectées pour réaliser correctement un trajet ou une manœuvre venant d'être effectué(e) ;
- [0016] - dans sa troisième étape la modification du réseau de neurones par apprentissage de type renforcement peut consister à modifier une représentation personnelle via une adaptation d'une structure interne ;
- [0017] - dans sa deuxième étape on peut estimer les première et seconde valeurs en fonction d'au moins un critère de performance choisi ;
- [0018] + chaque critère de performance peut, par exemple, être choisi parmi un confort de passagers du véhicule, une influence sur un temps de trajet ou de manœuvre du véhicule, un risque encouru par le véhicule, une stabilité des décisions (comme par exemple le fait que l'on n'oscille pas entre une décision de changer de voie et une décision de rester dans la voie), et une utilisation d'accélération maximales.
- [0019] L'invention propose également un produit programme d'ordinateur comprenant un jeu d'instructions qui, lorsqu'il est exécuté par des moyens de traitement, est propre à mettre en œuvre un procédé d'apprentissage du type de celui présenté ci-avant pour former un réseau de neurones chargé de déterminer des décisions de conduite pour un véhicule virtuel à conduite au moins partiellement autonome.
- [0020] L'invention propose également un dispositif d'apprentissage destiné à former un réseau de neurones chargé de déterminer des décisions de conduite pour un véhicule virtuel à conduite au moins partiellement autonome.
- [0021] Ce dispositif d'apprentissage se caractérise par le fait qu'il comprend au moins un processeur et au moins une mémoire agencés pour effectuer les opérations consistant à :
- [0022] - déclencher le fonctionnement en parallèle, avec un même ensemble de données d'environnement définissant un environnement du véhicule et une définition d'un trajet ou d'une manœuvre, d'une part, d'un simulateur de conduite pour qu'il détermine au moins une première décision de conduite pour le véhicule avec un ensemble de règles prédéfinies, et, d'autre part, du réseau de neurones pour qu'il détermine au moins une seconde décision de conduite pour le véhicule,
- [0023] - estimer des première et seconde valeurs représentatives respectivement de per-

formances des première(s) et seconde(s) décisions de conduite,

[0024] - modifier le réseau de neurones soit par apprentissage de type supervisé lorsque la première valeur est supérieure à la seconde valeur, soit par apprentissage de type renforcement lorsque la première valeur est inférieure à la seconde valeur, puis

[0025] - réitérer ces déclenchement, estimation et modification, avec à chaque fois un nouvel ensemble de données d'environnement et une nouvelle définition de trajet ou de manœuvre, tant qu'un critère d'arrêt n'est pas satisfait.

[0026] L'invention propose également un ordinateur comprenant, d'une part, un simulateur de conduite destiné à déterminer, avec un ensemble de règles prédéfinies, des décisions de conduite pour un véhicule virtuel à conduite au moins partiellement autonome, et, d'autre part, un dispositif d'apprentissage du type de celui présenté ci-avant et couplé à un réseau de neurones chargé de déterminer des décisions de conduite pour le véhicule.

### **Brève description des figures**

[0027] D'autres caractéristiques et avantages de l'invention apparaîtront à l'examen de la description détaillée ci-après, et des dessins annexés, sur lesquels :

[0028] [fig.1]

illustre schématiquement et fonctionnellement un exemple de réalisation d'un ordinateur de simulation équipé d'un exemple de réalisation d'un dispositif d'apprentissage selon l'invention et d'un réseau de neurones à former, et

[0029] [fig.2]

illustre schématiquement un exemple d'algorithme mettant en œuvre un procédé d'apprentissage selon l'invention.

### **Description détaillée de l'invention**

[0030] L'invention a notamment pour but de proposer un procédé d'apprentissage, et un dispositif d'apprentissage DA associé, destinés à permettre l'apprentissage (ou la formation ou encore l'entraînement) d'un réseau de neurones RN chargé de déterminer des décisions de conduite pour un véhicule virtuel à conduite au moins partiellement autonome. On comprendra qu'une fois le réseau de neurones RN formé (ou entraîné) il est destiné à être implanté dans un dispositif d'assistance destiné à assister la conduite d'un véhicule réel du même type que le véhicule virtuel avec lequel il a été formé (ou entraîné).

[0031] Dans ce qui suit, on considère, à titre d'exemple non limitatif, que le véhicule virtuel à conduite automatisée (ou autonome) est de type automobile. Il s'agit par exemple d'une voiture. Mais l'invention n'est pas limitée à ce type de véhicule virtuel. Elle concerne en effet tout type de véhicule terrestre à conduite au moins partiellement automatisée et pouvant circuler sur des voies de circulation terrestres. Ainsi, il pourra aussi s'agir d'un véhicule utilitaire, d'un cyclomoteur, d'une motocyclette, d'un

minibus, d'un car, d'un robot de stockage dans un entrepôt, ou d'un engin de voirie, par exemple.

- [0032] On notera que l'invention est destinée à permettre la conduite automatisée (au moins partiellement) de véhicules sur toute portion de route comprenant au moins une voie de circulation délimitée par deux délimitations.
- [0033] Comme évoqué plus haut, l'invention propose notamment un procédé d'apprentissage destiné à permettre l'apprentissage (ou la formation ou encore l'entraînement) d'un réseau de neurones RN chargé de déterminer des décisions de conduite pour un véhicule virtuel à conduite au moins partiellement autonome.
- [0034] Ce procédé d'apprentissage peut être au moins partiellement mis en œuvre par le dispositif d'apprentissage DA qui comprend à cet effet au moins un processeur PR, par exemple de signal numérique (ou DSP (« Digital Signal Processor »)), et une mémoire MD.
- [0035] Dans l'exemple illustré non limitativement sur la figure 1, le dispositif d'apprentissage DA fait partie d'un ordinateur OS qui comprend également un simulateur de conduite SC et le réseau de neurones RN devant être formé (ou entraîné). Mais cela n'est pas obligatoire. En effet, le dispositif d'apprentissage DA pourrait comprendre son propre calculateur et être couplé à un ordinateur OS comprenant un simulateur de conduite SC et le réseau de neurones RN devant être formé (ou entraîné). Par conséquent, le dispositif d'apprentissage DA peut être réalisé sous la forme d'une combinaison de circuits ou composants électriques ou électroniques (ou « hardware ») et de modules logiciels (ou « software »). La mémoire MD est vive afin de stocker des instructions pour la mise en œuvre par le processeur PR du procédé d'apprentissage.
- [0036] On notera que l'ordinateur OS peut être fixe ou portable ou bien peut faire partie d'un serveur. Il peut être matériel ou virtuel (par exemple réparti dans un nuage (ou « cloud »)).
- [0037] Le simulateur de conduite SC est agencé de manière à déterminer, en fonction d'un ensemble de règles prédéfinies, une première décision de conduite pour le véhicule virtuel chaque fois qu'il est alimenté avec des données d'environnement et une définition d'un trajet ou d'une manœuvre. Ces données d'environnement sont similaires à celles qui sont acquises par le(s) capteur(s) équipant un véhicule réel. Elles sont représentatives de l'environnement qui est situé au moins devant ce véhicule et, de préférence, sur une partie au moins des deux côtés latéraux de ce véhicule, ainsi qu'éventuellement derrière ce véhicule.
- [0038] Chaque capteur peut comprendre au moins une caméra installée dans une partie avant du véhicule (par exemple sur le pare-brise ou sur le rétroviseur intérieur) ou dans une partie arrière, ou au moins un capteur à ultrasons ou au moins un radar ou lidar.
- [0039] Comme illustré sur la figure 2, le procédé d'apprentissage, selon l'invention,

comprend trois étapes effectuées en série de façon réitérée.

- [0040] Dans une première étape 10-20 du procédé d'apprentissage, on (le processeur PR et la mémoire MD) alimente(nt) le simulateur de conduite SC et le réseau de neurones RN à former (ou entraîner) avec un même ensemble de données d'environnement définissant un environnement du véhicule virtuel (au moins devant ce dernier) et une définition d'un trajet ou d'une manœuvre. Cette alimentation est destinée à faire fonctionner en parallèle avec cet ensemble de données d'environnement sur ce trajet défini ou cette manœuvre définie, d'une part, le simulateur de conduite SC pour qu'il détermine au moins une première décision de conduite d1 pour le véhicule virtuel (avec son ensemble de règles prédéfinies), et, d'autre part, le réseau de neurones RN pour qu'il détermine au moins une seconde décision de conduite d2 pour le véhicule virtuel.
- [0041] Dans une deuxième étape 30 du procédé d'apprentissage, on (le processeur PR et la mémoire MD) estime(nt) des première v1 et seconde v2 valeurs qui sont représentatives respectivement de performances des première(s) d1 et seconde(s) d2 décisions de conduite venant d'être déterminées dans la première étape 10.
- [0042] On notera que les première v1 et seconde v2 valeurs peuvent être numériques ou alphanumériques.
- [0043] Par exemple, dans cette deuxième étape 30 on (le processeur PR et la mémoire MD) peu(ven)t estimer les première v1 et seconde v2 valeurs en fonction d'au moins un critère de performance choisi.
- [0044] A titre d'exemple, chaque critère de performance peut être choisi parmi le confort des passagers du véhicule, l'influence sur le temps de trajet ou de manœuvre du véhicule, le risque encouru par le véhicule, la stabilité des décisions (comme par exemple le fait que l'on n'oscille pas entre une décision de changer de voie et une décision de rester dans la voie), et l'utilisation d'accélération maximales. Mais tout autre critère de performance connu de l'homme de l'art peut être ici utilisé.
- [0045] Plus le confort des passagers du véhicule est respecté, plus la performance est élevée. Pour ce faire, on doit avoir une bonne gestion des accélérations (au moins longitudinales et si possible aussi transversales), des freinages et de l'angle de braquage des roues directrices du véhicule. D'une manière générale, on cherche à éviter autant que possible les oscillations longitudinales ou latérales qui sont sources de malaise, et les accélérations ou secousses trop importantes qui sont sources de stress ou d'inconfort. Par exemple, on peut vérifier si les accélérations décidées sont restées dans un intervalle de valeurs prédéfini et/ou si les intensités des freinages sont restées dans un intervalle de valeurs prédéfini et/ou si les angles de braquage sont restés dans un intervalle de valeurs prédéfini.
- [0046] Plus le temps de trajet ou de manœuvre est court, plus la performance est a priori

élevée sous réserve, bien entendu, que le confort ne soit pas (ou peu) diminué. Pour ce faire, on doit avoir une bonne gestion des dépassements, de la vitesse et de l'accélération.

- [0047] Plus le risque encouru est faible, plus la performance est élevée, sous réserve, bien entendu, que le véhicule bouge. Pour ce faire, on doit avoir une bonne gestion des évitements des obstacles présents dans l'environnement du véhicule et des distances entre le véhicule et les obstacles qui l'environnent, ainsi qu'une bonne maîtrise des règles de conduite et sociales.
- [0048] Dans une troisième étape 40-80 du procédé d'apprentissage, on (le processeur PR et la mémoire MD) modifie(nt) (ou corrige(nt)) de façon automatisée le réseau de neurones RN soit par apprentissage de type supervisé lorsque la première valeur  $v_1$  est supérieure à la seconde valeur  $v_2$ , soit par apprentissage de type renforcement lorsque la première valeur  $v_1$  est inférieure à la seconde valeur  $v_2$ .
- [0049] Puis, on (le processeur PR et la mémoire MD) réitère(nt) les première 10-20 à troisième 40-80 étapes, avec à chaque fois un nouvel ensemble de données d'environnement et une nouvelle définition de trajet ou de manœuvre, tant qu'un critère d'arrêt n'est pas satisfait.
- [0050] Grâce à l'invention, il est désormais possible de former (ou entraîner) un réseau de neurones de sorte qu'il présente un niveau minimum de performance (supérieur à celui du simulateur de conduite SC). En effet, lorsque le critère d'arrêt est satisfait (hormis en cas de collision) on considère que le réseau de neurones est devenu suffisamment performant. Lorsque le critère d'arrêt est une collision et qu'il est satisfait, on n'a pas besoin d'aller plus loin dans la simulation, car on sait que la suite de commandes choisie n'est pas la bonne et donc qu'il n'est pas nécessaire de poursuivre. Ainsi, plus le système de décision rencontrera de situations différentes lors de son fonctionnement en parallèle avec le simulateur de conduite SC, plus il saura s'adapter à de nouveaux cas.
- [0051] Par exemple, le critère d'arrêt qui est utilisé dans chaque troisième étape 40-80 peut être choisi parmi :
- [0052] - un niveau de performance global, fonction des secondes valeurs  $v_2$  qui ont été déterminées successivement lors des différentes itérations, et dépassant un seuil prédéfini,
- [0053] - un taux d'erreur devenant inférieur à un seuil prédéfini,
- [0054] - une évolution des secondes valeurs  $v_2$ , déterminées successivement lors des différentes itérations, qui tend vers une valeur prédéfinie et dépasse un seuil prédéfini, et
- [0055] - un nombre  $n$ , de fois où les secondes valeurs  $v_2$  déterminées ont été supérieures aux premières valeurs  $v_1$  correspondantes, qui atteint un seuil  $N$  prédéfini.
- [0056] Le niveau de performance global peut, par exemple, être égal à la somme ou la

moyenne des secondes valeurs  $v_2$  déterminées successivement lors des différentes itérations. Le seuil prédéfini est au moins égal au niveau de performance global du simulateur de conduite SC. On comprendra que plus ce seuil sera élevé, plus il sera difficile de l'atteindre. Après chaque itération on met à jour le niveau de performance global, et lorsqu'il devient supérieur au seuil prédéfini on arrête les itérations et donc on cesse l'apprentissage du réseau de neurones RN. On comprendra que dans ce cas, on doit effectuer un nombre minimal  $P$  d'itérations avant de décider d'arrêter le procédé. En effet, on pourrait avoir au début du procédé un niveau de performance global supérieur au seuil prédéfini. Par exemple (et non limitativement), le nombre minimal  $P$  peut être compris entre 100 et 200.

[0057] Le taux d'erreur peut, par exemple, être égal au nombre total d'erreurs divisé par le nombre d'itérations. Après chaque itération on met à jour le taux d'erreur, et lorsqu'il devient inférieur au seuil prédéfini on arrête les itérations et donc on cesse l'apprentissage du réseau de neurones RN. On comprendra que dans ce cas, on doit effectuer un nombre minimal  $P$  d'itérations avant de décider d'arrêter le procédé. En effet, on pourrait avoir au début du procédé un taux d'erreur inférieur au seuil prédéfini. Par exemple (et non limitativement), le nombre minimal  $P$  peut être compris entre 100 et 200.

[0058] C'est en comparant entre elles les secondes valeurs  $v_2$  qui sont déterminées successivement lors des différentes itérations que l'on peut déterminer si elles tendent progressivement (asymptotiquement) vers une valeur prédéfinie, et c'est en comparant chaque seconde valeur  $v_2$  au seuil prédéfini que l'on peut déterminer que lors d'une itération les secondes valeurs  $v_2$  ont dépassé ce seuil prédéfini. L'évolution des secondes valeurs  $v_2$  peut se faire par le calcul, par exemple en calculant des gradients. Après chaque itération on met à jour l'évolution et lorsqu'elle tend vers la valeur prédéfinie et que dans le même temps la dernière seconde valeur  $v_2$  devient supérieure au seuil prédéfini on arrête les itérations et donc on cesse l'apprentissage du réseau de neurones RN.

[0059] Le nombre  $n$  peut être déterminé lors de chaque itération en incrémentant d'une unité un compteur ( $n$ ) chaque fois qu'une seconde valeur  $v_2$  déterminée est supérieure à la première valeur  $v_1$  correspondante. Après chaque incrémentation on compare la valeur en cours du compteur  $n$  au seuil  $N$  prédéfini, et lorsque cette valeur en cours du compteur  $n$  devient égale à  $N$  on arrête les itérations et donc on cesse l'apprentissage du réseau de neurones RN. Par exemple (et non limitativement), le nombre  $N$  peut être compris entre 100 et 200.

[0060] On notera que dans la troisième étape 40-80 la modification du réseau de neurones RN par apprentissage de type supervisé peut, par exemple, se faire par une modification de sa représentation personnelle via une adaptation de sa structure interne (par

exemple en changeant des pondérations entre ses neurones).

- [0061] On notera également que dans la troisième étape 40-80 la modification du réseau de neurones RN par apprentissage de type renforcement peut, par exemple, consister à fournir à ce dernier (RN) une tendance représentative de sa performance sur un trajet ou une manœuvre venant d'être effectué(e). Par exemple, on peut indiquer au réseau de neurones RN si le trajet (ou la manœuvre) était bon(ne) ou mauvais(e) ou encore moyen(ne), ou bien on peut fournir au réseau de neurones RN une valeur (par exemple comprise entre 1 et 5 ou 1 et 10) représentative de la qualité du trajet ou de la manœuvre effectué(e), ou encore on peut fournir au réseau de neurones RN le résultat de la comparaison de chaque seconde  $v_2$  à la première valeur  $v_1$  correspondante durant un trajet ou une manœuvre.
- [0062] On va décrire ci-après, en référence à la figure 2, un exemple d'algorithme mettant en œuvre les trois étapes du procédé d'apprentissage décrit ci-avant.
- [0063] Dans une première sous-étape 10 de la première étape 10-20, on (le processeur PR et la mémoire MD) alimente(nt) le simulateur de conduite SC et le réseau de neurones RN avec un même ensemble de données d'environnement définissant un environnement du véhicule virtuel (au moins devant ce dernier) et une définition d'un trajet ou d'une manœuvre (à effectuer), afin de les faire fonctionner en parallèle avec cet ensemble de données d'environnement sur le trajet (ou la manœuvre) défini(e). Dans une deuxième sous-étape 20 de la première étape 10-20, le simulateur de conduite SC détermine au moins une première décision de conduite  $d_1$  pour le véhicule virtuel (avec son ensemble de règles prédéfinies), et le réseau de neurones RN détermine au moins une seconde décision de conduite  $d_2$  pour le véhicule virtuel.
- [0064] Dans une troisième sous-étape 30 de la deuxième étape 30, on (le processeur PR et la mémoire MD) estime(nt) des première  $v_1$  et seconde  $v_2$  valeurs représentatives respectivement des performances des première(s)  $d_1$  et seconde(s)  $d_2$  décisions de conduite sur le trajet (ou la manœuvre) défini(e).
- [0065] Dans une quatrième sous-étape 40 de la troisième étape 40-80, on (le processeur PR et la mémoire MD) effectue(nt) un test pour comparer entre elles les première  $v_1$  et seconde  $v_2$  valeurs. Si la première valeur  $v_1$  est supérieure à la seconde valeur  $v_2$  (soit  $v_1 > v_2$ ), on (le processeur PR et la mémoire MD) modifie(nt) de façon automatisée le réseau de neurones RN par apprentissage de type supervisé dans une cinquième sous-étape 50 de la troisième étape 40-80. En revanche, si la première valeur  $v_1$  est inférieure à la seconde valeur  $v_2$ , on (le processeur PR et la mémoire MD) modifie(nt) de façon automatisée le réseau de neurones RN par apprentissage de type renforcement dans une sixième sous-étape 60 de la troisième étape 40-80.
- [0066] Puis, dans une septième sous-étape 70 de la troisième étape 40-80, on (le processeur PR et la mémoire MD) effectue(nt) un test pour déterminer si le critère d'arrêt est

satisfait. Dans la négative (« non »), on (le processeur PR et la mémoire MD) réitère(nt) les première 10-20 à troisième 40-80 étapes, avec à chaque fois un nouvel ensemble de données d'environnement et une nouvelle définition de trajet (ou de manœuvre), tant que le critère d'arrêt n'est pas satisfait. En revanche, dans l'affirmative (« oui »), on (le processeur PR et la mémoire MD) mettent fin à l'apprentissage du réseau de neurones RN dans une huitième sous-étape 80 de la troisième étape 40-80.

[0067] On notera également que l'invention propose aussi un produit programme d'ordinateur comprenant un jeu d'instructions qui, lorsqu'il est exécuté par des moyens de traitement de type circuits électroniques (ou hardware), comme par exemple le processeur PR, est propre à mettre en œuvre le procédé d'apprentissage pour former un réseau de neurones RN chargé de déterminer des décisions de conduite pour un véhicule virtuel à conduite au moins partiellement autonome.

[0068] On notera également que sur la figure 1 le dispositif d'apprentissage DA est très schématiquement illustré avec sa mémoire vive MD et son processeur PR qui peut comprendre des circuits intégrés (ou imprimés), ou bien plusieurs circuits intégrés (ou imprimés) reliés par des connections filaires ou non filaires. On entend par circuit intégré (ou imprimé) tout type de dispositif apte à effectuer au moins une opération électrique ou électronique. Mais, comme illustré non limitativement sur la figure 1, le dispositif d'apprentissage DA peut aussi comprendre une mémoire de masse MM, notamment pour le stockage des données d'environnement des ensembles qu'il fournit au réseau de neurones RN à former, des définitions de trajets ou de manœuvre, des premières d1 et secondes d2 décisions de conduite des différentes itérations, des premières v1 et secondes v2 valeurs déterminées dans les différentes itérations, et d'éventuelles données intermédiaires intervenant dans tous ses calculs et traitements. Par ailleurs, ce dispositif d'apprentissage DA peut aussi comprendre une interface d'entrée IE pour la réception d'au moins les premières d1 et secondes d2 décisions de conduite, pour les utiliser dans des calculs ou traitements, éventuellement après les avoir mises en forme et/ou démodulées et/ou amplifiées, de façon connue en soi, au moyen d'un processeur de signal numérique PR'. De plus, ce dispositif d'apprentissage DA peut aussi comprendre une interface de sortie IS, notamment pour la transmission, d'une part, au simulateur de conduite SC et au réseau de neurones RN des ensembles de données d'environnement et des définitions de trajet ou de manœuvre, et, d'autre part, au réseau de neurones RN des modifications (ou corrections) d'apprentissage déterminées à chaque itération.

[0069] On notera également qu'une ou plusieurs étapes et/ou une ou plusieurs sous-étapes du procédé d'apprentissage à la conduite peuvent être effectuées par des composants différents. Ainsi, le procédé d'apprentissage à la conduite peut-être mis en œuvre par

une pluralité de processeurs de signal numérique, mémoire vive, mémoire de masse, interface d'entrée, interface de sortie.

## Revendications

- [Revendication 1] Procédé d'apprentissage d'un réseau de neurones (RN) chargé de déterminer des décisions de conduite pour un véhicule virtuel à conduite au moins partiellement autonome, caractérisé en ce qu'il comprend :
- une première étape (10-20) dans laquelle on fait fonctionner en parallèle, avec un même ensemble de données d'environnement définissant un environnement dudit véhicule et une définition d'un trajet ou d'une manœuvre, i) un simulateur de conduite (SC) pour qu'il détermine au moins une première décision de conduite pour ledit véhicule avec un ensemble de règles prédéfinies, et ii) ledit réseau de neurones (RN) pour qu'il détermine au moins une seconde décision de conduite pour ledit véhicule,
  - une deuxième étape (30) dans laquelle on estime des première et seconde valeurs représentatives respectivement de performances desdites première(s) et seconde(s) décisions de conduite, et
  - une troisième étape (40-80) dans laquelle on modifie de façon automatisée ledit réseau de neurones (RN) soit par apprentissage de type supervisé lorsque ladite première valeur est supérieure à ladite seconde valeur, soit par apprentissage de type renforcement lorsque ladite première valeur est inférieure à ladite seconde valeur, puis on réitère lesdites première (10-20) à troisième (40-80) étapes, avec à chaque fois un nouvel ensemble de données d'environnement et une nouvelle définition de trajet ou de manœuvre, tant qu'un critère d'arrêt n'est pas satisfait.
- [Revendication 2] Procédé selon la revendication 1, caractérisé en ce que ledit critère d'arrêt est choisi parmi i) un niveau de performance global, fonction desdites secondes valeurs déterminées successivement lors desdites itérations, et dépassant un seuil prédéfini, ii) un taux d'erreur devenant inférieur à un seuil prédéfini, iii) une évolution desdites secondes valeurs, déterminées successivement lors desdites itérations, qui tend vers une valeur prédéfinie et dépasse un seuil prédéfini, et iv) un nombre n, de fois où lesdites secondes valeurs déterminées ont été supérieures aux premières valeurs correspondantes, qui atteint un seuil prédéfini.
- [Revendication 3] Procédé selon la revendication 1 ou 2, caractérisé en ce que dans ladite troisième étape (40-80) ladite modification du réseau de neurones (RN) par apprentissage de type supervisé consiste à modifier une repré-

- sentation personnelle via une adaptation d'une structure interne.
- [Revendication 4] Procédé selon l'une des revendications 1 à 3, caractérisé en ce que dans ladite troisième étape (40-80) ladite modification du réseau de neurones (RN) par apprentissage de type renforcement consiste à fournir à ce dernier (RN) une tendance représentative de sa performance sur un trajet ou une manœuvre venant d'être effectué(e).
- [Revendication 5] Procédé selon l'une des revendications 1 à 4, caractérisé en ce que dans ladite deuxième étape (30) on estime lesdites première et seconde valeurs en fonction d'au moins un critère de performance choisi.
- [Revendication 6] Procédé selon la revendication 5, caractérisé en ce que chaque critère de performance est choisi parmi un confort de passagers dudit véhicule, une influence sur un temps de trajet ou de manœuvre dudit véhicule, un risque encouru par ledit véhicule, une stabilité des décisions et une utilisation d'accélération maximales.
- [Revendication 7] Produit programme d'ordinateur comprenant un jeu d'instructions qui, lorsqu'il est exécuté par des moyens de traitement, est propre à mettre en œuvre le procédé d'apprentissage selon l'une des revendications précédentes pour former un réseau de neurones (RN) chargé de déterminer des décisions de conduite pour un véhicule virtuel à conduite au moins partiellement autonome.
- [Revendication 8] Dispositif d'apprentissage (DA) pour former un réseau de neurones (RN) chargé de déterminer des décisions de conduite pour un véhicule virtuel à conduite au moins partiellement autonome, caractérisé en ce qu'il comprend au moins un processeur (PR) et au moins une mémoire (MD) agencés pour effectuer les opérations consistant à :
- déclencher le fonctionnement en parallèle, avec un même ensemble de données d'environnement définissant un environnement dudit véhicule et une définition d'un trajet ou d'une manœuvre, i) d'un simulateur de conduite (SC) pour qu'il détermine au moins une première décision de conduite pour ledit véhicule avec un ensemble de règles prédéfinies, et ii) dudit réseau de neurones (RN) pour qu'il détermine au moins une seconde décision de conduite pour ledit véhicule,
  - estimer des première et seconde valeurs représentatives respectivement de performances desdites première(s) et seconde(s) décisions de conduite,
  - modifier ledit réseau de neurones (RN) soit par apprentissage de type supervisé lorsque ladite première valeur est supérieure à ladite seconde valeur, soit par apprentissage de type renforcement lorsque ladite

première valeur est inférieure à ladite seconde valeur, puis  
- réitérer lesdits déclenchement, estimation et modification, avec à  
chaque fois un nouvel ensemble de données d'environnement et une  
nouvelle définition de trajet ou de manœuvre, tant qu'un critère d'arrêt  
n'est pas satisfait.

[Revendication 9] Ordinateur (OS) comprenant un simulateur de conduite (SC) destiné à  
déterminer, avec un ensemble de règles prédéfinies, des décisions de  
conduite pour un véhicule virtuel à conduite au moins partiellement  
autonome, caractérisé en ce qu'il comprend en outre un dispositif  
d'apprentissage (DA) selon la revendication 8, couplé à un réseau de  
neurones (RN) chargé de déterminer des décisions de conduite pour  
ledit véhicule.

[Fig. 1]

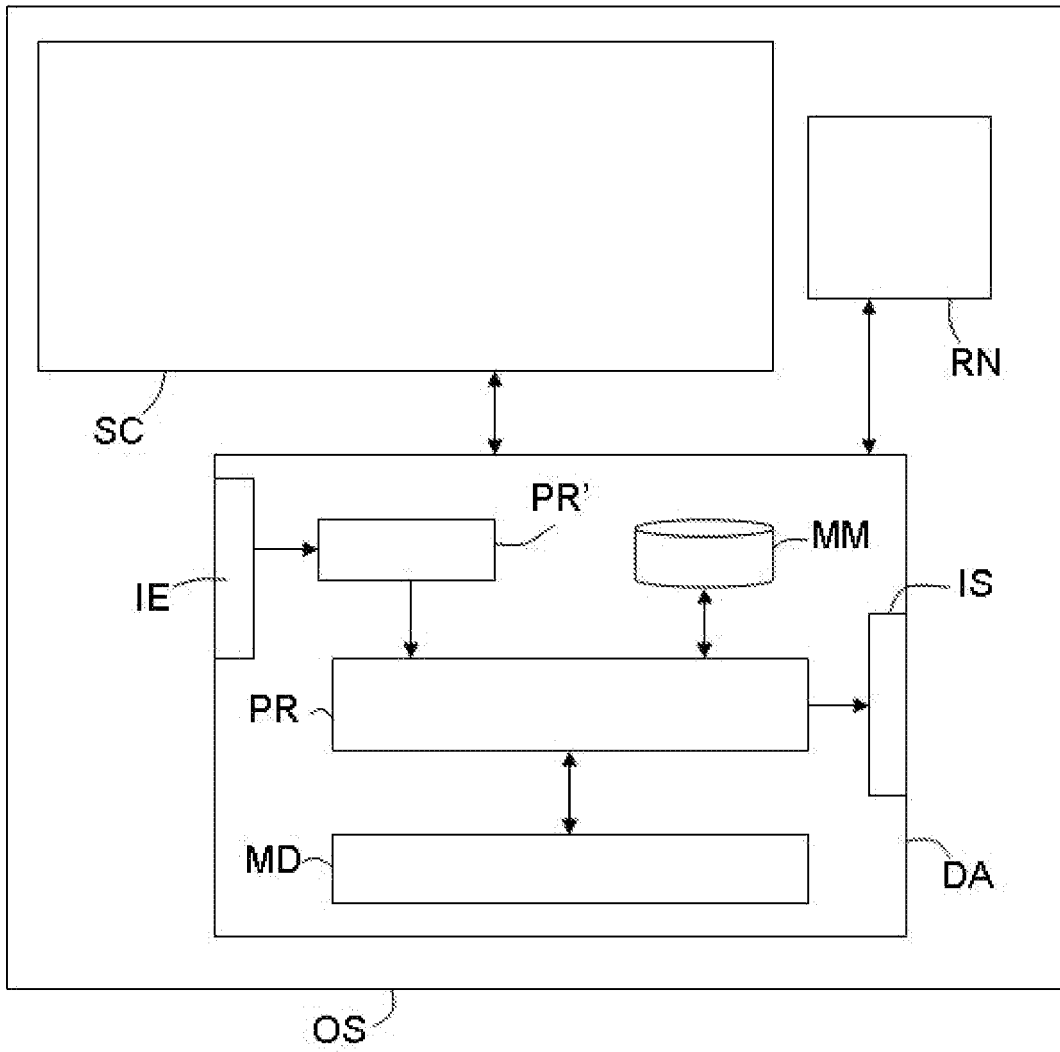


Fig.1

[Fig. 2]

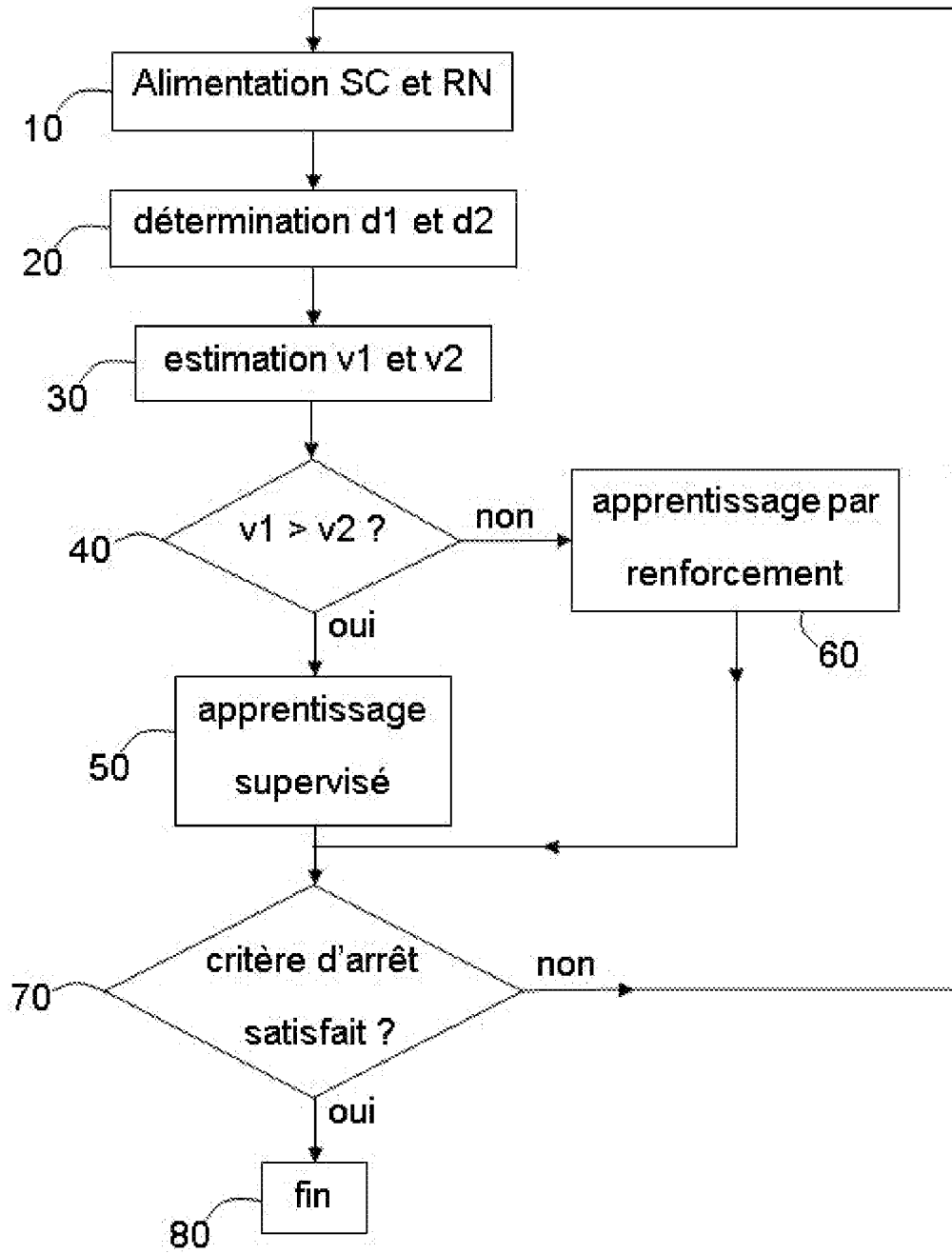


Fig.2

**RAPPORT DE RECHERCHE  
PRÉLIMINAIRE**

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

N° d'enregistrement  
national

FA 865784  
FR 1902086

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
X	<p>WANG BIN ET AL: "Design and implementation of an adaptive cruise control system based on supervised actor-critic learning", 2015 5TH INTERNATIONAL CONFERENCE ON INFORMATION SCIENCE AND TECHNOLOGY (ICIST), IEEE, 24 avril 2015 (2015-04-24), pages 243-248, XP033225634, DOI: 10.1109/ICIST.2015.7288976 [extrait le 2015-10-02] * abrégé * * chapitres I - III; page 1 - page 4; figure 1 *</p>	1-9	G06N3/08 B60W30/18
X	<p>DONGBIN ZHAO ET AL: "A supervised Actor-Critic approach for adaptive cruise control", SOFT COMPUTING, vol. 17, no. 11, 22 août 2013 (2013-08-22) , pages 2089-2099, XP055651888, DE ISSN: 1432-7643, DOI: 10.1007/s00500-013-1110-y * abrégé * * chapitres 1-3; page 2 - page 6; figures 1-3 *</p>	1-9	DOMAINES TECHNIQUES RECHERCHÉS (IPC) G06N
X	<p>Jeffrey A. Clouse ET AL: "A teaching method for reinforcement learning", 22 janvier 1992 (1992-01-22), XP055651799, Extrait de l'Internet: URL:https://web.cs.umass.edu/publication/docs/1992/UM-CS-1992-007.pdf [extrait le 2019-12-11] * abrégé * * chapitres 1-3; page 1 - page 4 *</p>	1-9	
Date d'achèvement de la recherche		Examineur	
12 décembre 2019		Hasnas, Sergiu	
CATÉGORIE DES DOCUMENTS CITÉS		T : théorie ou principe à la base de l'invention	
X : particulièrement pertinent à lui seul		E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure.	
Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie		D : cité dans la demande	
A : arrière-plan technologique		L : cité pour d'autres raisons	
O : divulgation non-écrite		.....	
P : document intercalaire		& : membre de la même famille, document correspondant	