



HAL
open science

Reconnaissance et localisation spatiale d'objets. Travaux Pratiques de Vision par Ordinateur en Licence Professionnelle en Mécatronique.

Jean-Louis Dillenseger, Carine Kulik, Yann Dieulangard

► **To cite this version:**

Jean-Louis Dillenseger, Carine Kulik, Yann Dieulangard. Reconnaissance et localisation spatiale d'objets. Travaux Pratiques de Vision par Ordinateur en Licence Professionnelle en Mécatronique.. Cetsis, Mar 2010, Grenoble, France. papier 83, 5 p. <hal-00805310>

HAL Id: hal-00805310

<https://univ-rennes.hal.science/hal-00805310v1>

Submitted on 27 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Reconnaissance et localisation spatiale d'objets. Travaux Pratiques de Vision par Ordinateur en Licence Professionnelle en Mécatronique.

Jean-Louis Dillenseger, Carine Kulik, Virginie Le Rolle, Yann Dieulangard*

Jean-louis.dillenseger@univ-rennes1.fr

IUT Rennes, dpt GEII, rue du Clos Courtel, Rennes

*EDIXIA, ZART des Perrières, Vern/Seiche

RESUME : L'objectif global de l'enseignement est de faire comprendre et de mettre en place une chaîne de traitements permettant de reconnaître et de localiser un objet spécifique dans l'espace. L'étude de cette chaîne de traitement englobe l'apprentissage de procédures d'acquisition d'images, de calibration de caméra, de segmentation/reconnaissance de formes et de vision stéréoscopique. D'un point de vue matériel, le dispositif de vision est constitué de deux caméras (monochrome et couleur) pouvant faire des acquisitions stéréoscopiques. Les caméras et les traitements qui y sont rattachés sont pilotés par le logiciel LabVIEW associé à la bibliothèque graphique NI-Vision. Les traitements mathématiques se font à l'aide du logiciel Matlab.

Mots clés : vision stéréoscopique, calibration, Licence Professionnelle.

1 INTRODUCTION

La Mécatronique est une discipline encore en émergence qui résulte de l'association de la mécanique, de l'électronique, de l'automatique et de l'informatique industrielle. La Licence Professionnelle a pour objectif de compléter la formation des techniciens possédant une spécialité centrée sur la mécanique, l'électronique ou l'automatisme pour leur apporter les connaissances technologiques transversales essentielles dans un contexte de conception et fabrication intégrée de systèmes automatisés et robotiques. Plus concrètement, elle a comme vocation de former des techniciens à haute qualification susceptibles de gérer des plateformes multitechniques (suivi de l'installation, mise en œuvre et réglage des différents modules, maintenance et prospective d'évolution).

Un des enseignements de cette licence porte sur la commande des systèmes industriels et plus particulièrement sur des aspects Robotique/Vision. L'objectif global de cet enseignement est de faire comprendre les différents éléments d'une chaîne de traitements d'images. Afin d'illustrer plus concrètement ces différents aspects, nous avons décidé de réaliser une série de Travaux Pratiques portant sur la reconnaissance et la localisation spatiale d'objets simples. Cette application est suffisamment complexe pour intégrer et associer différents modules fondamentaux de la vision par ordinateur allant des aspects d'acquisition d'images, de calibration de caméras, de segmentation/reconnaissance de formes et de vision stéréoscopique. L'originalité de la démarche que nous proposons est d'aller au delà de la simple démonstration de quelques modules de traitement d'images ou d'extraction de formes, généralement proposée dans le cadre d'enseignement en Licence Professionnelle, en montrant les liens réels qui existent entre l'image et son environnement. C'est pour cela que nous insisterons sur les notions de calibration (plane et 3D) ainsi que sur les notions de reconstruction 3D.

De telles notions se retrouvent dans le milieu industriel lors de procédures de contrôle par la vision ou de commande de robots. Les procédures de calibration liant l'image et l'espace sont primordiales pour ces types d'applications,

L'intérêt de ces modules et plus encore le formalisme de ces notions sont toutefois assez difficiles à comprendre. Nous nous proposons donc d'aborder ces notions sous l'aspect de leur utilité dans la chaîne plutôt que sous un aspect mathématique trop complexe pour le niveau d'étude des étudiants qui suivent cette formation de Licence Professionnelle. En effet, ces étudiants ont un niveau de Technicien Supérieur (titulaire d'un DUT GEII, GIM ou GMP ou d'un BTS MAI, MI, électrotechnique, ...). L'objectif attendu des TP n'est donc pas qu'ils maîtrisent parfaitement les aspects mathématiques ou algorithmiques des différentes procédures d'acquisition et de traitement d'images, mais plutôt qu'ils en comprennent les principes généraux et la façon de les associer. Un logiciel de programmation graphique de type LabVIEW¹ de National Instruments associé à une bibliothèque graphique de type NIVision est donc parfaitement adapté à répondre aux objectifs de ces Travaux Pratiques. Ce papier décrit donc les différents thèmes abordés, leurs aspects matériels et également logiciels.

2 APPLICATION

L'objectif applicatif de cette série de TP est de reconnaître, de localiser spatialement un cylindre vert parmi d'autres cylindres (voir figure 1).

L'objectif à terme est de coupler cette partie vision à une partie opérative, un bras de robot, se saisissant de cet objet. La localisation 3D se fait par vision stéréoscopique à partir de 2 caméras (une caméra monochrome et une caméra couleur).

Nous avons choisi de reconnaître et localiser des cylindres car ces objets simplifient l'étape de reconnaissance de formes. En effet, l'image d'un cylindre vue par une

¹ <http://www.ni.com/labview/f/>

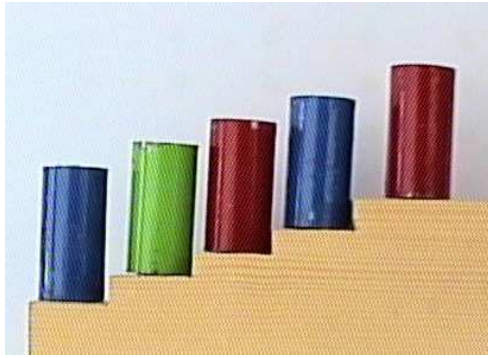


fig 1 : Cylindre vert (deuxième à partir de la gauche) parmi d'autres cylindres de couleur rouge ou bleu. Cible à reconnaître et localiser dans l'espace.

caméra se présente en 2D sous la forme d'une silhouette de couler de forme rectangulaire. Si la caméra se déplace dans un plan perpendiculaire à l'axe des cylindres et quelle que soit sa position, les cylindres seront toujours perçus sur les images comme des silhouettes rectangulaires. La phase de reconnaissance se résume alors à segmenter et reconnaître des formes rectangulaires sur les images.

L'intérêt pédagogique de ce sujet réside dans le fait qu'il illustre une chaîne complète de vision par ordinateur avec des parties matérielles (caméras couleur et monochrome, carte d'acquisition, banc de calibration) et des suites de traitements adaptés dans des domaines variés (segmentation d'images en niveau de gris et couleurs –plans de couleurs-, reconnaissance de formes simples, calibration 2D et 3D et vision stéréoscopique). Cette chaîne sera véritablement complète avec la réalisation de la partie opérative.

2.1 Partie matérielle

Le double objectif de reconnaissance de formes et de localisation spatiale nous a tout naturellement conduit à envisager une solution de deux caméras en vision stéréoscopique, dont une caméra couleur pour illustrer la segmentation d'images basée sur les plans couleurs. Pour le choix des caméras nous avons voulu privilégier la résolution par rapport à la vitesse d'acquisition. Nous avons opté pour deux caméras Philips analogiques de "haute résolution" (de l'ordre de 750x500).

Afin de faciliter l'intégration de la capture d'images dans LabVIEW, les caméras seront pilotées par deux cartes d'acquisition de National Instrument IMAQ PCI 1407 et IMAQ PCI 1411 pour, respectivement, l'acquisition monochrome et couleur.

Ces caméras sont fixées sur un rail sur lequel coulisse une grille de calibration et sur lequel est posé un plateau avec les cylindres (figure 2). Cette installation fixe ainsi le contexte de fonctionnement et de calibration des caméras.

2.2 Mise en œuvre logicielle

Comme il a été dit dans l'introduction, l'utilisation d'un logiciel de programmation graphique de type LabVIEW associé à une bibliothèque graphique de type

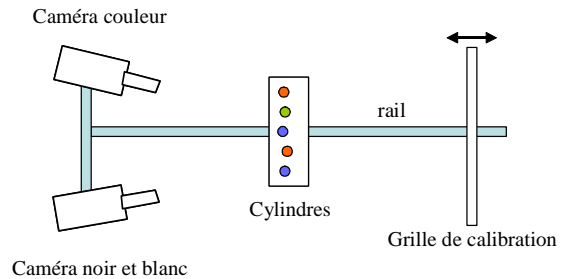


fig 2 : Disposition matérielle vue du dessus. Deux caméras fixées sur un rail en position stéréoscopiques. Une grille de calibration peut se déplacer sur le rail. Un plateau portant les cylindres peut être posé sur le rail.

NIVision nous a paru vraiment adaptée pour faire ressortir les principes fonctionnels ou mathématiques des différents traitements sans être gêné par des difficultés de programmation. Les différentes routines de LabVIEW et NIVision se présentent sous la forme d'objets graphiques (Virtual Instrument, VI, selon la terminologie de National Instruments) qui sont associés entre eux pour former le traitement final. Lors de l'élaboration des séances, nous nous sommes heurtés à une des grandes difficultés de LabVIEW que représente la conversion de données (clusters, tableaux, variables) entre les différents VI. Pour des raisons d'efficacité, nous ne pouvions demander aux étudiants d'utiliser directement les VIs de labVIEW et NIVision. Nous avons donc dû réaliser nos propres VIs intégrant les VIs natifs et la conversion de données. Les étudiants ont donc à leur disposition une bibliothèque documentée de VIs spécifiques à l'application. Leur travail consiste à associer ces VIs en fonction de la notion étudiée.

Pour des soucis d'économie, nous ne disposons que d'un seul jeu de caméras associé à un PC serveur. La communication entre les PCs de travail et le PC serveur s'est faite par envoi de commandes via les modules de communications TCP/IP de LABVIEW. Sur les PCs de travail, nous avons créé des VIs d'acquisition d'images dont la fonction est d'envoyer une requête d'acquisition au PC serveur et de récupérer l'image acquise par partage de fichiers entre PCs. L'utilisation de ces VIs rend l'acquisition complètement transparente pour les étudiants.

2.3 Notions étudiées

La chaîne complète de traitements peut se décomposer de la façon suivante (figure 3) :

- La procédure d'acquisition : Nous voulions que les étudiants prennent conscience des notions suivantes :
 - Les paramètres d'acquisitions d'images : choix de la caméra, format des images délivrées par la carte, séquençement des acquisitions (snap, grab, etc.). Une démonstration de ces différents paramètres leur est faite.

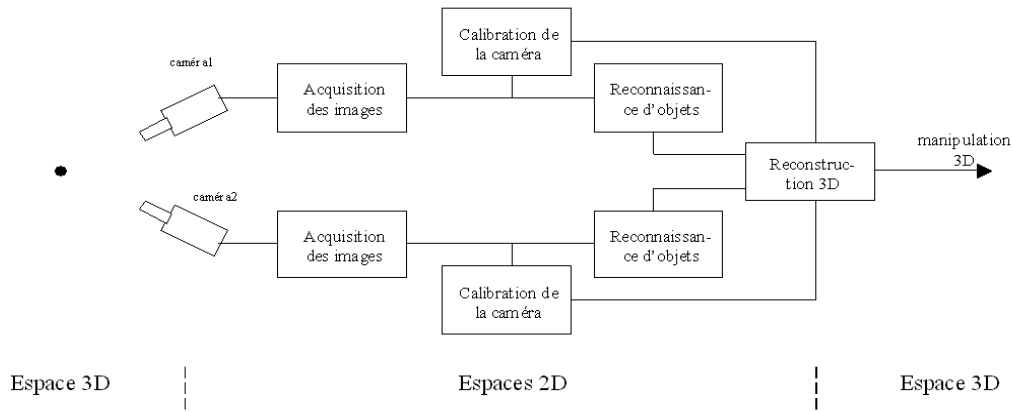


fig3 : Chaîne complète de traitements

- Le codage des images (information codée sur 8 bits pour un pixel d'image monochrome, 3x8bits pour une image couleur, codage des couleurs en RGB, en HSL, etc.). L'objectif à terme étant de préparer la segmentation d'images couleurs.
- Les paramètres d'une caméra (position, orientation, focale, distorsion). Pour cela, un VI de modèle Pin Hole de caméra leur est proposée. Les étudiants font varier les paramètres afin d'ajuster la projection d'une grille virtuelle sur l'image d'une grille réelle.
- La procédure de calibration d'une caméra qui consiste à définir la relation qui existe entre un pixel de l'écran et l'espace 3D. Les étudiants abordent dans un premier temps la calibration plane puis étendent cette calibration plane à une calibration 3D par une méthode multi-plans. Ces calibrations se font sur un banc de calibration à l'aide de grilles de points dont les coordonnées 3D sont connues dans le repère réel (figure 4). Les croix indiquent des positions connues sur le plan et serviront pour l'évaluation de calibrations.

- Calibration plane : elle est utilisée pour effectuer des mesures de distances réelles sur une image 2D d'objets se trouvant dans un plan (mesure de distances en microscopie par exemple).
- Calibration 3D multiplans [1] (bi-plans dans notre cas). L'idée est d'estimer l'intersection de la droite de projection avec plusieurs plans connus. Ceci nous donne alors la relation directe entre pixel et droite de projection, sans connaissances ou hypothèses a priori. Le principe est le suivant (figure 5) :

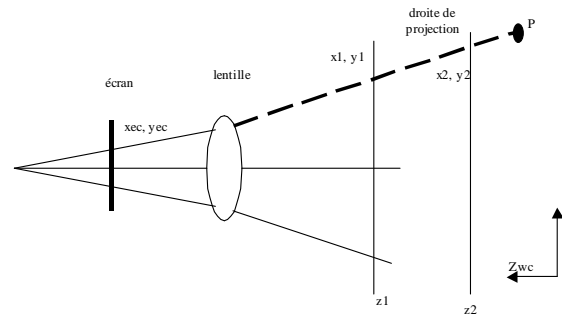


fig.5 : Principe de la calibration multi-plans

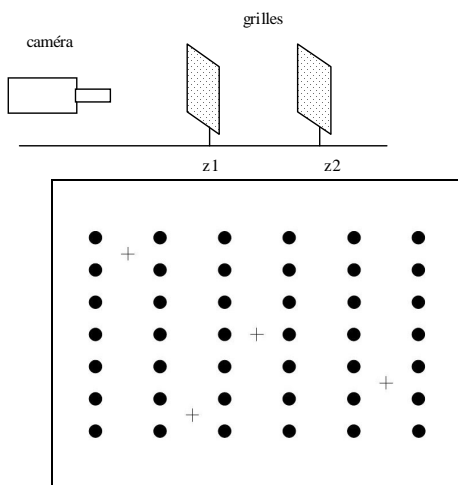


fig 4 : Banc et grille de calibration

Prenons deux plans parallèles et perpendiculaires à l'axe des z_{wc} . Le premier plan, le plan 1, est caractérisé par $z=z_1$, le second, le plan 2, par $z=z_2$. Prenons la droite de projection qui aboutit à la formation du pixel $\mathbf{P}_{ec}(x_{ec}, y_{ec})$. La droite de projection va traverser les plan 1 et 2 respectivement aux coordonnées x_1, y_1 et x_2, y_2 respectivement aux coordonnées $\mathbf{P}_1(x_1, y_1, z_1)$ et $\mathbf{P}_2(x_2, y_2, z_2)$. Ces deux points \mathbf{P}_1 et \mathbf{P}_2 permettent de définir la droite de projection. L'équation paramétrique de la droite de projection est alors : $\mathbf{P}_1 + (\mathbf{P}_2 - \mathbf{P}_1).t$. Le but de la calibration est d'associer pour tous les pixels $\mathbf{P}_{ec}(x_{ec}, y_{ec})$ de l'écran, les points \mathbf{P}_1 et \mathbf{P}_2 . correspondant aux deux plans connus Ceci peut se faire par deux calibrations planes, l'une pour le plan 1 l'autre pour le plan 2.

- Les objets d'intérêts seront isolés et localisés dans l'image à l'aide de procédures de reconnaissances de formes. Les objets à reconnaître sont assez simples. Les étudiants utilisent des méthodes de segmentation par seuillage (seuillage de niveaux de gris pour la caméra N&B, seuillage dans l'espace de couleur HSL pour l'image couleur). La reconnaissance et localisation des formes sur l'image sont ensuite menées à l'aide de VIs existant dans la bibliothèque NI Vision.
- L'information issue de la calibration des caméras et de la reconnaissance d'objets permet de retrouver la position spatiale des objets d'intérêts. Des procédures de calculs d'intersection de droites de projection sont développées par les étudiants sous Matlab et intégrées comme script dans LabVIEW. De même, l'utilisation de droites épipolaires permet d'illustrer la mise en correspondance d'entités entre images (figure 6).

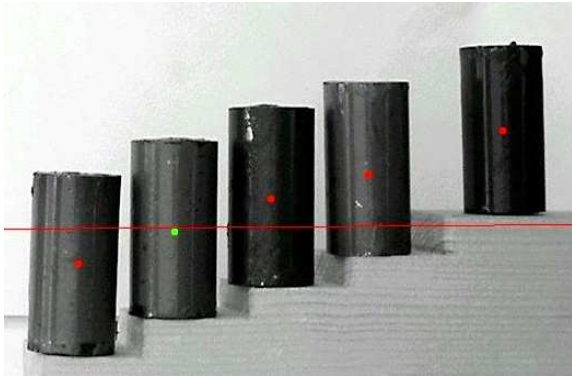


fig 6 : Mise en correspondance par droite épipolaire

2.4 Organisation de l'enseignement

Dans un premier temps, les étudiants suivent 6h de cours-TD. L'enseignement porte sur les différentes notions d'une chaîne d'acquisitions, de traitements d'images et de vision : constitution et information contenue dans une image (pixel, valeurs, plans couleurs – RGB, HSL, etc.), notion de convolution, restauration, segmentation, interprétation, calibration 2D et 3D, stéréovision et mise en correspondance. Ces différentes notions sont présentées de manière assez qualitative, l'objectif étant que les étudiants comprennent les notions des différents traitements sans toutefois rentrer dans un formalisme trop théorique. Par exemple, pour la calibration plane, seule l'intérêt et le résultat de la calibration sont présentés sans entrer dans les détails (modèle de projection, distorsion, interpolation, etc.). Ces différentes notions sont abondamment illustrées afin que les étudiants perçoivent au mieux la problématique et l'effet de chaque type de traitement.

En parallèle les étudiants suivent une série de 4 séances de 3 heures de Travaux Pratiques sur le matériel pré-

senté dans les paragraphes précédents. Le contenu de ces séances est assez progressif pour que les étudiants, d'une part, apprennent les notions de base et pour que, dans un second temps, ils réinvestissent ces connaissances dans un schéma plus général. Le contenu de ces séances suit la progression suivante² :

- Séance 1 : Acquisition, restauration et amélioration des images. Durant cette séance, les étudiants apprennent les différents modes d'acquisition (monocoup, séquence, continu), la notion d'espace de couleurs (RGB et HSL), quelques notions d'analyse d'images (profil, histogrammes, etc.), de restauration d'images (ils mettent en œuvre des filtres moyenneurs, Gaussiens et médians) et de amélioration d'images (fenêtrage, fonction de transfert, égalisation d'histogrammes). Durant cette séance, les étudiants utilisent soit des routines toutes faites, soit partent de la documentation de Labview et associent des routines de la librairie graphique. Les traitements sont appliqués sur les images acquises par les caméras ou sur des images représentatives d'un problème particulier. Par exemple une image d'un tableau de Rembrandt est utilisée afin de « rendre plus visible » l'information cachée par le clair-obscur.
- Séance 2 : Segmentation reconnaissance de formes. Cette séance n'a pas pour vocation de présenter les différentes classes de méthodes existantes. Il est demandé aux étudiants de proposer une stratégie de segmentation et de reconnaissance de formes directement liées à notre problématique finale, c'est-à-dire l'extraction de la position des centres des cylindres sur l'image acquise par la caméra noir et blanc et l'extraction de la position du seul centre du cylindre vert sur l'image acquise par la caméra couleur (figure 1). Les étudiants utilisent des notions de choix de plans de couleur, de choix de seuil de binarisation à l'aide d'histogrammes, d'utilisation d'outils de reconnaissance de formes (des rectangles dans notre cas particulier) et d'exploitation des résultats de cette reconnaissance. À la fin de cette séance, les étudiants ont tous élaborés une routine d'extraction des centres de cylindres à partir des images acquises par les deux caméras. La notion de pattern recognition à partir de templates et également illustrée lors de cette séance. Les étudiants définissent un template sur une image du tableau « jour et nuit » de Escher (figure 7). Ce tableau permet d'illustrer la notion de degré de ressemblance entre le template et l'information recherchée dans l'image.

² Les sujets des différents TP sont disponibles sur le site : http://perso.univ-rennes1.fr/jean-louis.dillenseger/home_enseignement.htm

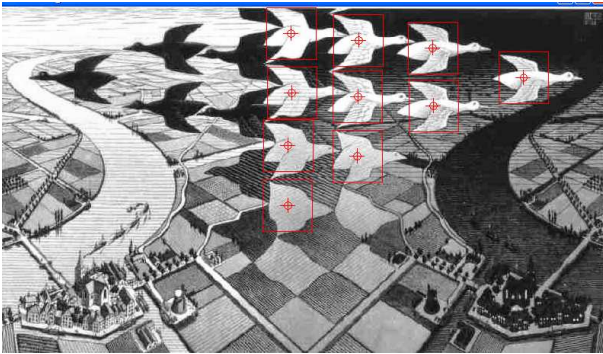


fig 7 : Recherche de motifs

- Séance 3 : Calibration de caméras. Les notions de calibrations plane et 3D sont abordées durant cette séance.

Calibration plane. Les étudiants font l'acquisition d'une image de la grille de calibration (figure 4), entrent les paramètres réelles de la grille (seuil de détection, origine et géométrie réelle de la grille) et effectuent la calibration plane à l'aide d'un VI de la librairie graphique de Labview [2]. Dans un second temps, les étudiants chargent les paramètres de calibration et les vérifient en cliquant dans l'image sur des points dont les coordonnées réelles sont connues (les croix dans l'image de la grille).

Calibration 3D. Les étudiants suivent le protocole décrit au paragraphe 2.3 : calibration plane de deux plans, l'un à $z=z_1$, l'autre à $z=z_2$. Le travail des étudiants consiste à effectuer la calibration plane pour deux positions z_1 et z_2 de la grille de calibration. Cette calibration multi-plans est ensuite vérifiée à l'aide d'une troisième position z_3 du plan de calibration. Lorsque sur cette image, l'étudiant clique sur une des croix de vérification, la calibration multi-plans nous donnent deux points dans l'espace $P_1(x_1, y_1, z_1)$ et $P_2(x_2, y_2, z_2)$ qui définissent la droite de projection. Cette droite de projection passera forcément par les coordonnées réelles de la croix.

- Séance 4 : Localisation 3D. Durant cette séance, les étudiants assemblent les parties développées précédemment : calibration 3D des deux caméras, reconnaissance de formes et localisation 3D. Pour une paire d'images, les étudiants sont alors capables d'extraire les droites de projections des centres de cylindres. L'intégration de routines développées par l'enseignant sous Matlab leur permet de mettre en correspondance les cylindres et d'estimer l'intersection entre les droites de projection et donc la position 3D du cylindre cible.

Durant ces séances, les étudiants sont en binômes et de surcroît les enseignants incitent les différents groupes à

communiquer entre eux, ceci pour favoriser l'échange de savoir et de savoir faire. Généralement les étudiants arrivent au bout de ces quatre séances. L'évaluation porte d'une part sur leur implication durant les Travaux Pratiques et, d'autre part, sur un devoir plus formel où il est demandé à l'étudiant de proposer une stratégie de traitements à une étude de cas.

3 CONCLUSION

L'enseignement suivi par les étudiants leur a permis de comprendre les différents éléments d'une chaîne de traitements en vision par ordinateur. L'utilisation d'un langage de programmation graphique comme Labview associé à une bibliothèque de routine de Vision permet aux étudiants d'appréhender des notions assez complexes comme la calibration ou la reconstruction 3D sans qu'il soit nécessaire d'en maîtriser les aspects mathématiques ou formels.

Bibliographie

- [1] H. A. Martins, J. R. Birk and R. B. Kelley, "Camera models based on data from two calibration planes", *Computer Graphics and Image Processing* 17, 2, 1981, pp. 173-180.
- [2] NI Vision Concepts Manual, NI Vision Concepts Manual.