

**Fusion de données multicapteurs :
application à la détection d'obstacle en temps réel.**

**Abdennasser CHEBIRA[°], Roger REYNAUD[°]
& Guy DEMOMENT^{°°}**

[°] IEF, bât 220, 91405 ORSAY Cédex

^{°°}LSS, ESE, Plateau de MOULON, 91192 GIF-sur -YVETTE Cédex

RÉSUMÉ

On présente dans cet article un algorithme de fusion de données qui construit une représentation d'une scène observée par deux capteurs identiques. L'algorithme, utilisé comme un algorithme de type Kalman, manipule récursivement une série d'objets et utilise des règles pour la remise à jour de cette représentation. Cet algorithme, basé sur la triangulation, présente trois originalités : — il limite la complexité en limitant la précision; — il coopère avec un système de décision, en conséquence, il doit fournir l'information au format du système expert et doit être contrôlable par ce dernier; — le temps d'exécution dépend des données et du suivi temporel. La décomposition du traitement en plusieurs phases, et particulièrement l'introduction d'une classification d'événements, permet d'obtenir un algorithme incrémental vis-à-vis du nombre d'événements à traiter.

ABSTRACT

We present in this paper a data fusion algorithm which build a representation of a scene observed by two identical sensors. The algorithm is used as a Kalman-like algorithm, it manipulates recursively two dedicated data bases and uses rules to update this representation. This algorithm is based on three original points: — as in fuzzy logic, decreasing complexity is done by decreasing precision; — the fusion algorithm can cooperate with an expert system and can be controlled by it; — the computation should be incremental according to the number of events to be processed.

1. INTRODUCTION

Nous nous intéressons ici à la reconstruction d'une image représentant l'environnement d'un véhicule équipé de deux capteurs [1] ayant une couverture angulaire de 120 degrés, et en absence de télémétrie qui permette une mesure directe de la distance. Les paramètres à estimer sont représentés sur la Figure 1.1.

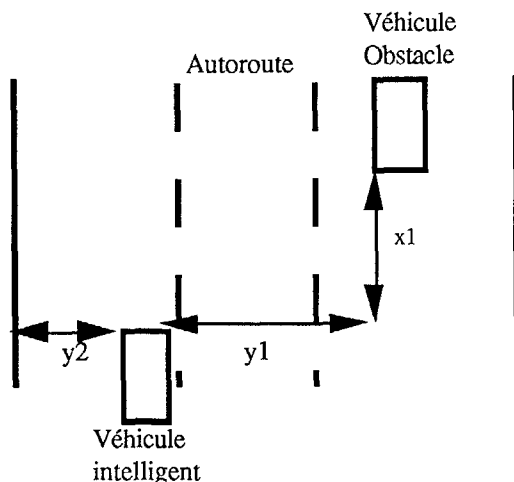


Figure 1.1 Paramètres définissant
la scène

imprécisions, difficultés d'apparier certaines données capteur, ce qui conduit à des artéfacts. Le but n'est pas d'aboutir à une estimation précise des distances, mais à une estimation qui permette une analyse correcte de la scène. Les résultats des estimations sont utilisés par un système expert [3] pour diagnostiquer des situations dangereuses et pour en informer le conducteur (assistance en conduite automobile).

L'originalité de notre méthode réside dans la façon dont nous nous sommes servi du suivi temporel des objets reconstruits, associé à une classification évolutive de ces objets, qui conditionne les traitements qu'ils vont subir.

Les contraintes qui pèsent sur la méthode sont nombreuses. Il faut pouvoir : (i) — oublier lentement pour être adaptatif ; (ii) — générer un nouvel obstacle d'autant plus vite qu'il est proche; (iii) — retarder la décision d'éliminer un artéfact pour améliorer les décisions ultérieures, ce qui amène à associer un degré de confiance à l'estimation de chaque obstacle; (iv) — et bien entendu réaliser le suivi des obstacles.

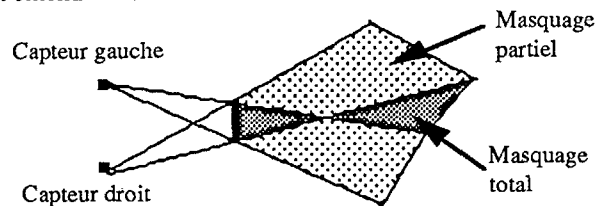


Figure 1.2 Les zones d'ombre

L'utilisation d'une paire de capteurs conduit à une reconstruction par triangulation [2] (stéréovision) dont les défauts sont bien connus (Figure 1.2) : zones d'ombre,

A la différence de [4] qui concatène tous les paramètres à identifier dans un vecteur d'état avec pour équation de transition d'état : "état à l'instant n+1 = état à l'instant n + bruit générateur", et qui utilise alors un filtre de Kalman classique,



nous avons décidé d'opter pour la solution suivante : une première phase de triangulation permet de reconstruire des obstacles potentiels, elle est suivie d'une deuxième phase de classification dont le but est d'éliminer les artéfacts de la phase précédente, en essayant d'associer à chaque obstacle potentiel nouvellement reconstruit un obstacle déjà identifié à l'itération temporelle précédente, une dernière phase de prise de décision vient alors compléter le traitement. Cette prise de décision s'insère dans une récurrence temporelle comme dans [5].

2. DESCRIPTION DE L'ALGORITHME

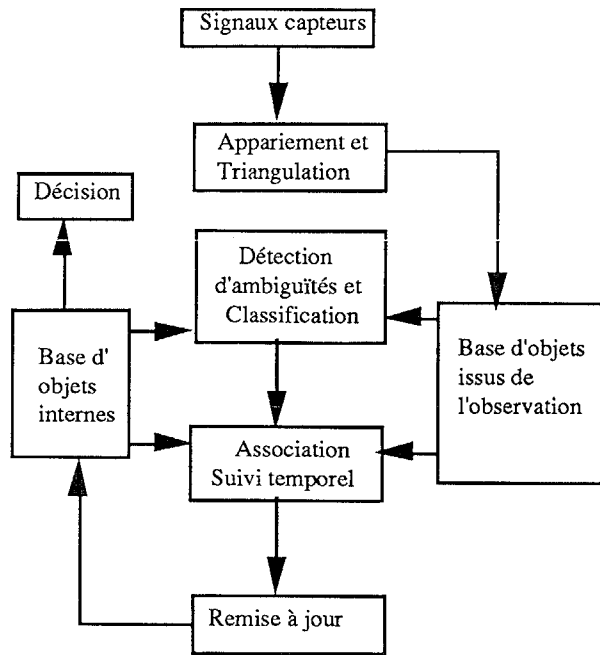


Figure 2.1 Algorithme général

Le traitement global effectué se décompose en une séquence de traitements successifs décrits dans la Figure 2.1. Une reconstruction par triangulation est insuffisante, et l'on est obligé d'utiliser un certain nombre d'informations géométriques a priori pour réduire le nombre de cas à prendre en compte, mais il est nécessaire aussi d'utiliser un suivi temporel en définissant à un instant t un état probable de la scène à reconstituer. Cet état est remis à jour dans un système bouclé à chaque récurrence temporelle comme le fait un filtre de Kalman en calculant une erreur entre un objet potentiel reconstruit par triangulation et l'état prédit d'un objet de la représentation interne.

Les traitements de fusion sont donc les suivants.

2.1 Appariement et triangulation

Cette phase consiste à appairer les données renvoyées par le capteur de gauche et le capteur de droite, et à calculer les positions en x et y d'obstacles potentiels. Les difficultés de la triangulation sont la présence de bruit sur les données, le manque d'information permettant d'effectuer correctement l'appariement entre les données fournies par le capteur «gauche» et celles fournies par le capteur «droit» et l'existence de zones d'ombre (Figure 1.1). Dans la zone de masquage total, aucun obstacle ne peut être détecté par les capteurs. Dans les zones de masquage partiel, les obstacles peuvent être partiellement observés, soit par le capteur gauche, soit par le capteur droit.

On utilise alors des connaissances a priori de nature géométrique qui permettent de minimiser le nombre de cas à traiter sur les positions reconstruites.

2.2 Classification

Des critères géométriques et le suivi temporel nous permettent d'attribuer aux paires de positions (x_i, y_i) fournies par la triangulation une description qualitative. Cette description consiste en un rattachement à l'une des quatre classes répertoriées. Cette opération est appelée classification, elle influe sur le calcul du degré de confiance que l'on accorde à l'objet ou à la paire (x, y) et rajoute une information de nature sémantique qui peut être utilisée par un niveau supérieur du traitement (prise de décision).

Nous avons répertorié quatre classes:

- classe «Bruit» : forte présomption de présence de bruit.
- classe «Ambiguë» : difficulté à partir de deux paires (x_i, y_i) et (x_j, y_j) d'y associer un obstacle.
- classe «Exacte» : reconstruction sans problème d'un obstacle.
- classe «Cachée» : non détection d'un objet déjà identifié dans la passé (objet caché)

Dans la phase de triangulation, on ne peut classer les objets que dans l'une des classes «Exacte» ou «Ambiguë». Pour arriver à détecter une reconstruction de type «Bruit» ou des objets «Cachés», il nous faut effectuer un suivi temporel. Ainsi en utilisant des lois de la physique telles que l'inertie, on peut classer un objet en classe «bruit» s'il apparaît brutalement et que l'on ne dispose pas de trace temporelle d'un objet se trouvant dans une zone voisine. On classe un objet en classe «cachée», s'il existe à l'instant t et disparaît à l'instant $t+1$.

2.3 Association et suivi temporel

Pour pouvoir effectuer la poursuite d'obstacles, on a choisi une représentation qui reflète l'environnement perçu par les capteurs binaires. Deux bases de données sont ainsi constituées. La première base \mathcal{O} , appelée base de données observation, est construite par la phase d'appariement/triangulation. Elle reflète donc la perception que l'on a de l'environnement à un instant t donné. La seconde \mathcal{I} , appelée base de données de la représentation interne, est l'image que l'on a de l'environnement à des instants précédant l'instant t . Elle représente donc une mémoire à horizon fini pour le système. Ces bases de données sont constituées d'objets identiques composés des champs suivants :

$I.x$	position longitudinale
$I.y$	position transversale
$I.classe$	classe
$I.Vx$	vitesse longitudinale
$I.Vy$	vitesse transversale
$I.p$	vraisemblance que l'objet soit l'image d'un obstacle réel présent dans la scène, ce qui traduit le degré de confiance que l'on accorde à l'existence de cet objet.

Soit \mathcal{I} la base de données a priori à l'instant t (base de connaissance initiale avant prise en compte des nouvelles observations). Soit \mathcal{O} la base de données reconstruite à partir des observations. On cherche pour chaque objet $O \in \mathcal{O}$, l'objet $I \in \mathcal{I}$ que l'on pourra lui associer de façon la plus probable. Une façon naturelle de faire serait de prédire la position de l'objet I , puis de calculer les distances entre la position de l'objet O et celles prédites de tous les objets I . Il y aurait association entre l'objet O et tous les objets I suffisamment proches de O .

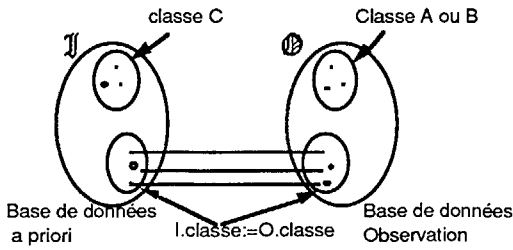


Figure 2.1 Association et Classification

L'association est faite par l'intermédiaire d'une fonction F . La fonction F est définie de $\mathbb{I} \times \mathbb{O}$ dans $[0,1]$ par :

$$F(I,O) = 2 \text{Sint} / (\text{SI} + \text{SO})$$

où SI et SO sont les surfaces respectivement des régions RI et RO où sont localisés les objets I et O , et Sint est la surface de la région intersection des deux régions précédentes (Figure 2.2).

Du fait de la faible résolution des capteurs utilisés, la précision sur les positions estimées devient insuffisante pour pouvoir estimer correctement les vitesses relatives des objets. Il n'est donc plus question de pouvoir prédire les nouvelles positions à l'instant $t+\Delta t$ d'un objet I , et cette prédiction est remplacée par un rectangle d'incertitude suffisamment grand pour prendre en compte l'incertitude des positions, mais aussi l'absence de modélisation de la dynamique.

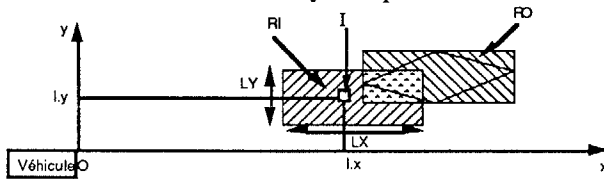


Figure 2.2 Localisation des objets et fonction Association

La scène est découpée en zones trapézoïdales par le maillage imposé par le balayage angulaire avec une résolution de 1 degré et par la méthode de reconstruction. Plus on s'éloigne des capteurs, plus l'incertitude sur la localisation des obstacles augmente. La surface de la maille en forme de trapèze s'agrandit (Figure 2.3).

La région RO est un rectangle qui englobe le trapèze d'incertitude associé aux coordonnées de l'objet reconstruit. Le fait de s'étendre au delà de la surface du trapèze est nécessaire si l'on veut pouvoir suivre un objet qui change de maille.

La région RI est aussi rectangulaire, elle délimite la zone où l'on pense devoir localiser un obstacle probable de la représentation interne à l'instant $t+1$. Elle est centrée autour de la position de l'objet interne à l'instant t car nous ne sommes pas capable de modéliser la dynamique de la scène de façon exacte et donc de prédire la nouvelle position de l'objet.

La longueur des côtés de RI est fonction des positions relatives de l'objet I par rapport au véhicule intelligent. La largeur et la longueur de ce rectangle sont des fonctions positives, strictement croissantes soit de x , soit de y .

La fonction F autorise l'association entre objet O et objet I si sa valeur est supérieure à un seuil. Ce seuil est obligatoirement faible si l'on veut autoriser l'association et donc le suivi d'un objet qui quitterait une maille du maillage imposé par la triangulation pour sauter dans une maille adjacente. Ce seuil devrait être encore plus faible et la région RO plus grande si l'on désire suivre un saut vers une maille non adjacente.

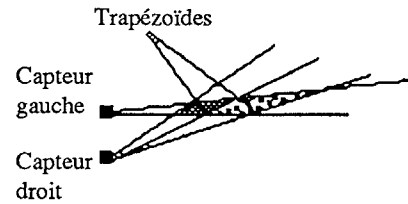


Figure 2.3 Résolutions des capteurs et maillage du plan

2.4 Mise à jour de la base de données a priori

La remise à jour de la base de données a priori se fait de la manière suivante.

2.4.1 Création d'un objet

En cas d'apparition d'un nouvel obstacle dans le champ de vision des capteurs, la phase de triangulation fournit une position en x (p_x), une position en y (p_y), et une classe c . Un nouvel objet I est alors créé et a pour attributs :

- position longitudinale $I.x = p_x$
- position transversale $I.y = p_y$
- classe $I.classe = \text{correct ou ambiguë}$
- vitesse longitudinale $I.V_x = 0$
- vitesse transversale $I.V_y = 0$
- Probabilité que I soit l'image d'un obstacle réel présent dans la scène $I.p = 0.5$

2.4.2 Poursuite d'obstacles en présence d'association

Les attributs des objets I qui ont été associés à des objets O sont remis à jour de la manière suivante:

- $I.x = O.x$
- $I.y = O.y$
- $I.classe = O.classe$
- estimation de vitesses transversale et longitudinale
- $I.\text{prob}_{t+1} = \lambda(\text{classe})(I.\text{prob}_t + (1 - I.\text{prob}_t) C(I,O))$
- où

λ : facteur d'oubli inférieur à un, il peut être fonction de la classe

$I.\text{prob}_t$: probabilité d'existence de l'objet I à l'instant t . Il s'agit d'un terme cumulatif interprété ici en terme de probabilité Il est donc borné à 1. On désire oublier les effets antérieurs, ce qui est le rôle du facteur d'oubli sur ce terme. Cet oubli est compensé par une génération proportionnelle à une fonction $C(I,O)$ qui représente la qualité de l'association des objets I et O . Ce coût est une quantité comprise entre 0 et 1, car on veut donner à $I.\text{prob}$ une interprétation de probabilité avec donc valeur sur $[0,1]$. Ce coût donne une information sur la qualité de l'association effectuée (probabilité d'avoir effectué la bonne association). On affectera donc la valeur 1 à une association de bonne qualité, zéro sinon.

Remarque : on aurait pu prendre comme coût la fonction $F(I,O)$, en effet les valeurs de cette fonction sont comprises entre 0 et 1. La définition d'une fonction coût différente de F est guidée par les objectifs suivants :

- moduler la fonction $F(I,O)$,
- permettre un choix plus vaste et donner une interprétation plus forte à cette notion. Dans la prochaine version de cet algorithme, cette fonction aura comme attributs supplémentaires les classes des objets I et O . Il paraît naturel de donner des coûts différents aux associations selon que les classes des objets soient «Correct», «Ambiguë»,... Une valeur nulle permettra un oubli sans compensation (classe = «bruit»).



2.4.3 Suivi d'un objet en absence d'association avec des observations

Les attributs des objets I n'ayant pas été associés à des objets O sont remis à jour de la manière suivante :

$$\begin{aligned} I.x_{t+1} &= I.x_t + I.V_x \text{ DT} \\ I.y_{t+1} &= I.y_t + I.V_y \text{ DT} \\ I.V_{xt+1} &= I.V_{xt} \\ I.V_{yt+1} &= I.V_{yt} \\ I.\text{probt}+1 &= l(\text{classe})I.\text{probt} \\ I.\text{classe} &= \text{«Caché» ou «Bruit»} \end{aligned}$$

Ce cas nous permet de suivre un certain temps des objets masqués ou des artefacts, le facteur d'oubli λ et le seuil de destruction conditionnent la durée de vie de tels objets. A ce stade là on effectue une deuxième classification. On attribue aux objets I n'ayant pas été associés la classe Cachée ou Bruit, un objet de classe «ambiguë» ou «exacte» est créé pour chaque objet n'ayant pas d'associé (Figure 2.1).

2.4.4 Destruction d'un objet de la base de données a priori

Tous les objets $I \in \mathbb{J}$ dont la probabilité d'existence est inférieure à un seuil pré-défini sont *enlevés* de la base \mathbb{J} (élimination des obstacles sortis du champ de vision des capteurs, filtrage).

Remarque : la remise à jour des attributs dépend de la qualité de la mise en correspondance ($C(I,O)$), de la confiance accordée à l'événement I (classe et probabilité d'existence). De cette manière on arrive à faire apparaître les ambiguïtés, à éliminer les objets dus au bruit (filtrage) et à tenir compte de l'inertie des obstacles présents dans la scène (persistance, suivi temporel).

Les contraintes temps réel de traitement en ligne imposent une sélection des objets I les plus intéressants de l'ensemble \mathbb{J} . Cette phase est incrémentale vis à vis du nombre de cas à traiter. En effet la durée disponible avant l'arrivée des échantillons suivants est variable, et on doit se doter de critères pour décider quels premiers cas traiter. Deux critères sont utilisés, le premier est la proximité (synonyme de danger), le second l'appartenance à la classe Correcte.

3. SIMULATIONS ET RESULTATS

Plusieurs scénarios ont été simulés. Trois véhicules (obstacles) ayant des trajectoires prédéfinies sont placés dans le champs de vision des capteurs. L'algorithme arrive toujours à détecter et estimer les paramètres du véhicule le plus proche même si parfois l'objet correspondant appartient à la classe «Ambiguë». Les hyperparamètres ont été fixés lors des simulations, les valeurs suivantes ont été choisies :

facteur d'oubli	Classe
0.9	«Correcte»
0.8	«Ambiguë»
0.7	«Cachée»

— seuil de destruction 0.4

— seuil d'association 0.4

Ces valeurs permettent de suivre un obstacle caché après trois périodes d'échantillonnage, de rejeter un artefact après 2 périodes et de faire tendre vers le maximum la probabilité associée à un objet de classe «Correcte» après 2 périodes.

4. CONCLUSION ET PERSPECTIVES

En utilisant la technique de triangulation, on a pu développer un algorithme qui permet d'estimer les positions d'obstacles potentiels présents dans le champs de vision de capteurs. Ces positions sont utilisées par un système expert temps réel pour remplir la fonction d'assistance à la conduite automobile.

La classification et le suivi temporel ont permis d'une part de lever des ambiguïtés et d'améliorer les qualités d'estimation par rapport au cas statique, et d'autre part de rajouter des informations (Classe et probabilité) qui peuvent être utilisées par un système expert pour améliorer les qualités de poursuite de l'algorithme.

Les différentes techniques utilisées en fusion de données peuvent être trouvées dans [6].

La durée du traitement est faible, découlant de l'utilisation de capteur à faible débit par rapport à des caméras CCD. La précision sur les positions estimées est suffisante pour notre application. La durée de traitement de chaque récursion n'est pas fixe, elle croît avec le nombre d'événements présents sur les deux signaux de balayage angulaire. Pour conserver un traitement en ligne, cette durée doit être limitée en imposant des contraintes supplémentaires, nous avons choisi la classification pour introduire un traitement orienté objectif et non une simple limitation du nombre de cas. De plus, l'objectif peut être spécifié par le système expert.

Dans le futur, l'interface entre le système expert et le module de fusion sera prépondérant. Un protocole de coopération aura pour objectif le contrôle de la partie fusion par l'introduction d'un boucle supplémentaire. Ce contrôle doit être dynamique, dépendant de l'environnement incertain dans lequel évolue le véhicule et interviendra à différents niveaux du traitement (au niveau de la remise à jour, de l'association, de la détection d'ambiguïtés, de la classification ou de la triangulation). Cet objectif a conditionné l'organisation de l'algorithme.

5 BIBLIOGRAPHIE

- [1] D. Estève, M. Alaoui, R. Gillet, T. Maurin, R. Reynaud & M. Dang. " I.R Detection sensor for obstacle avoidance" Fourth Workshop PROMETHEUS PROCHIP, Florence, December 3-4 1990, pp 112-114.
- [2] B. Thiesse, P. Marthon, M. Bellamine, P.Darcos & A. Bruel "Etude comparée des différentes approches algorithmiques de la vision binoculaire", COGNITIVA, Paris, 4-7 juin 1985, pp 151-158.
- [3] D. Berschandy, T. Maurin, F. Poirier, R. Reynaud & J.C. Heudin, "On-board real time expert system : Hardware and Software for danger detection demonstrator", Third Workshop PROMETHEUS PROCHIP, Orsay, MAY 14 ,1990, pp 301-307.
- [4] O. D. Faugeras, N. Ayache & F. Lustman, "Depth and motion analysis: the machine developed within esprit project 940" , IAPR Workshop on CV -Special Hardware and Industrial Applications, Tokyo, oct. 12-14, 1988.
- [5] J. Dézert, "Vers un nouveau concept de navigation autonome d'engin. Un lien entre le filtrage à association probabiliste de données et la théorie de l'évidence", Thèse de l'Université Paris XI, Sept 1990.
- [6] E. Waltz & J. Llinas, "Multisensor data fusion", Artech House, 1990.