

# Cyclope: Une méthode de discrimination spectrale innovante dans le cadre d'un capteur de vision 3D intégré temps réel sans fils

Anthony KOLAR<sup>1</sup>, Olivier ROMAIN<sup>1</sup>, Bertrand GRANADO<sup>2</sup>

<sup>1</sup>Groupe SYEL- EA2385  
Université Pierre et Marie CURIE - Paris VI  
4 place Jussieu, Paris France

<sup>2</sup>ETIS  
CNRS - ENSEA  
Université Cergy Pontoise  
F95000 Cergy, France

kolar@lisif.jussieu.fr, romain@upmc.fr, bertrand.granado@ensea.fr

**Résumé** – Le travail présenté ici est lié à Cyclope, un système de vision 3D active intégré. Le défi est de réaliser un capteur intégré effectuant une reconstruction 3D temps réel avec des contraintes d'encombrement, de consommation et de ressource de calcul. Dans ce document, nous présentons une méthode d'acquisition qui permet d'acquérir à la fois la texture et le motif à 25 image/s. Cette méthode utilise un projecteur de motif infrarouge et un imageur CMOS avec un temps d'intégration programmable. Ce capteur est réalisé en technologie CMOS standard avec aucun filtrage optique. Nous présentons plusieurs résultats avec la réalisation d'un démonstrateur macroscopique.

**Abstract** – The work presented here is related to Cyclope, an integrated active 3D vision system. The challenge is to realize an integrated sensor which achieves real-time depth vision with constraints on size, consumption and computational resources. In this paper, we present acquisition method which allows to grab both texture and pattern images at least at 25 frames/s. The method uses a IR laser pattern projector and CMOS image sensor with programmable integration time. This sensor is realized in a standard CMOS technology and does not involve any special optical filtering. We present some experimental results with the realization of a large scale demonstrator.

## 1 Introduction

La nécessité d'avoir des systèmes capables de percevoir le relief est apparue dès le début de la photographie. Cette nécessité a conduit à la réalisation de dispositifs macroscopiques qui avaient tous comme point commun la perception du relief d'un point de vue humain. Ces systèmes utilisent le principe de la stéréoscopie, c'est à dire l'appariement de deux images provenant de deux capteurs acquérant la même scène sous un angle différent à la manière de la vision animale. Les domaines d'utilisation de ces systèmes sont par exemple la sécurité, le multi-média, la description de scène ou navigation pour robot [1]. La perception d'une image en trois dimensions d'un environnement microscopique est une autre affaire. Les techniques utilisées sont basées sur la prise de plusieurs images successives. Cette technique fonctionne si et seulement si rien n'a bougé dans la scène entre les différentes acquisitions, les images en relief réalisées par ce type de méthode sont entachées d'une certaine incertitude.

Les avancées technologiques laissent entrevoir la possibilité de réaliser des systèmes intégrés de vision en trois dimensions sur un seul circuit de silicium, permettant ainsi de réaliser

un système de perception du relief pour des environnements microscopiques. Cela ouvre de nouveaux champs applicatifs, comme la vision précise en trois dimensions in vivo dans le corps humain lors d'examen de type endoscopique. A ce jour peu de travaux tendent à concevoir un capteur intégré polyvalent [2] malgré les avantages évidents pour ces nouvelles applications. Jusque là, la principale restriction était principalement due à la taille et à la vitesse de fonctionnement des appareils.

C'est tout l'objet du projet *Cyclope* dans lequel est menée une étude pour déterminer les moyens d'arriver à intégrer les traitements nécessaires à la reconstruction de la troisième dimension dans un seul circuit. Afin de parvenir à cet objectif, nous présentons ici le capteur que nous avons conçu, *Cyclope* est un capteur de vision 3D temps réel intégré avec des capacités de communication sans fil [4].

## 2 Cyclope

Cyclope est un capteur de vision 3D intégré basé sur la stéréoscopie active. Il est composé de trois parties essentielles :

1. Un bloc d'instrumentation permettant à la fois la capture

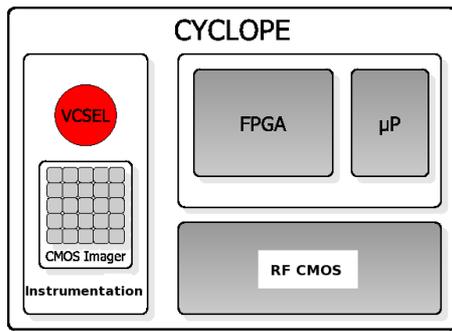


FIGURE 1 – Diagramme de capteur Cyclope

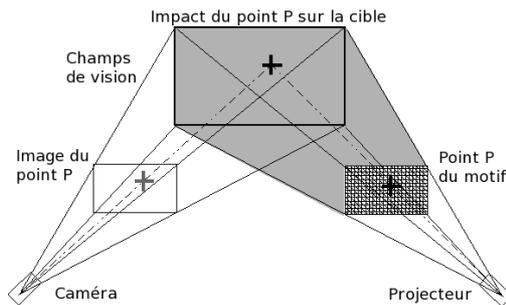


FIGURE 2 – Système stéréoscopique

d'images et la génération du motif laser.

2. Un bloc de traitement pour le contrôle du système et le traitement des données. Ce bloc est l'union d'une partie reconfigurable de type FPGA destinée à l'exécution des traitements parallèles et d'un micro-processeur pour les traitements séquentiels.
3. Un bloc RF servant à la transmission des résultats.

Pour toutes ces parties, nous avons choisi d'utiliser les technologies les plus adéquates et de les intégrer à *Cyclope* sous la forme d'un *SiP* (System in Package). Nous utilisons une technologie CMOS pour le capteur d'images et les unités de transformation, une technologie GaAs pour la structure VCSEL du projecteur et enfin une technologie RF-CMOS pour l'unité de communication. Ces éléments peuvent être intégrés dans une puce à l'échelle du capteur [5].

Le capteur d'images est multi-spectral et permet de saisir simultanément une image projetée du motif dans le spectre infrarouge et une image de la texture des objets dans le spectre visible pour générer une représentation en 3D texturée. Le motif est une matrice de 49 points régulièrement espacés obtenus en couplant le projecteur laser à une optique de diffraction. Les éléments de ce couple stéréoscopique (voir figure 2) sont fixe l'un par rapport à l'autre.

La connaissance a priori des caractéristiques du système stéréoscopique, ainsi que de la structure du motif projeté nous permet de déterminer la distance d'un point du motif sur la scène

à partir de ces coordonnées dans l'image.

## 3 Reconstruction et acquisition

### 3.1 Principe de reconstruction

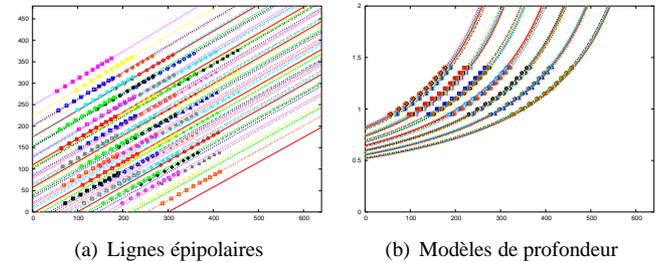


FIGURE 3 – Résultats de calibration

La reconstruction 3D est rendue possible grâce à la connaissance des paramètres du couple stéréoscopique (voir figure 3) déterminés lors d'une phase de calibration préalablement réalisée [6]. Il est alors possible d'extraire un modèle de profondeur nous permettant d'exprimer la distance entre le couple stéréoscopique et le point d'impact du laser en fonction de ses coordonnées dans l'image. Ces modèles sont utilisés en temps réel afin d'effectuer l'appariement de chaque point laser et de calculer la profondeur en utilisant les paramètres appropriés. L'architecture globale est divisée en trois parties principales :

- Une unité pour les traitements d'images bas niveaux et l'extraction du motif infrarouge projeté.
- Une unité d'appariement et de calcul de profondeur.
- Une unité de communication sans fil.

En outre, une mémoire double port est utilisée pour le stockage des images permettant l'acquisition asynchrone des images. Les traitements bas niveaux mis en oeuvre sont le seuillage, la segmentation et la détermination du barycentre de chaque point laser. La communication sans fils est gérée par un coeur de processeur logiciel *Picoblaze*.

La principale contrainte de notre application consiste en l'acquisition quasi simultanée de la texture de l'image dans le spectre du visible et du motif projeté, et de concevoir une architecture de contrôle assurant un traitement de l'information en temps réel. Ayant opté pour un laser dont la longueur d'onde est dans le proche infrarouge, pour ne pas dénaturer la scène, plusieurs solutions étaient possible [7] [8]. Malheureusement, aucune d'entre elles n'étaient pleinement satisfaisante. En effet, nous voulions une solution capable d'assurer l'adaptabilité à des systèmes existant en technologie standard de Silicium afin de leur fournir la capacité à déterminer une information 3D sans changement majeur. De plus, l'utilisation de filtre optique est à proscrire afin de ne pas diminuer la résolution spectrale de l'imageur. Dans ce contexte, nous avons opté pour une approche d'acquisition originale basée sur l'existence du gap énergétique entre la scène et le motif projeté à cette longueur d'onde.

### 3.2 Méthode de discrimination spectrale énergétique

Pour être capable d'opérer une discrimination spectrale sur un capteur réalisé en technologie CMOS standard sans traitement vidéo, une nouvelle méthode a été développée [9]. L'approche spectrale normalement utilisée dans ce genre d'application a été mise de côté pour laisser place à une approche énergétique et temporelle.

En fait, la méthode se rapproche de celle utilisée dans le monde radar. En simplifiant le processus, les radaristes déterminent la présence et dans certain cas aussi la nature d'une cible par sa signature énergétique. Le radar envoie une onde qui sera réfléchié lorsque cette dernière touchera un objet.

C'est sur ce principe que repose la discrimination de *Cyclope*. Il est possible de faire une analogie entre l'émission d'une onde radio et celle d'une onde lumineuse. L'approche adoptée est à la fois énergétique et temporelle :

**Energétique** car se basant sur le gap énergétique entre la texture et le motif. En gardant l'analogie avec le radar, deux ondes sont réfléchies sur le capteur : celle émise par le projecteur laser, et celle émise par les sources de lumière ambiante.

**Temporelle** car la discrimination sur le capteur par une telle méthode, c'est à dire en modifiant le temps d'intégration, implique deux prises d'images consécutives : une pour la capture du motif avec un temps d'intégration court, et une pour l'acquisition de la texture avec un temps d'intégration long. Pour cela, l'acquisition complète s'effectue en deux temps (figure 4) :

1. la première phase consiste en l'acquisition du motif et en émettant une impulsion laser. Le résultat est alors mémorisé et traité en temps réel pour fournir l'information de profondeur;
2. la seconde phase permet la capture de la texture sans laser.

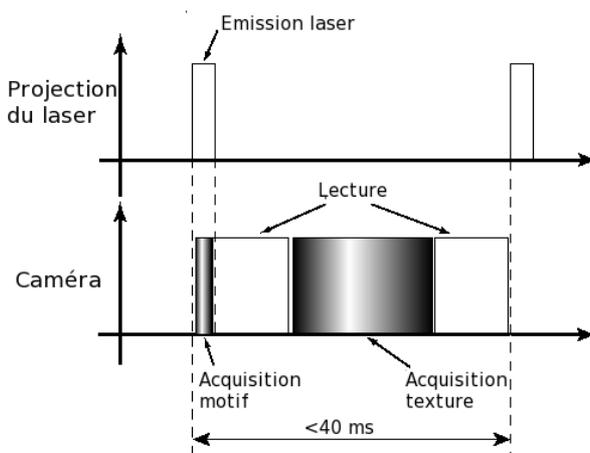


FIGURE 4 – Séquençage temporel de l'acquisition

Cependant, pour remplir la contrainte du temps réel, c'est à dire 25 images par seconde, il est obligatoire que l'acquisition complète, ainsi que le traitement, soient exécutés en moins de 40 ms. Le temps global d'acquisition est donné par l'équation 1.

$$T_{total} = 2 \cdot T_{rst} + 2 \cdot T_{rd} + T_{intVI} + T_{intIR} \quad (1)$$

La mise en oeuvre de cette méthode est assurée par l'architecture présentée sur la figure 5.

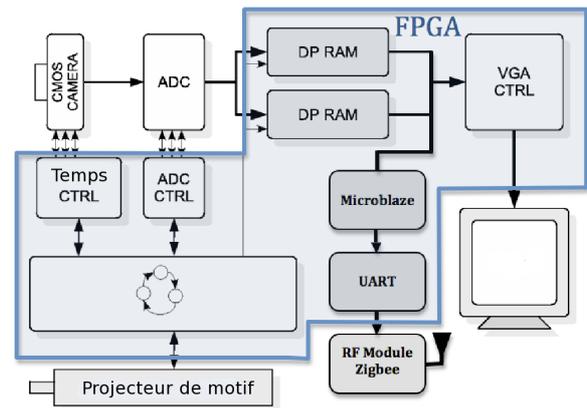


FIGURE 5 – Architecture d'acquisition et de synchronisation du laser de Cyclope

## 4 Démonstrateur macroscopique expérimental et résultats



FIGURE 6 – Démonstrateur macroscopique de Cyclope

Une étude de faisabilité de l'intégration sous la forme d'un *SiP* [6] a été faite avant la conception du démonstrateur. Ce démonstrateur a pour objectif de valider notre approche sur le principe et ne représente pas une version finale du projet.

Ainsi, ce dernier a été réalisé avec un facteur d'échelle dans l'objectif d'être compatible avec ce que pourrait être l'intégration en *SiP*. Le démonstrateur est présenté sur la figure 6, il est composé d'une caméra CMOS d'une résolution de 4096 pixels dotée de CAN 8 bits externe, d'un projecteur de motif infrarouge générant une matrice de 49 points, d'une carte FPGA Xilinx Virtex2Pro et d'un module de communication sans fil ZIGBEE. Sur le FPGA est mis en oeuvre les architectures de contrôle et de synchronisation de l'acquisition des images, du projecteur laser, des convertisseurs analogique numérique, de la communication sans fils et de l'affichage via une interface VGA. La dimension de la base stéréoscopique est de 5 cm.

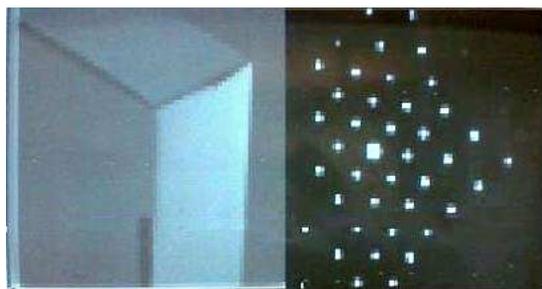


FIGURE 7 – Acquisition d'un objet avec texture et motif séparés

Le résultat est présenté sur la figure 7, l'image de gauche est l'acquisition de la texture de la scène, tandis que celle de droite est l'acquisition du motif projeté. De plus, le tableau 1 démontre le respect de la contrainte temps-réel que nous nous sommes fixé. La consommation des ressources du FPGA par l'architecture est présentée dans le tableau 2.

49 points laser - taux d'occupation du FPGA de 37%	
Temps d'acquisition total	< 25ms
Temps de lecture total	< 1ms
Cadence vidéo	38fps

TABLE 1 – Tableau récapitulatif des performances temporelles

Architecture	Cib slices	Latches	LUT	RAM
Acquisition	309	337	618	4
Total libre	13693	29060	27392	136

TABLE 2 – Utilisation des ressources matérielles

## 5 Conclusion et perspectives

Dans le but d'achever une reconstruction 3D en temps réel dans un système de vision active intégré, nous avons présenté dans ce papier *Cyclope*. Il s'agit d'un capteur de vision 3D intégré regroupant tous les éléments nécessaires à la vision et

à communiquer les résultats. Nous avons également présenté une solution afin de réaliser un tel capteur à l'échelle d'une micro-puce via une approche *SiP*. Ensuite, nous avons présenté une méthode d'acquisition originale basée sur l'existence du gap énergétique existant entre la scène et le motif qui permet l'emploi d'une technologie standard en Silicium et sans filtre optique. Pour finir, nous avons réalisé un démonstrateur macroscopique de *Cyclope* afin de valider notre approche et présenter quelques résultats d'acquisition et d'implémentation avérant de la viabilité de notre approche.

A court terme, il sera nécessaire de réaliser l'étude approfondie de la consommation de l'architecture globale. Puis à moyen terme, la réalisation d'une intégration monolithique devra être faite afin d'effectuer une étude de consommation sur le système complet.

## Références

- [1] E. Cortizo, A. Morene Yeras, J.R. Lepore, and M. Garavaglia, "Application of structured illumination method to study the topography of the sole of the foot during a walk," *Optics and Laser in Engineering*, vol. 40, pp. 117–132, 2003.
- [2] O. Elkhaili, O.M. Schrey, P. Mengel, M. Petermann, W. Brockherde, and B.J. Hosticka, "A 4 x 64 pixel cmos image sensor for 3d measurement applications," *IEEE journal of solid state circuits*, vol. 39, pp. 1208–1212, 2004.
- [3] GivenImaging, *Given Diagnostic System, The Platform for PillCam Endoscopy*, 2005.
- [4] T. Graba, B. Granado, O. Romain, T. Ea, A. Pinna, and P. Garda, "Cyclope : an integrated real-time 3d image sensor," in *XIX Conference on design of circuits and integrated systems*, 2004.
- [5] H. Machida, J. Nitta, A. Seko, and H. Kobayashi, "High efficiency fiber grating for producing multiple beams of uniform intensity," *Applied Optics*, vol. 23, pp. 330–332, 1984.
- [6] T. Graba, *Etude d'une architecture de traitement pour un capteur intégré de vision 3D*, Ph.D. thesis, Université Pierre et Marie Curie, 2006.
- [7] Z. Pei, L.S. Lai, H.P. Hwang, Y.T. Tseng, and M.-J. Tsai C.S. Liang, "Si1 xgex=si multi-quantum well phototransistor for near-infrared operation," *Physical*, 2003.
- [8] Rajeev Pillai David Starikov, Chris Boney and Abdelhak Bensaoula, "Dual-band uv/ir optical sensors for fire and flame detection and target recognition," *Senson for Industry Conference*, 2004.
- [9] A. Kolar, T. Graba, A. Pinna, O. Romain, B. Granado, and E. Belhaire, "Smart bi-spectral image sensor for 3d vision," *Ieee Sensors*, pp. 577–580, Oct 2007.