

Conception d'un capteur de vision active basé sur un imageur CMOS

Use of a CMOS imager to design an active vision sensor

P. Chalimbaud

F. Berry

LABoratoire des Sciences et Matériaux pour l'Electronique et d'Automatique
GRoupe Automatique et VIsion pour la Robotique
24 avenue des landais
63177 Aubiere cedex - FRANCE
chalimba,berry@lasmea.univ-bpclermont.fr

Résumé

Le but de cet article est de démontrer l'intérêt d'un capteur de vision basé sur les principes de vision active. Afin d'intégrer les concepts de vision active, ce capteur exploite les technologies FPGA et les nouveaux capteurs de vision CMOS. L'objectif d'un tel système est de rétroagir sur lui même de façon afin d'optimiser continuellement ses paramètres de perception. Cette capacité permet, d'une part d'améliorer la perception, mais aussi de se focaliser sur les régions dites "intéressantes". Ainsi, le goulot d'étranglement entre la caméra et le système de traitement, qui est un problème recurrent au sein des systèmes de vision, peut être évité.

La notion de *détecteur actif* est proposée afin de formaliser notre approche. En guise d'illustration, un suivi rapide de zone texturée est proposé. Cette méthode est basée sur la corrélation d'un motif de référence avec une région d'intérêt centrée sur la cible. Suivant les dimensions du motif considéré, la fréquence de fonctionnement du dispositif de suivi varie entre 200 et 2000 acquisitions par seconde.

Mots Clef

Capteur de vision, imageur CMOS, vision active et suivi de motif.

Abstract

The goal of this paper is to show the interest of a vision sensor based on active vision principles. In order to integrate active vision concepts, this sensor uses FPGA technologies and new CMOS vision sensors. Objective of a such system is to update its perception parameters, in order to improve its performances. This capacity allows to improve the perception and to focalise the system on areas of interest. Thus, the bottleneck between camera and the processing unit can be avoided.

Then, *active detector* notion is proposed in order to formalized our work. To illustrate our approach, a high speed

tracking method of a grey level template is proposed. This method is based on a correlation between a reference template and an area of interest centered on the target. According to the size of the area of interest, the acquisition rate varies from 200 fr/s to 2000 fr/s.

Keywords

Vision sensor, CMOS imager, active vision and template tracking.

1 Introduction

Le but des recherches en vision artificielle est d'exploiter des images ou des séquences d'images, générées par des capteurs, afin d'interpréter un environnement. A partir de cette interprétation, différentes méthodes permettent d'identifier ou d'inspecter des objets, guider des robots, etc... Le problème consiste à identifier l'ensemble des objets présents dans la scène. Le processus d'interprétation utilise les informations fournies par le capteur et généralement, un modèle de l'environnement ou des modèles d'objets.

La première manière d'aborder un problème de vision est l'approche par vision passive ou par vision dynamique. La vision passive est la méthode classique d'analyse des images. Elle consiste à étudier les images d'une scène suivant un ou plusieurs angles d'observation en supposant un environnement statique. Par opposition, la vision dynamique correspond à l'étude des informations visuelles, mais dans une séquence d'images non-finie. Le temps est introduit dans le système, donc la notion de mouvement apparaît dans le processus de perception, ce que l'on caractérise par le flot optique. Le flot optique peut être défini comme un ensemble de vecteurs qui modélise la direction et l'intensité du mouvement dans l'image pendant un intervalle de temps. Il existe un grand nombre de méthodes qui permettent d'extraire le flot optique : les méthodes différentielles, des méthodes basées sur la

mise en correspondance, des méthodes basées sur la théorie des ondelettes [5, 7]. Bien que la notion de mouvement enrichisse grandement l'approche passive, on peut reprocher à la vision passive ou dynamique de n'utiliser le capteur que de manière statique¹. En effet, jusqu'ici la chaîne d'acquisition est exploitée de façon non contrôlée. A partir de cette observation, Aloimonos [1], Bajcsy [3] et Ballard [4] ont été les premiers à définir une nouvelle approche du problème de vision nommée : la vision active. Les techniques de vision active sont issues d'une tentative d'imitation du dispositif de vision humain. La vision humaine utilise les mouvements de la tête et des yeux dans le mécanisme de fixation, la sensibilité des yeux s'adapte aux variations de luminosité, etc... Par conséquent, le but de la vision active est de reproduire ces facultés d'adaptation. Les points faibles qui limitent les performances des systèmes de vision sont généralement leur sensibilité aux bruits, leur faible précision due à un manque de réactivité. L'un des buts de la vision active est d'élaborer des stratégies pour adapter les paramètres de la ou des caméras (position, vitesse, zoom, ouverture ...) afin d'améliorer la perception. Cette réactivité provient du retour de l'information visuelle sur les paramètres de la caméra et les paramètres de traitement de l'information. Notre approche est basée sur cette notion de vision active. Nous désirons développer un capteur qui intègre un dispositif actif. L'objectif est de permettre au capteur d'adapter sa conduite en fonction des informations visuelles détectées. On introduit ainsi une rétroaction sur la capteur conditionnée par la tâche à réaliser et les informations visuelles.

La suite de cet article est organisée de la façon suivante. Tout d'abord, nous proposons une brève revue des systèmes de vision active récents. Ensuite, nous justifions notre intérêt pour les imageurs CMOS en comparant les deux principales technologies d'imagerie. Puis, nous développons notre approche basée sur l'association d'un imageur CMOS avec le concept de vision active. Nous proposons, alors, une méthode de suivi rapide dont les performances illustrent l'intérêt de notre approche. En conclusion, nous proposons diverses améliorations de notre algorithme de suivi, et nous présentons nos directions de recherche futures.

2 Vision active

Plusieurs systèmes basés sur le paradigme de vision active ont été développés. Ils ont contribué au développement de nouvelles idées qui ont permis de mettre au point des techniques très efficaces pour résoudre des problèmes de vision variés. Mais, d'après Vieville [17], la vision active ne doit pas simplement se limiter à piloter l'orientation d'une caméra. La vision active peut être définie comme un contrôle actif de tous les paramètres du capteur afin de réaliser une tâche donnée. Cette définition montre la place

importante qu'occupe le capteur dans un système de vision active.

K. Pahlavan et *al.*, dans [15], divisent les composants d'un système de vision active en quatre catégories :

- les paramètres optiques qui définissent la projection d'une scène 3D en une image 2D (zoom, éclairage...),
- les paramètres sensoriels qui conditionnent le signal électrique porteur de l'information image 2D (sensibilité, échantillonnage...),
- les paramètres mécaniques qui déterminent le positionnement et les mouvements de la caméra,
- les paramètres algorithmiques qui caractérisent le traitement et les transformations de l'information.

L'ensemble de ces paramètres définissent le trajet de l'information visuelle de son état brut vers une représentation adaptée aux mécanismes d'interprétation.

Un grand nombre de travaux basés sur la vision active utilisent une ou plusieurs caméras fixées sur une plate-forme robotique dont on contrôle l'orientation panoramique et l'inclinaison. Une revue sur des systèmes de vision basés sur des têtes robotiques est d'ailleurs proposée dans la thèse de Andersen intitulée "*A Framework for control of a camera head*" [2]. Ces dispositifs mettent en œuvre différents types de fonction tels que des processus d'attention, de contrôle du point d'observation ou encore des fonctions élémentaires de robotique tel que le suivi. Les lecteurs intéressés peuvent également consulter le livre de Viéville intitulé "*A few steps towards 3D active vision*" [17] qui comporte une analyse approfondie des systèmes stéréoscopiques munis d'une orientation panoramique et d'une inclinaison contrôlable. Plus récemment, J. Batista [6] a proposé un système de suivi binoculaire exploitant le flot optique. Ces travaux tirent profit du contrôle des paramètres mécaniques, optiques et de l'intégration algorithmique. Les paramètres sensoriels ne sont pas exploités par le dispositif de contrôle du système de vision. Ceci est dû à l'utilisation de capteurs basés sur une technologie CCD qui ne permet pas d'asservir le conditionnement du signal électrique porteur de l'information image. En d'autres termes, ces paramètres ne peuvent être introduits dans la boucle de perception. Pour cette raison et afin d'exploiter ces paramètres, nous nous sommes intéressés à une autre technologie d'imageur pour développer notre approche.

3 Approche proposée

3.1 Introduction

Dans le chapitre précédent, les travaux présentés utilisent des têtes robotiques monoculaires ou stéréoscopiques afin de contrôler la position, l'orientation et les paramètres optiques de la caméra. Ces systèmes sont intéressants pour réaliser des tâches de reconstruction, de reconnaissance ou encore de suivi. Mais, le lien étroit entre la perception et l'action proposée dans le concept de vision active ne se

¹statique : dans le sens où le capteur n'est pas contrôlé.

limite pas à la conduite d'une caméra dans l'espace. Le traitement de l'information est lui aussi directement lié aux activités visées par le système (navigation, manipulation, signalisation de danger et d'opportunité, ...). L'adaptation de la partie algorithmique du mécanisme de vision peut engendrer une simplification des méthodes de contrôle et de représentation de la scène. De ce fait, le temps de réponse du processus de vision peut alors être minimisé. De plus, l'adaptation ou l'ajustement du système aux caractéristiques de la scène et aux objectifs de la tâche à réaliser, augmente sa fiabilité.

Des études récentes [12, 8, 11] ont été menées sur l'asservissement du dispositif de vision selon le concept de vision active sans considérer d'action mécanique sur le système. Dans ces travaux, l'objectif est d'intégrer le retour de l'information visuelle dans le capteur. L'essentiel du problème consiste à sélectionner des régions d'intérêt² dans l'image afin de concentrer les ressources de calcul sur l'information utile. De ce fait, la notion d'étude locale devient alors prédominante.

En ce qui concerne notre approche, nous proposons d'inclure dans la boucle de vision les paramètres sensoriels en introduisant la notion de *détecteur actif*. Un détecteur actif est une entité autonome constituée d'un ensemble de traitements associé à une zone d'intérêt sur l'imageur. La fonction de cette entité consiste à gérer et à faire évoluer, dans le temps, deux types de données :

- les caractéristiques de la zone d'intérêt sur l'imageur,
- un ensemble de traitements associé à cette région de l'image.

L'objectif de ce détecteur est d'extraire une information locale et de déceler continuellement son évolution durant la perception. Par principe, la définition d'un détecteur actif est fixée par la tâche qui lui est confiée. Celle-ci doit conditionner d'une part les traitements locaux à effectuer, mais aussi les caractéristiques géométriques de sa fenêtre d'acquisition. Cette notion de détecteur actif, précédemment introduite dans [10] peut être considérée comme une macro-fonction visuelle destinée à réaliser une tâche locale. Une notion similaire a été proposée par S. Ullman en 1984 avec la notion de routine visuelle. Le concept de routine visuelle a été introduit pour la première fois par S.M. Kosslyn en 1973, mais la formalisation qui a permis sa mise œuvre, a été définie par S. Ullman [16]. L'idée principale est d'établir une collection de routines qui permettent de réaliser une large gamme d'opérations élémentaires de traitements d'images. L'objectif de cette approche est de faciliter l'élaboration de tâches spécifiques.

La différence majeure avec nos travaux réside dans le fait qu'un détecteur actif associe un élément matériel (le capteur) et un élément logiciel (contrôle et algorithmes)

²Il est important de souligner la différence entre le suivi d'une zone d'intérêt dans l'image et le suivi réalisé à partir d'une caméra mobile. Dans le premier cas, cela revient à un mouvement dans l'espace image, tandis que dans le second cas, on considère alors le mouvement dans l'espace opérationnel.

alors que les routines visuelles manipulent seulement des éléments logiciels. De ce fait, il est essentiel de noter l'importance de l'adéquation entre le concept de détecteur actif et l'architecture matérielle. Comme cela est expliqué ci-dessous, l'exploitation de la partie matérielle, qui particularise le concept, sera facilitée par les caractéristiques des imageurs CMOS.

3.2 Architecture du capteur

Généralement, une image est mesurée par un capteur, convertie, stockée (carte d'acquisition) et traitée par un ordinateur. Dans notre approche nous proposons de réaliser l'ensemble des traitements de vision précoce à l'aide de détecteurs actifs directement intégrés dans le capteur. Ce type de capteur, intégrant des pré-traitements, voir de l'"intelligence" est souvent appelé "smart sensor". Traditionnellement, un tel capteur est constitué d'une caméra et d'une unité de traitement embarquée (fig. 1).

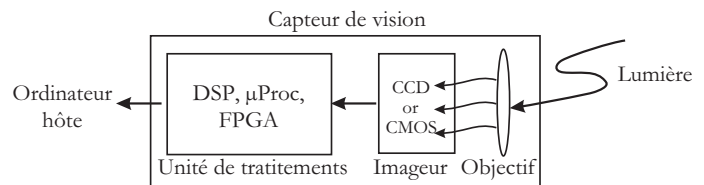


FIG. 1 – Capteur de vision embarqué traditionnel

Dans de telles architectures, il existe de nombreuses solutions concernant le choix l'imageur et le choix de l'unité de traitement. Dans notre cas, il est évident qu'une adéquation optimale entre le capteur et les détecteurs actifs est nécessaire. De plus, il est important de considérer chaque détecteur actif comme une boucle d'asservissement interne au capteur. Au sein de cet asservissement l'image représente la mesure et le système à contrôler est le capteur. Ainsi pour un problème de vision donné, le détecteur actif doit asservir le capteur afin d'effectuer la tâche. Dans ce contexte, l'architecture du capteur peut être considérée comme un ensemble de boucles d'asservissement fonctionnant en parallèle. Ce parallélisme va déporter le goulot d'étranglement du système au niveau de l'imageur. Ainsi la rapidité du dispositif sera conditionné par les performances de l'imageur. D'un point de vue technologique, notre capteur (fig. 3) peut être décomposé en trois modules :

- Le module imageur constitué par imageur CMOS, un convertisseur analogique numérique et un dispositif optique.
- Le module de traitement, qui s'avère être le cœur de notre système, est constitué d'un FPGA Stratix EP1S60 de la société Altera et de différents blocs mémoires. A l'heure actuelle nous disposons de 64 Mo de SDRAM et de 5×1 Mo de SRAM ayant chacune des bus d'adresses et de données privés. Ce module offre la possibilité de connecter un module optionnel tel qu'un

DSP pour effectuer des calculs particuliers.

- Le module de Communication permet de connecter le capteur à une station hôte (fig. 2) à l'aide du protocole USB2.0. Ce module permet d'interfacer un accéléromètre, un contrôleur de zoom, un contrôleur de moteur dans le cas d'un embarquement sur une tourelle pan/tilt.

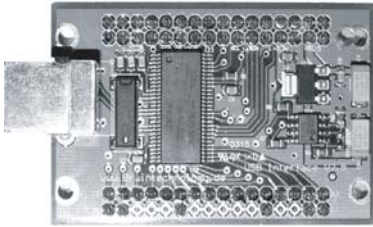


FIG. 2 – Carte de communication USB2.0



FIG. 3 – Capteur de vision embarqué

Dans le paragraphe suivant nos choix technologiques sont développés et comparés à d'autres solutions.

3.3 Les imageurs CMOS

Actuellement, les nouveaux imageurs CMOS-APS³ et les imageurs CCD⁴ présentent des performances de plus en plus similaires. Mais les imageurs de technologie CMOS présentent une propriété supplémentaire intéressante. La structure de ces capteurs (fig. fig. 5) permet d'accéder aux valeurs des pixels, de la matrice photo-sensible, de manière aléatoire.

Les capteurs d'image CCD et CMOS sont deux technologies très différentes pour capturer les images. Bien qu'elles soient souvent considérées comme deux technologies concurrentes ; les imageurs CMOS et CCD possèdent chacun leurs avantages et inconvénients qui font d'eux des capteurs adaptés à des applications différentes. Les principaux avantages des capteurs CMOS [9] sont :

³Complementary metal Oxide Semiconductor - Active Pixel Sensor

⁴Charge-Coupled Device

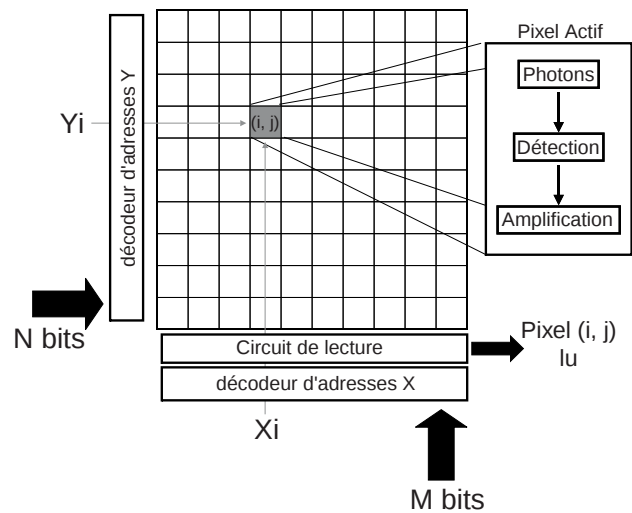


FIG. 5 – Architecture élémentaire d'un imageur CMOS

- leur large gamme de dynamique,
- l'accès aléatoire aux pixels,
- la maturité de la technologie d'intégration CMOS,
- la faible sensibilité au "Blooming",
- l'absence de "Smearing".

Le deuxième de ces avantages, nous intéresse tout particulièrement. En effet, l'accès aléatoire permet de contrôler et sélectionner les pixels que l'on désire exploiter. Ainsi, contrairement aux imageurs CCD, il est possible d'acquérir uniquement une faible quantité de pixels dans le but d'obtenir des vitesses d'acquisition image élevées. Ceci peut expliquer l'intérêt actuelle de la communauté, travaillant sur les systèmes de vision artificielle, pour cette technologie. Il est important de noter que la cadence d'acquisition des caméras CCD est exprimée en images par secondes tandis qu'en ce qui concerne les caméras CMOS, il est plus représentatif d'exprimer la cadence d'acquisition en pixels par secondes. Bien que les capteurs d'images CMOS soient, maintenant, très répandus dans le commerce, peu d'entre eux permettent un accès au pixel réellement aléatoire dans l'espace et le temps. En d'autres termes, il n'est pas possible de lire, à n'importe quel moment n'importe quel pixel de la matrice photo-sensible. De plus, la plupart des dispositifs d'adressage intégrés dans les capteurs commerciaux, induisent des contraintes supplémentaires. Ces contraintes proviennent de leur architecture interne qui est étudiée pour minimiser le nombre d'entrées du circuit. L'accès aléatoire aux pixels devrait mener à la mise en œuvre de solutions efficaces pour des applications variées, lorsque la cadence d'acquisition peut être augmentée aux dépens de la taille et de la résolution de l'image. Un autre avantage des imageurs CMOS est la capacité à contrôler la dynamique de l'image. Dans la plupart des imageurs CMOS le signal image est directement converti en numérique dans le circuit. Il

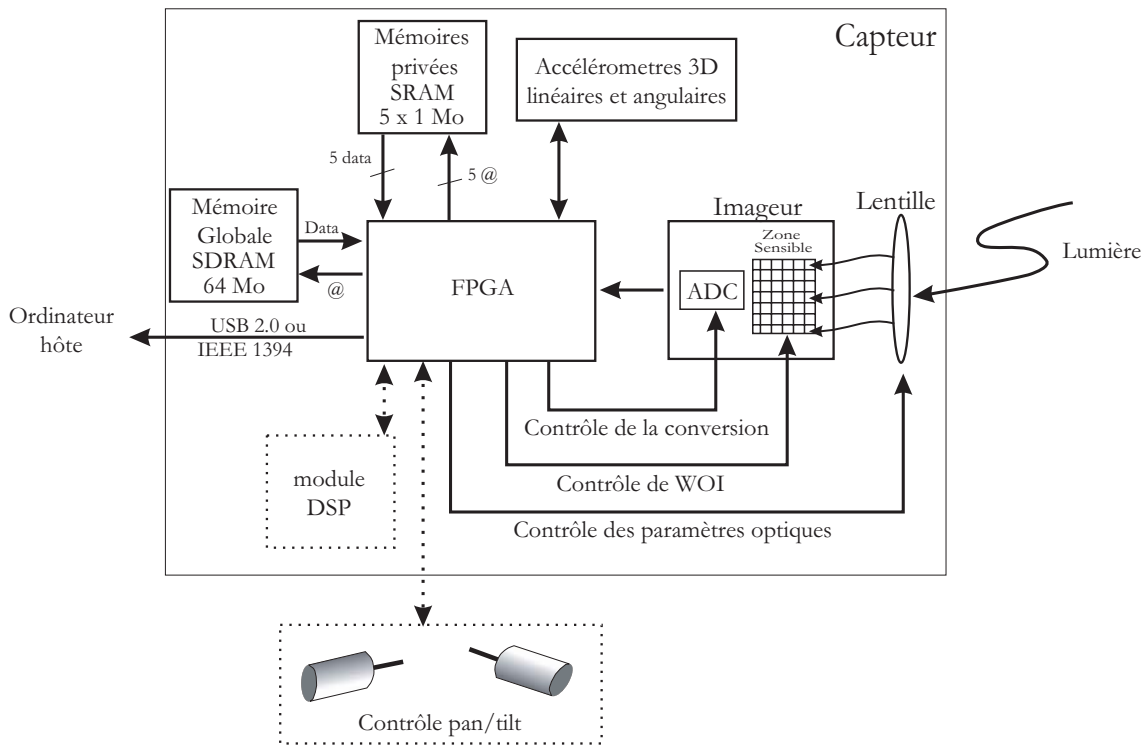


FIG. 4 – Architecture globale du capteur

est possible de régler dynamiquement les paramètres du convertisseur analogique numérique. Ceci est très utile pour considérer des régions de l'image dont la luminosité varie de manière spatio-temporelle.

3.4 FPGA

Les récentes avancées dans le domaine de la logique programmable ont conduit à la réalisation de nombreuses plate-formes de prototypage. En effet parmi ses nombreux avantages, les FPGAs permettent un développement à bas coût et un temps de développement très court et présentent une grande facilité de reconfiguration par l'utilisateur. De tels composants sont constitués d'un ensemble de blocs logiques (LUT, flip-flop), d'une RAM conséquente et d'un vaste système d'interconnection. Ces avantages sont un facteur déterminant pour le design de notre capteur et sont particulièrement bien adaptés au concept de détecteurs actifs. Dans notre cas le composant choisi (Stratix EP1S60) est basé sur une technologie 1,5v, 0,13 μ m et offre 57120 éléments logiques avec 122 broches d'entrées sorties. Ce composant est optimisé pour l'implantation de systèmes programmables sur puce (System On Programmable Chip) en intégrant la notion de cœur de processeur logiciel. Pour ce type de FPGA, Altera propose un outil de développement (NIOS) qui permet d'instancier un processeur RISC avec 5 niveaux de pipeline cadencé à 80 MHz avec une largeur de bus de 16 ou 32 bits. Le terme "cœur de processeur logiciel" correspond à une descrip-

tion matérielle (en VHDL ou Verilog) qui est synthétisée dans le FPGA. Il est à noter la possibilité d'accélérer certains traitements, en ajoutant des instructions personnelles (décrites en VHDL) au cœur NIOS. De cette manière, il est alors possible de réaliser de la surcharge d'opérateurs ou simplement d'étendre le jeu d'instruction (fig. 6).

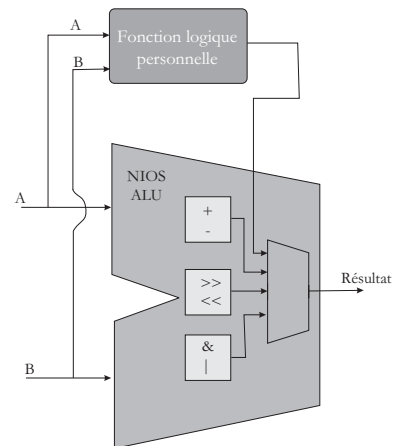


FIG. 6 – Instructions personnalisées d'un processeur NIOS

4 Exemple de tâche : Suivi de motifs

4.1 Technique de suivi

Afin d'illustrer notre méthode, nous proposons, dans ce paragraphe, les premiers résultats d'un algorithme de suivi. Cet algorithme s'appuie sur une méthode de corrélation dans le domaine fréquentiel. Ce suivi est réalisé en asservissant une fenêtre d'acquisition sur une zone de niveau de gris pré-sélectionnée. Suivant cet objectif, le détecteur actif est constitué d'une zone d'intérêt adressée dynamiquement et d'un asservissement de sa position. Le schéma global de cet asservissement est décrit par la figure 7. Cette tâche de suivi peut être formalisée par l'équation suivante :

$$\mathbf{P}_{\mathbf{Woi}}^{\mathcal{R}_i}(t+1) = \mathbf{P}_{\mathbf{Woi}}^{\mathcal{R}_i}(t) - \mathbf{P}_{\mathbf{Wref}}^{\mathcal{R}_f}(t)$$

$\mathbf{P}_{\mathbf{Woi}}^{\mathcal{R}_i}(t)$ est la position de la zone d'intérêt ($\mathbf{Woi}$) dans le repère image ($\mathcal{R}_i$), à l'instant t . $\mathbf{P}_{\mathbf{Wref}}^{\mathcal{R}_f}(t)$ est la position du motif de référence ($\mathbf{Wref}$), à l'instant t , dans le repère centré ($\mathcal{R}_f$), associé à la fenêtre d'intérêt.

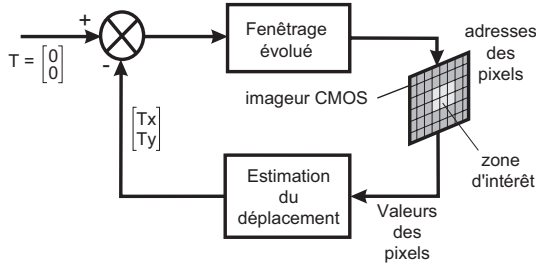


FIG. 7 – Schéma d'asservissement du détecteur actif

L'estimation de la position $\mathbf{P}_{\mathbf{Wref}}^{\mathcal{R}_f}(t)$ est évaluée comme le déplacement d'un motif entre deux images. Cette estimation est basée sur les méthodes présentées dans l'article de LL. Lucchese et al.[13] et l'article de M. Marcel et al. [14]. Elles sont basées sur la représentation fréquentielle qui permet d'estimer efficacement les déplacements planaires entre deux images. Évidemment, cet algorithme nécessite le calcul de transformées de Fourier discrètes des images.

4.2 Estimation du déplacement

L'estimation de la translation $\mathbf{T} = \begin{pmatrix} T_x \\ T_y \end{pmatrix}$ est effectuée à partir d'une corrélation de phase entre les images $i_1(\mathbf{x})$ et $i_2(\mathbf{x})$, en considérant la relation :

$$i_2(\mathbf{x}) = i_1(\mathbf{x} - \mathbf{T}) \quad (1)$$

avec $\mathbf{x}, \mathbf{T} \in \mathbb{R}^2$.

On note les transformées de Fourier des images i_n ($n = 1, 2$) :

$$I_n(\mathbf{k}) = \mathcal{F}(i_n(\mathbf{x})) = \int_{\mathbb{R}^2} i_n(\mathbf{x}).e^{-j.2\pi.\mathbf{k}^t.\mathbf{x}} d\mathbf{x}$$

avec $\mathbf{k} \in \mathbb{R}^2$.

L'équation 1 dans le domaine fréquentiel devient :

$$I_2(\mathbf{k}) = \int_{\mathbb{R}^2} i_1(\mathbf{x} - \mathbf{T}).e^{-j.2\pi.\mathbf{k}^t.(\mathbf{x} - \mathbf{T})} d\mathbf{x} = I_1(\mathbf{k}).e^{+j.2\pi.\mathbf{k}^t.\mathbf{T}} \quad (2)$$

soit $\Psi(\mathbf{k})$ le rapport des spectres des deux images i_1 et i_2 , on obtient :

$$\Psi(\mathbf{k}) = \frac{I_2(\mathbf{k})}{I_1(\mathbf{k})} = e^{j.2\pi.\mathbf{k}^t.\mathbf{T}} \quad (3)$$

Ensuite, il est simple d'estimer le vecteur translation $\mathbf{T}$ en calculant la transformée de Fourier $\Psi(\mathbf{k})$:

$$\mathcal{F}(\Psi(\mathbf{k})) = \mathcal{F}(e^{j.2\pi.\mathbf{k}^t.\mathbf{T}}) = \delta(\mathbf{T}) \quad (4)$$

4.3 Résultats Expérimentaux

Dans ce paragraphe, nous présentons les résultats temporels de notre approche. Afin d'illustrer, les performances de la méthode précédente, nous avons introduit dans le mécanisme de suivi, une phase d'acquisition de l'image entière. Cette phase permet de visualiser la scène globale toutes les 720 ms (fig. 9). Pour interpréter cette séquence, il faut noter, que durant le suivi, uniquement la zone repérée par le carré coloré est exploitée. Pour l'expérimentation de la figure 9, la taille de la fenêtre d'intérêt de 32×32 pixels permet d'obtenir un temps d'acquisition de $570 \mu\text{s}$ et un temps de cycle total de 1.8 ms (temps acquisition + temps de calcul). Ainsi, la fréquence de travail du dispositif est de 550 Hz. Pour préciser les performances de notre méthode, nous donnons dans le tableau suivant, une évaluation des temps, qui caractérise le système, en fonction de la taille de la fenêtre d'intérêt :

taille de la fenêtre	16x16	32x32	64x64	128x128
temps d'acquisition	200 μs	570 μs	2 ms	8 ms
temps de cycle	450 μs	1.8 ms	8.6 ms	78 ms
cadence de suivi	2220 f/s	550 f/s	117 f/s	12 f/s

Ces résultats révèlent la rapidité de réaction du dispositif qui est une propriété déterminante pour un système de suivi. Il est clair que cette approche est lucrative pour développer des systèmes de vision active efficaces.

4.4 Amélioration de l'algorithme

La méthode précédente est très efficace pour suivre une cible en translation planaire pure. Mais, le signal $\delta(\mathbf{T})$ (équation 4), qui permet d'évaluer le déplacement, est fortement distordu lorsque les mouvements du motif ne peuvent plus être considérés comme planaires ou lorsque des rotations apparaissent.

Afin d'éviter ce problème, une étape supplémentaire est introduite dans l'algorithme. Une estimation de la variation

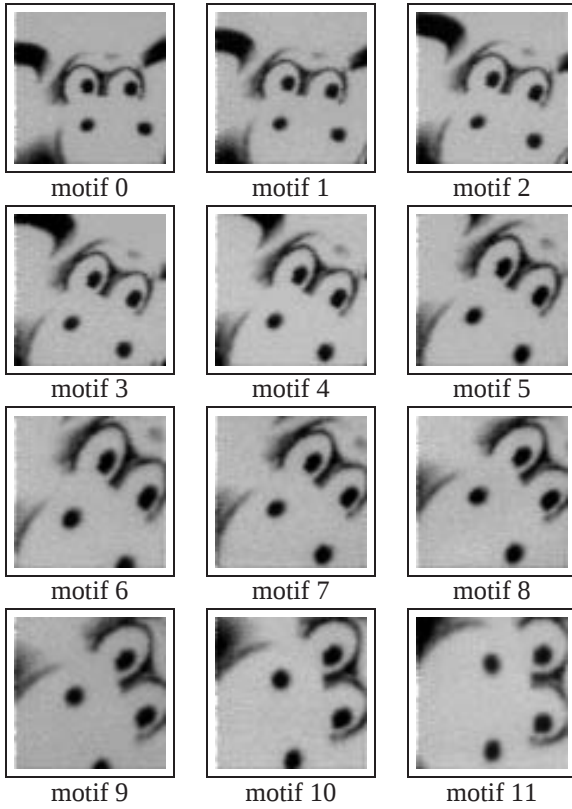


FIG. 8 – Évolution du motif de référence de taille 64×64 pixels. Pendant le suivi, le motif de référence est mis à jour suivant la dynamique des distorsions du signal de corrélation $\delta(\mathbf{T})$.

du signal $\delta(\mathbf{T})$ est réalisée à partir de l'équation suivante :

$$\Upsilon = \frac{\text{Max}(\delta(\mathbf{0})) - \text{Min}(\delta(\mathbf{0}))}{\text{Max}(\delta(\mathbf{T})) - \text{Min}(\delta(\mathbf{T}))} \quad (5)$$

ensuite une mise à jour du motif de référence est effectuée lorsque cette variation devient trop importante :

$$\Upsilon \geq \text{Seuil} \Rightarrow \mathbf{W}_{\text{ref}} = \mathbf{W}_{\text{oi}}(\mathbf{t})$$

Grâce à cette simple modification, le suivi peut être réalisé malgré l'apparition de rotations et de changements d'échelle modérés. Une exemple d'évolution du motif de référence est présenté figure 8. L'efficacité de cette méthode est directement dépendante de la cadence de mise à jour des informations visuelles. Cette cadence varie avec la taille des images considérées. Lorsque la cible est très rapide ou lorsque les dimensions de la fenêtre d'intérêt sont trop importantes, on note une erreur de position dans le motif de référence jusqu'à ce que cette référence soit totalement faussée. Le succès de cette méthode impose donc l'utilisation d'images de faible taille.

5 Conclusion

Dans cet article, nous avons utilisé la notion de *Détecteur Actif*, et proposé une architecture basée sur un imageur CMOS. L'état actuel de nos travaux, nous permet d'affirmer que l'exploitation de l'accès aléatoire aux pixels et du concept de vision active dans le processus de perception, doit mener au développement de nouveaux capteurs de vision active. Une application de suivi par corrélation a été développée à partir du concept de détecteur actif. Cette approche a été testée sur une grande variété de cibles avec succès.

Nos futures travaux consisteront à développer d'autres détecteurs actifs, et à inclure dans notre stratégie de vision le contrôle du zoom et de l'iris de la caméra.

Références

- [1] Aloimonos J.Y., Weiss I. and Bandyopadhyay A. Active vision. In *Proceedings of the 1st International Conference on Computer Vision*, pages 35–54, London, England, June 1987.
- [2] Andersen C.S. *A Framework for control of a camera head*. PhD thesis, Laboratory of image analysis, Aalborg University, Denmark, 1996.
- [3] Bajcsy R. Active perception. *IEEE Proceedings*, 76(8) :996–1006, August 1988.
- [4] Ballard D. Animate vision. *Artificial Intelligence*, 48(1) :1–27, February 1991.
- [5] Barron J.L., Fleet D.J. and Beauchemin S.S. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1) :43–77, 1994.
- [6] Batista J., Peixoto P. and Araújo H. Binocular tracking and accommodation controlled by retinal motion flow. In *International Conference on Pattern Recognition*, pages 1171–1174, Barcelona, Spain, September 2000.
- [7] Bernard Ch. *Ondelettes et Problèmes mal posés : la mesure du flot optique et l'interpolation irrégulière*. PhD thesis, Thèse de l'École Polytechnique, 24 novembre 1999.
- [8] Brajovic V. and Kanade T. Sensory attention : Computational sensor paradigm for low-latency adaptive vision. *Image Understanding Workshop*, May 1997.
- [9] Cavadore C. Design and characterization of cmos-aps. *SUPAERO (Toulouse) and Thomson TSC CCD, Ph.D Thesis*, 1998.
- [10] Chalimbaud P., Berry F. and Martinet P. Active vision based on cmos camera : an active detector approach. In *Proceedings of the IEEE International Symposium on Robots*, Stockholm, Sweden, October 2002.
- [11] Indiveri G., Mürer R. and Kramer J. Active vision using an analog vlsi model of selective attention. *IEEE Transactions on Circuits and Systems-II : Analog and digital signal processing*, 48 :492–500, May 2001.

- [12] Kato T. and Floreano D. An evolutionary active-vision system. *In Proceedings of the Congress on Evolutionary Computation*, 2001.
- [13] Lucchese, L. and Cortelazzo, G.M. A noise-robust frequency domain techniques for estimating planar rotations. *IEEE Transactions on signal processing*, 48(6) :1769–1786, 2000.
- [14] Marcel B., Briot M. and Murrieta R. Calcul de translation et rotation par la transformée de fourier. *Traitement du signal*, 14(2) :135–150, 1997.
- [15] Pahlavan K., Uhlin T. and Eklund J.O. *Active vision as a methodology*, pages 19–46. Robotics and Automated Systems, ISBN 981-02-1364-6. Lawrence Erlbaum Associates Publishers, 1993.
- [16] Ullman S. Visual routines. *Cognition*, 18 :97–160, 1984.
- [17] Vieville T. A few steps towards 3d active vision. *Springer Series in Information Sciences*, 33, 1997.



FIG. 9 – Suivi d'un motif de niveaux de gris de dimensions 32×32 pixels. L'intervalle de temps entre deux images entières vaut 720 ms. La position du motif est actualisée toutes les 1.8 ms.
