Sens visuel pour robots industriels

Autor(en): **Dessimoz, J.-D.**

Objekttyp: **Article**

Zeitschrift: **Bulletin des Schweizerischen Elektrotechnischen Vereins, des Verbandes Schweizerischer Elektrizitätsunternehmen = Bulletin de l'Association Suisse des Electriciens, de l'Association des Entreprises électriques suisses**

Band (Jahr): **74 (1983)**

Heft 9

PDF erstellt am: **23.05.2024**

Persistenter Link: https://doi.org/10.5169/seals-904793

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.
Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.
Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Ein Dienst der *ETH-Bibliothek*
ETH Zürich, Rämistrasse 101, 8092 Zürich, Schweiz, www.library.ethz.ch

http://www.e-periodica.ch

Sens visuel pour robots industriels

J.-D. Dessimoz

Le sens visuel, basé sur l'usage de capteurs d'images à tube ou à semi-conducteur, permet le contrôle de robots industriels dont l'environnement n'est pas déterministe. Son emploi est particulièrement judicieux dans les applications de manutention de pièces et d'outils, puisqu'il est capable à distance de percevoir les grandeurs géométriques. L'article présente quelques algorithmes pour l'acquisition de pièces en vrac, pour l'estimation de position et d'orientation d'objets, pour l'évaluation de distance et pour le contrôle de qualité.

Röhren oder Halbleitersensoren erlauben es, Roboter mit Sehvermögen auszustatten, das die Überwachung und Steuerung von Industrierobotern ermöglicht, deren Umwelt nicht genau festgelegt ist. Speziell geeignet ist der Einsatz derartiger Roboter für die Handhabung von Bestandteilen und Werkzeugen, da sie geometrische Abmessungen auf Distanz wahrnehmen können. Der Aufsatz zeigt einige Algorithmen zur Ergreifung von Rohteilen, für die Feststellung des Standortes und der Orientierung von Gegenständen, für die Abschätzung des Abstandes und für die Qualitätskontrolle.

L'essentiel de cet article a été présenté lors de la conférence à la Journée d'information ASE «Les capteurs – base de l'automatique de demain», le 8 mars 1983 à l'EPFL Ecublens.

Adresse de l'auteur
Jean-Daniel Dessimoz, Ecole d'ingénieurs de l'Etat de Vaud EINEV, 1401 Yverdon.

Introduction

Les robots doivent, par définition, s'adapter à une grande variété de tâches. Mais jusqu'à présent, l'absence de capteurs extéroceptifs a présenté un obstacle majeur à leur emploi généralisé. Un robot sans capteur – c'est une situation actuellement très courante dans l'industrie – a besoin d'un environnement parfaitement déterministe pour pouvoir travailler. Ainsi les objets à manipuler doivent lui être fournis avec une grande précision de positionnement. Toute situation imprévue, telle que l'alimentation de fausses pièces ou la présence d'obstacles dans le champ de travail, peut conduire à des détériorations de l'outillage, voire à des accidents. L'organisation d'une place de travail pour un robot sans capteur s'avère généralement coûteuse : on cite souvent le rapport un tiers entre le prix du robot (bras et contrôleur) et celui de l'installation en fonctionnement qui comprend de plus les «accessoires» qui entourent le robot (par exemple : magasins, organe de préhension et bols vibreurs pour l'alimentation de pièces).

La première section de cet article présente le capteur d'images et ses particularités par rapport aux autres capteurs des robots. Elle situe également la vision artificielle pour robots industriels parmi les disciplines qui ont en commun le traitement d'images. Dans la deuxième section, le sens visuel d'un robot est étudié dans le contexte de la manutention d'objets.

1. Capteurs et sens visuel

Les capteurs disponibles pour un robot sont de types variés. Ils perçoivent le contact, la proximité, les forces, les moments, les images et donc les formes et les distances. Le sens visuel des robots repose sur l'utilisation de capteurs d'images. Ceux-ci ont en propre les caractéristiques suivantes:

– Très grand débit d'information. Un débit d'information de l'ordre de 100 000 bit/s est courant pour les applications en vision artificielle. Ceci dépasse largement tous les autres capteurs utilisés en robotique. Ainsi, les capteurs de contact ont un débit allant de 10 (interrupteur) à quelques centaines de bit/s (peau artificielle basée sur l'emploi de caoutchouc à résistance électrique dépendant de la pression [1;2] ou de capteurs de pression intégrés [3]). Les capteurs de forces et de moments délivrent de l'ordre de 1000 bit/s; cette information est généralement distribuée selon six composantes. D'autre part, plusieurs capteurs de proximité peuvent être montés sur la main des robots. L'amplitude des signaux de proximité reçus n'est généralement codée que de manière binaire (présence ou absence d'objets), car elle est influencé par de nombreux facteurs qu'il est difficile de démêler (réflectance, distance, angle des surfaces de réflexion . . .). La fréquence maximale de ces signaux se confond avec celle des déplacements mécaniques. Les capteurs de proximité peuvent être capacitifs, inductifs, à ultrasons ou électromagnétiques (lumière visible ou infrarouge). Leur débit d'information est de l'ordre de 100 bit/s.

– Perception à distance de la géométrie d'une scène (indirectement, c'est-à-dire par interprétation de la réflectance des objets). Ceci est avantageux, car une optimisation des mouvements du robot peut en résulter. Par contraste, la plupart des autres capteurs nécessitent le déplacement du bras de robot pour explorer l'environnement, ce qui est très lent. Bien que les capteurs à ultrason puissent aussi estimer la géométrie des objets à distance, ils sont peu utilisés car ils n'ont, dans l'air, qu'une résolution très grossière[4]. Les ultrasons sont rapidement absorbés lorsque leur fréquence dépasse 400 kHz. La résolution est approximativement gouvernée par l'équation $d = \lambda/2$, où λ est

la longeur d'onde de la porteuse. Les distances entre capteurs et objets sont estimées par différents moyens indirects. Une méthode parfois utilisée est basée sur la focalisation de l'image. Lorsque l'image est nette, la distance entre l'objet et le capteur est proportionnelle à la distance focale. Beaucoup plus utilisée, la triangulation se base sur la propriété que les six éléments d'un triangle (angles et côtés) sont fixés si un côté et deux angles sont connus. Le triangle est formé par deux capteurs et l'objet (vision stéréoscopique) ou par un capteur, l'objet et une source lumineuse directionnelle. Ce point est développé plus loin.

- Les limitations d'utilisation des capteurs visuels sont liés à l'impossibilité de percevoir les grandeurs dynamiques (forces et moments). Même les grandeurs géométriques ne sont pas toujours perceptibles. Par exemple, la réflectance d'une scène peut être trop uniforme ou, au contraire, trop variée pour permettre l'analyse. Des obstacles peuvent aussi se trouver entre capteur et objet. Typiquement, durant la phase terminale des mouvements du robot, la main cache une partie de la zone de travail.

1.1 Capteurs d'images à tube et à semi-conducteur

Le sens visuel repose sur l'utilisation de capteurs qui se classent en deux catégories: caméras à tube et caméras à semi-conducteur.

Les caméras à tube ont en commun une bonne résolution. On peut typiquement distinguer jusqu'à 300 lignes noires sur fond blanc dans une image. Ainsi, on acquiert aisément 512 x 512 points significatifs par image (c'est-à-dire 512 lignes et 512 colonnes). Ces caméras ont aussi l'avantage d'être compatibles avec la majorité des équipements vidéo grand public (moniteurs, etc.). Une abondance de tubes à caractéristique de transfert spécifiques sont disponibles. Les défauts sont principalement dans les distorsions géométriques de l'image (typiquement 2%) et dans la cadence d'acquisition. Celle-ci est de 25 images par seconde pour la pleine résolution et de 50 images par seconde si l'on supprime l'entrelaçage.

La seconde catégorie de capteurs d'images regroupe les caméras à semi-conducteur. Lorsque le capteur est bidimensionnel, elles ont une résolution bien moindre que les caméras à tube. Elles sont disponibles en 32×32 éléments, mais les plus populaires ont 128×128 points (p.ex. General Electric) ou 256×256 environ (p.ex. Hitachi, NEC ou RCA). Peut-être parce que le marché potentiel est trop petit, NEC et RCA ne livrent plus ce type de caméra. Par contre, on note une entrée en force des capteurs à semi-conducteur dans la vidéo pour amateur. Le nombre de cellules est assez grand, mais les cellules du capteur sont recouvertes d'un filtre rouge, vert ou bleu.

Beaucoup de capteurs à semi-conducteur sont linéaires. Les plus denses comptent jusqu'à 4000 cellules (Fairchild, Reticon). Ces capteurs sont particulièrement intéressants lorsque les objets bougent, car ceux-ci assurent le balayage dans une direction et, de plus, le temps d'intégration court propre aux capteurs linéaires est alors un avantage. Pratiquement exemptes de distorsions géométriques, les caméras à semi-conducteurs s'avèrent aussi plus sensibles, légères et compactes que les caméras à tube. De plus, elles sont moins sujettes à l'éblouissement («blooming») et se satisfont de tensions d'alimentations basses. Elles sont cependant encore relativement chères.

1.2 Contexte du sens visuel des robots

Ce paragraphe illustre les disciplines qui côtoient le sens visuel des robots, par quelques réalisations (Fig. 1). Les capteurs visuels délivrent des images qui doivent être traitées. En traitement d'image, on peut distingueur deux catégories principales: la vision artificielle et le graphique. La première catégorie traite de l'analyse et de la compréhension de scènes par des machines. La deuxième s'occupe de synthèse et de traitement d'images pour des observateurs humains. A ce deuxième type de technique correspond la majorité de l'équipement matériel existant. Les systèmes les plus performants, tels ceux de COMTAL ou de VICOM, traitent simultanément, en temps réel, plusieurs images de 4096×4096 points (p.ex. générées par LANDSAT). Le traitement peut inclure des agrandissements, des rotations, des filtrages et des changements de couleurs.

En vision artificielle, de nombreuses applications se démarquent: inspection, classification, surveillance, contrôle de robots, etc. Pour la surveillance, des systèmes commerciaux (Octek aux USA) détectent le mouvement; à l'Université de Karlsruhe, un système visuel surveille la zone de travail d'un robot pour prévenir les accidents. En classification, il existe, par exemple des équipements de reconnaissance de caractères pour aveugles, des systèmes pour tri postal ou des machines à compter les différentes sortes de globules blancs du sang. En inspection, le domaine des applications industrielles potentielles semble énorme. Exemples d'équipement très performant: une machine inspecte les soudures de circuits imprimés chez Hitachi, à la cadence d'un circuit format double-Europe par seconde; une autre contrôle l'apparence de médicaments chez Fujitsu à la vitesse de vingt capsules par seconde.

Spécifiquement pour le contrôle des robots, on peut mentionner les applications en soudure à l'arc, en exploration et en manutention. Pour la soudure, des firmes (Unimation, Kawasaki, Automatix, etc.) commercialisent des systèmes de poursuite de joints en

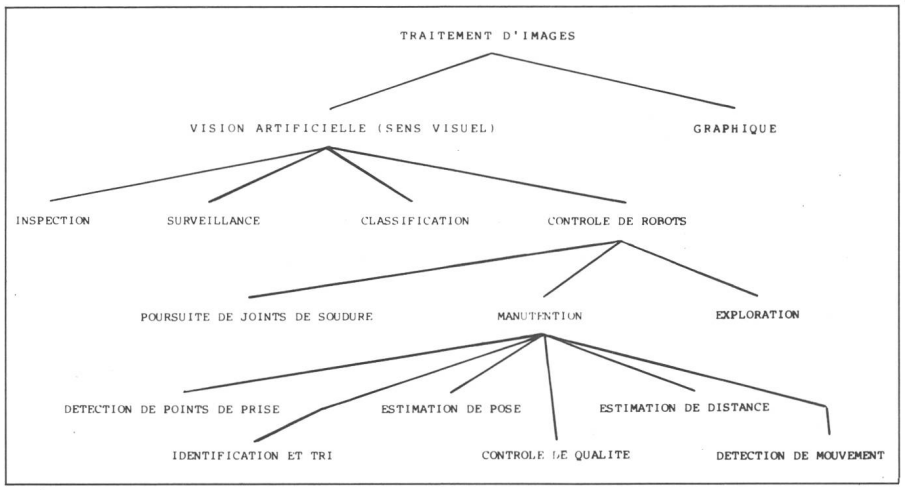

Fig. 1 Quelques domaines voisinant la vision artificielle pour robots industriels

deux passages (d'abord la reconnaissance visuelle, puis la soudure). A l'«Institut für Produktionstechnik und Automatisierung» de Stuttgart, un système spécialisé a été développé, qui observe le métal en fusion et contrôle les déplacements de la mèche en conséquence. Un sens visuel pour robots soudeurs est très demandé, car ces robots sont déjà installés en grand nombre dans les ateliers industriels. Pour l'exploration, le sens visuel est fondamental: les robots mobiles doivent optimiser leur trajectoire à travers des environnements souvent imprévisibles. Entre autres, le Jet Propulsion Laboratory aux USA et le LAAS au «Centre national de recherches scientifiques» (CNRS) de Toulouse consacrent d'importantes ressources à la recherche dans ce domaine. Le sens visuel doit être particulièrement performant en termes d'abstraction et d'intelligence pour ce type d'application.

Quant à la manutention d'objets, c'est le terrain idéal pour les robots industriels dotés de vision, en atelier. La perception de la géométrie de la scène est alors déterminante. Lorsque, par contre, le robot touche des éléments rigidement fixés (assemblage, usinage, etc.), les grandeurs dynamiques, de contact ou de proximité sont plus indiquées pour le contrôle du robot. Localisation de but, prise d'objets, estimation de pose (position et orientation), identification et tri de pièces sont des tâches typiquement dévolues au sens visuel durant la manutention. De plus, un contrôle de qualité, au moins grossier, est souvent nécessaire simultanément.

2. Manutention d'objets

Les opérations de manutention de pièces se déroulent schématiquement en trois phases: acquisition, transport et positionnement.

Durant l'acquisition, le type et la pose des objets à prendre sont parfois parfaitement déterministes. Cette situation est particulièrement favorable au robot et le dispense d'avoir un sens visuel. Cependant, il n'est pas toujours économique d'organiser cet environnement (coûts de stockage et de transport des pièces orientées, magasins spéciaux, équipements de mise en ordre après ébavurage, p.ex.). Parfois, les pièces ont une pose aléatoire, mais elles sont présentées une à une. C'est souvent le cas lorsque les pièces arrivent sur une bande transporteuse. Le robot peut alors en estimer la pose et les acquérir en conséquence. Mais, il y a dans l'atelier, beaucoup de pièces stockées en vrac. Traditionellement, c'est de la main d'œuvre ou des bols vibreurs qui les manipulent pour en alimenter des machines, des magasins ou des supports divers. Dans bien des cas, il apparaît maintenant possible de céder cette tâche à un robot doté de vision. Le sens visuel se limite alors à la détection de points de prise, laissant l'estimation de pose à plus tard, lorsque la pièce est présentée de manière plus adéquate au capteur d'images.

Durant le transport, l'objet peut bouger dans l'organe de préhension et une estimation de pose peut être à nouveau requise. A l'opposé, l'incertitude de pose peut être restreinte si des contraintes mécaniques sont introduites. Ceci arrive, par exemple, lorsqu'une pièce doit être réorientée, voire changée de main pour pouvoir être déposée au suite désiré.

Le positionnement, aussi, a parfois recours aus sens visuel. Ainsi, le site où la pièce doit être déposée peut ne pas être localisé a priori. Une estimation de pose du but est alors nécessaire. Il arrive même que le but bouge et qu'ainsi le robot doivent procéder à une estimation de mouvement.

Lorsqu'on analyse une place de travail, on est souvent surpris de la diversité des contributions faites par l'opérateur. Ainsi, il y a presque toujours un certain contrôle de qualité (inspection) qui est accessoire à la tâche principale. Si l'on veut remplacer l'opérateur, il est alors judicieux de prévoir un substitut à toutes les fonctions essentielles.

2.1 Acquisition de pièces en vrac

Bien que l'on s'efforce de maintenir la position et l'orientation des pièces à travers les divers postes d'une usine, il arrive très souvent que le désordre s'installe. Parfois, pour des raisons technologiques (p.ex. pour l'opération de sablage), mais généralement pour des raisons économiques les pièces sont simplement mises en vrac dans des caisses. Les méthodes courantes pour l'utilisation de pièces ainsi entreposées se divisent principalement en deux classes: dans l'une, de la main d'œuvre prend les pièces à la main et en alimente les machines ou des magasins-tampons; dans l'autre, des bols vibreurs trient et orientent les pièces (fig. 2).

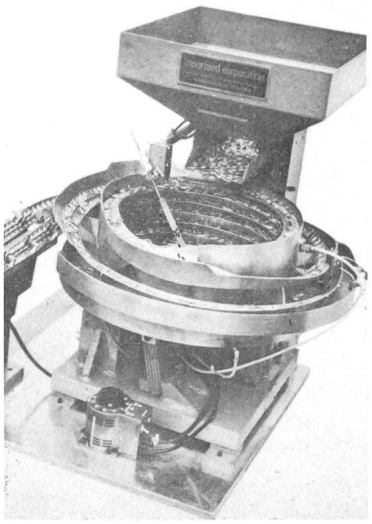

Fig. 2 Bol vibreur pour le positionnement de pièces stockées en vrac

On commence aujourd'hui à reconnaître l'acquisition de pièces en vrac comme une tâche importante pour les robots. Deux raisons à cela: d'une part, l'acquisition automatique de pièces apparaît souvent comme le dernier maillon à automatiser le long d'un processus de fabrication. Ainsi, dans plusieurs pays, on a développé des ateliers dits flexibles qui peuvent usiner et assembler des pièces manuellement préparées sur des plateaux.

D'autre part, des instituts de recherches ont prouvé la faisabilité de cette automatisation par des moyens robotisés [5, 6]. Pour être soluble, le problème de la prise de pièces en vrac doit se formuler sous une forme élémentaire: «Etant donné un organe de préhension, où dans un tas de pièces, est-il possible de prendre quelque chose?» Ce qui est fondamental, c'est d'identifier un point de prise. En conséquence, l'estimation d'orientation et l'identification de la pièce prise sont reportées à une phase ultérieure, après que la pièce acquise soit séparée des autres par le robot et que sa pose soit éventuellement mécaniquement limitée à quelques degrés d'incertitude [5].

Avec les méthodes publiées à ce jour, la probabilité d'acquisition d'une pièce (inconnue) d'un tas atteint souvent 90% après trois essais. Une des premières méthodes utilisées permet la détection de points de prise pour une ventouse (fig. 3). Elle se satisfait d'une représentation booléenne à 128×128 points par image traitée. A la suite d'un prétraitement qui segmente partiellement l'image et qui réduit la quantité de points à traiter, des zones

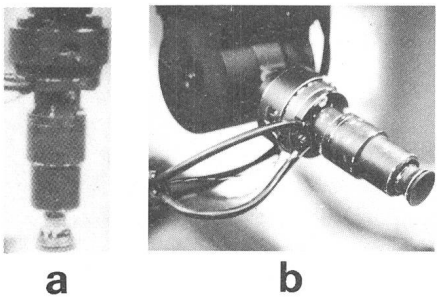

Fig. 3 Organe de préhension au moyen d'une ventouse

En (a), la ventouse peut s'accomoder à l'orientation de la surface de prise dans un domaine de 120°. En (b), la ventouse est rétractée pour forcer la surface de prise dans un plan perpendiculaire à son axe. La main est équipée d'un capteur de contact.

uniformes, libres de bords sont détectées. Les plus grandes zones sont alors désignées comme points de prise possible (fig. 4). Des expériences ont donné d'excellents résultats pour des pièces aussi variées que des cylindres de tailles diverses (acier), des pales de turbines pour avions à réaction (titanium forgé), des boîtes de raccordement (aluminium) et des stylos feutres (plastique).

Pour une autre classe d'applications où la surface (en termes de réflectance, de texture ou de forme) des pièces ne conduit pas à un bon contraste entre pièces superposées, une représentation à niveaux de gris multiples est nécessaire. Dans ce domaine, une technique basée sur l'usage de filtres adaptés s'est révélée intéressante. Une zone de préhension typique doit d'abord se définir. Celle-ci dépend à la fois des pièces

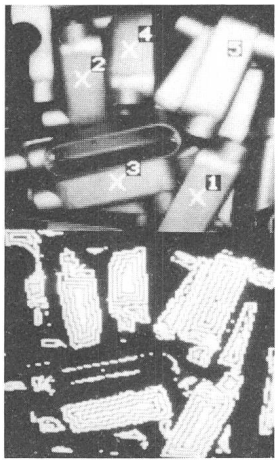

Fig. 4 Acquisition de pièces en vrac

En haut, cinq points de prise pour la main de la figure 3 ont été détectés et sont superposés à l'image numérisée. En bas, un détail du traitement est illustré.

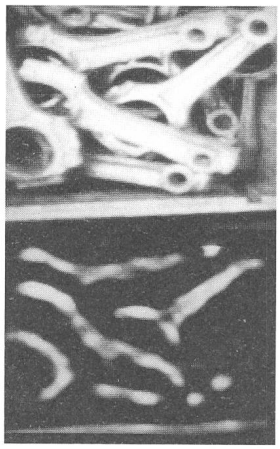

Fig. 5 Représentation à niveaux de gris multiples

En haut, l'image numérisée de pièces en vrac (128×128 points de 6 bit). En bas, l'image après filtrage adapté, l'intensité indique en chaque point la probabilité d'y avoir un bon point de prise, pour la pince de la figure 6.

à manipuler et de l'organe de préhension. Elle se définit naturellement en termes géométriques. Un modèle permet ensuite de passer de la représentation géométrique à une représentation lumineuse, telle que perçue sur les capteurs visuels. Ceci définit un filtre spatial (si possible récursif) qui s'applique numériquement à l'image. Les zones de préhension probable apparaissent alors comme des zones d'intensité élevée (fig. 5) et s'identifient donc par une simple détection de pics. Ainsi, par exemple, une pince à doigts parallèles (fig. 6) a des chances de prendre une pièce là où deux volumes vides entourent un volume occupé par un segment de pièce. Pour l'unité de vision artificielle, deux zones noires entourant une zone plus claire doivent se détecter.

2.2 Estimation de distance

La distance entre un objet et un capteur peut être estimée de nombreuses

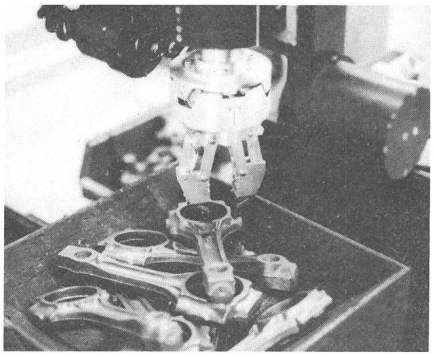

Fig. 6 Acquisition de bielles, avec l'algorithme illustré en figure 5

manières. La plus intuitive se base sur le temps de propagation d'une onde. Une onde acoustique est adéquate si l'objet n'est pas localisé avec une grande précision (cf. relation entre résolution et longueur d'onde). Cette méthode est utilisée avec succès en photographie amateur et peut être un complément utile au sens visuel des robots. Lorsque l'onde utilisée est électromagnétique (p.ex. rayon infrarouge), une bonne résolution spatiale est liée à des temps de propagation extrêmement courts, de l'ordre de la ns. L'équipement est donc compliqué, cher et, de plus, les mesures doivent être répétées («moyennées») un grand nombre de fois pour être significatives. Paradoxalement, ce type de mesure est donc lent. La littérature technique rapporte des cadences de l'ordre de 1 point par seconde (un facteur d'au moins 10^6 par rapport aux capteurs visuels) [7].

Les méthodes d'estimation de distance les plus courantes pour les objets proches sont basées sur la triangulation [6; 8]. On utilise un point lumineux qui balaie la scène. Si le point est visible, l'angle du rayon perçu sur le capteur, la distance entre le capteur et la source lumineuse et l'angle d'émission du rayon lumineux permettent de connaître la distance entre ce point et le capteur. En utilisant un laser et des filtres optiques appropriés, il est possible de faire des mesures rapides (10^4 points par seconde) en lumière ambiante même avec une puissance émise modeste de l'ordre du mW [8]. La limitation de la méthode vient des zones «d'ombres» où le spot lumineux n'est pas perceptible sous l'angle correspondant au capteur visuel. Un certain gain de vitesse est possible si l'on projette simultanément toute une ligne, voire toute une matrice de points au lieu de balayer l'objet à localiser par un seul point. Il faut toutefois noter que des ambiguités sont alors de plus en plus probables, entre points émis et points perçus.

Une méthode similaire mais qui se distingue par l'absence de lumière spécialement projetée sur l'objet est très utilisée en photogrammétrie: c'est la vision stéréoscopique. Les difficultés de mise en œuvre de cette technique en robotique sont liées à plusieurs facteurs. D'une part, une corrélation d'images est requise, qui est gourmande en ressources de calcul. De plus, cette corrélation doit se faire de manière locale («par morceaux»), car chaque zone d'image corrélée doit correspondre à tout ou partie d'un seul

objet. Enfin, on a des difficultés fondamentales:

1) éventualité de zones d'ombre où une zone perçue par un capteur ne l'est pas par l'autre;
2) nécessité d'avoir une texture suffisamment riche et aléatoire pour permettre la corrélation précise de toute zone visible;
3) sensibilité aux différences des deux capteurs utilisés en stéréoscopie et à l'éclairage dont la perception varie avec la position des capteurs.

Une dernière méthode a été proposée récemment, appelée radiométrie. Elle utilise le fait qu'un élément de surface d'objet réémet (la méthode est surtout intéressante pour des objets qui ne brillent pas, c'est-à-dire dont la surface diffuse la lumière) une partie de la lumière incidente non seulement proportionnelle au coefficient de réflectance, mais qui est également fonction de l'angle relatif entre la normale à l'élément de surface et les rayons de lumière incidents. Un capteur et trois sources lumineuses allumées tour à tour permettent ainsi de connaître l'orientation des surfaces visibles, ce qui est lié à la dérivée de la distance entre capteur et objet. Outre la quantité des calculs nécessaires, la méthode a le défaut de n'être applicable qu'à la zone linéaire du capteur. Celle-ci est limitée à moins de deux décades. D'autre part la radiométrie est similaire à l'interférométrie en ce qu'elle ne donne pas de valeur absolue, même grossière, de la distance.

2.3 Estimation de pose et identification d'objets

Au moment d'estimer la pose d'un objet, la distance de l'objet au capteur est, en général, approximativement connue. Ceci définit un facteur d'échelle entre l'objet dans l'espace et sa projection sur le capteur. Il sera pris égal à 1 dans cette discussion.

Une pose est définie comme un point dans un espace à six dimensions *P*, généré par toutes les translations et les orientations possibles d'un objet rigide. Un état, dans notre contexte, est défini comme l'une parmi plusieurs régions disjointes de l'espace *P*. Il est dit stable s'il est associé à une quantité d'énergie (localement) minimale. L'énergie est liée à la présence d'un champ (en pratique, la gravité ou un champ magnétique) et à des forces et moments transmis mécaniquement. Les problèmes d'estimation de pose sont caractérisés par la dimension de l'espace des poses possibles (EPP) correspondant. Par exemple, on définira comme un problème d'estimation de pose monodimensionnel, l'évaluation de la position d'une pièce qui ne peut bouger que le long d'un axe. On dira également que l'EPP est de dimension 1.

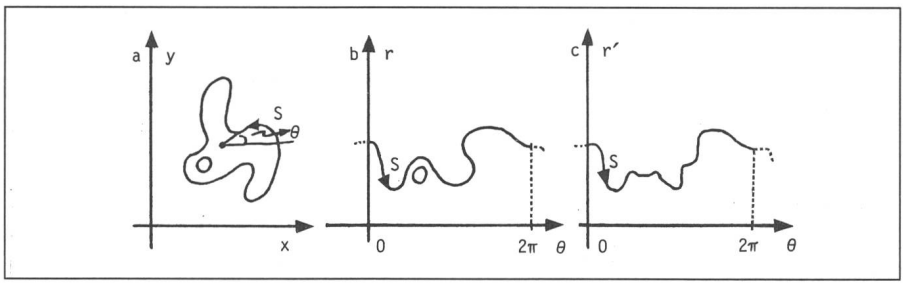

Fig. 7 Codage polaire d'une silhouette
Seuls les bords sont informatifs

L'estimation de pose varie largement en complexité en fonction de la multiplicité des objets dans le champ de vue, de leur superposition éventuelle, de la rigidité des objets ou de limites mécaniques qui peuvent restreindre leur EPP. Dans les applications industrielles, il apparaît que les pièces sont généralement analysées une-à-une, reposant sur une surface plane.

Lorsqu'un objet se trouve isolé sur une surface plane, il peut généralement se décrire par sa silhouette. Dans ce contexte, l'EPP est tridimensionnel: deux translations et une rotation dans le plan de travail. Un bon contraste peut, en général, s'obtenir et des solutions utilisant une représentation binaire de l'intensité lumineuse sont bien établies. Dans ce cas simple, des produits commerciaux pour l'estimation de pose sont disponibles [9]. Les techniques les plus puissantes sont basées sur un codage polaire [10; 11] (fig. 7). Elles sont rapides (environ 0,1 s par objet) lorsqu'implantées sur du matériel spécialisé. Une autre classe de solutions est particulièrement populaire pour réalisation logicielle. Celle-ci regroupe les méthodes d'estimation de moments d'inertie de l'image [12]. L'angle de l'axe d'inertie minimum definit l'orientation de la pièce observée dans un plan perpendiculaire à l'axe du capteur (fig. 8). Ces méthodes ont deux aspects particulièrement défavorables: l'axe d'inertie minimum peut exister en nombre infini (c'est, p.ex. le cas pour un carré) et il demeure une ambiguité de 180 degrés, qui doit être levée par des heuristiques.

Si les pièces reposent sur un plan mais sont partiellement cachées (superposition) des méthodes basées sur l'utilisation de caractéristiques locales et, en particulier, de la courbure du contour sont possibles [13] (fig. 9). Dans la mesure où les pièces ne sont visibles que partiellement, un risque d'erreur sur l'identification des pièces existe, qui est inévitable. Il n'existe pas de solution pratique pour les situations où l'EPP est supérieur à trois. Une difficulté fondamentale vient peut-être du fait que les capteurs visuels perçoi-

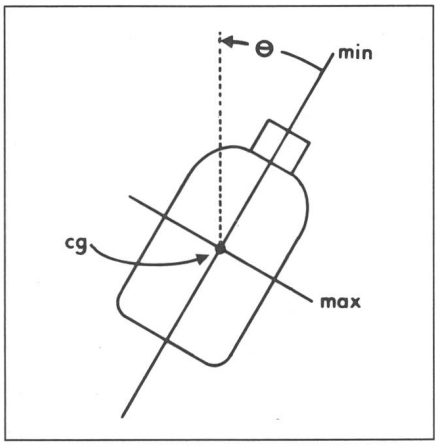

Fig. 8 Les moments d'inertie de l'image peuvent donner la position (cg) et l'orientation (axes de maximum ou de minimum d'inertie) de l'objet vu

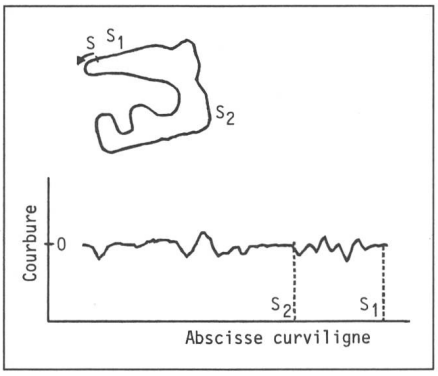

Fig. 9 Détermination de la pose d'une pièce au moyen de la courbure locale
La courbure est définie localement le long du contour. Elle est donc utile même si l'objet est partiellement caché.

vent une *surface* dans l'espace et non l'espace dans ses trois dimensions, comme en tomographie par exemple.

Le problème de base décrit peut aussi se compliquer énormément si la forme des pièces change, comme dans les opérations de forge, le démoulage ou lorsque les pièces ne sont pas rigides.

L'estimation de pose d'un objet n'est généralement pas indépendante de son identification. Si l'objet est perçu dans sa totalité (absence de recouvrement et d'ombre) des paramètres indépendants de la pose peuvent être définis, tels que la longueur du périmètre et la surface circonscrite par le périmètre (contour). Ces paramètres sont toutefois peu nombreux et modérément précis [14]. Ils ne permettent donc de différencier que quelques pièces de types très différents. Pour l'estimation de pose, par contre, un plus grand nombre de paramètres sont généralement requis. Ils peuvent alors aussi s'utiliser pour l'identification de pièces même similaires. Si une pièce n'est pas totalement visible, l'interdépendance est encore plus nette puisque c'est la position relative de détails identifiés (structures locales telles que coins, fentes ou trous) qui permettent l'estimation de pose [13; 14]. Dans ce contexte, la courbure du contour qui est une fonction invariante à la translation et à la rotation des objets est beaucoup utilisée.

Si un objet a une pose connue comme, par exemple, une pièce rigidement fixée sur son support, la tâche d'identification est alors considérée comme faisant partie du travail d'inspection.

2.4 Inspection

La tâche d'inspection se caractérise par différents traits dont les plus importants sont les suivants: pose de l'objet inspecté connue (au moins grossièrement), grande quantité d'information perçue par le capteur et, en conséquence, traitement hiérarchisé, analyse par régions et algorithmes heuristiques. L'orientation et la position connues a priori, cela permet de focaliser localement l'inspection. Ainsi une résolution poussée dans les quelques (petites) zones adéquates de l'image reste compatible avec un temps d'analyse admissible. Si la pose de l'objet n'est pas connue, l'échantillonnage doit être uniforme pour toute l'image, et une étape d'estimation de pose précède l'inspection.

Il apparaît que les applications en inspection nécessitent généralement une haute résolution et un vaste champ de vue (en grandeur relative). Par exemple, une pièce de quelques cm doit être mesurée au 1/100 mm, des rayures de l'ordre du µm doivent être détectées sur un cristal de quelques mm ou encore des trous de quelques 1/10 mm doivent être reconnus sur des pistons de 10 à 20 cm de diamètre. Il existe des exemples similaires dans l'industrie textile, électronique (circuits imprimés), etc.

Vu la quantité d'information perçue par le capteur, il est vital, en inspection, de hiérarchiser le traitement. Ainsi, dans une application industrielle d'inspection d'étiquettes, une première phase estime l'écart entre la position effective de l'étiquette et sa position nominale. Pour cela, l'état d'un bord de l'image perçue est analysé. Pour la suite du traitement, l'image perçue est déplacée pour correspondre exactement à une image de référence interne. Dans une autre application, la lecture de quelques colonnes de l'image suffit à affiner l'estimation de pose d'une dent se scie en carbure de tungstène (fig. 10). Par la suite, une fraction seulement de l'image à 256×256 points est analysée pour détecter des défauts de bord (bords cassés); en cas de test positif, une dernière zone de l'image est traitée pour vérification de l'état de surface.

Etant donné la variété des tâches possibles en inspection et le temps d'analyse limité, il n'existe pas de solution générale. Il est nécessaire de définir de cas en cas la stratégie adéquate. La communication entre opérateur et système d'analyse doit donc être particulièrement soignée. On définit interactivement des zones d'intérêt dans l'image (fenêtres), petites et peu nombreuses. Pour chaque fenêtre, une résolution aussi grossière que possible doit être choisie, pour chacun des deux axes du capteur indépendamment. A l'intérieur des fenêtres, le traitement le plus adéquat doit aussi être défini. Ce sera au besoin une estimation de dimension, de surface, de niveau de réflectance, éventuellement une estimation de pose ou encore d'autres paramètres.

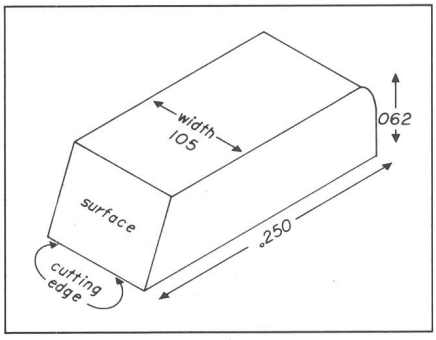

Fig. 10 Inspection d'une dent en carbure de tungstène

D'abord la largeur doit être vérifiée (dimension), puis le bord tranchant (cassures éventuelles) et, en cas de tests positifs, l'état de la surface de coupe.

La matériel idéal pour l'inspection comprend, en particulier, un capteur à haute résolution (à tube si le capteur est bidimensionnel ou à semiconducteur s'il est linéaire). Pour une grande vitesse de traitement, une unité câblée de prétraitement est nécessaire, qui soit capable de gérer des fenêtres (segments d'image), et à l'intérieur de chacune d'elles, de faire du filtrage, des rotations et des changements de résolution (interpolation et rééchantillonnage) programmables.

3. Conclusions

La vision artificielle a commencé récemment à permettre l'automation rentable de certains processus industriels. Des algorithmes ont été évoqués pour l'acquisition de pièces en vrac, pour l'estimation de pose d'objets, pour l'évaluation de distance et pour le contrôle de qualité. Dans bien des situations cependant, la vision artificielle n'offre pas encore de solution viable. Un obstacle est la quantité d'opérations qu'elle requiert. Celles-ci impliquent souvent une grande complexité du matériel nécessaire ou alors, si l'essentiel est fait par logiciel, beaucoup de temps de calcul. D'autres limitations viennent du manque de généralité des algorithmes existants, du défaut de modèle géométrique adéquat, et du peu d'expérience disponible à l'intérieur des entreprises, en vision artificielle.

Lorsqu'on allie le sens visuel à un robot, des caractéristiques spécifiques apparaissent, qui sont plus ou moins favorables [15]. Parmi les avantages, il faut noter la possibilité d'intervenir activement dans la scène, par exemple en montrant sous un angle adaptatif un objet à la caméra [16]; possibilité également de contraindre mécaniquement les objets de manière à limiter l'incertitude de leur pose; enfin, des capteurs complémentaires de contact et de proximité montés sur l'organe de préhension peuvent contrôler le robot durant la phase finale des mouvements,

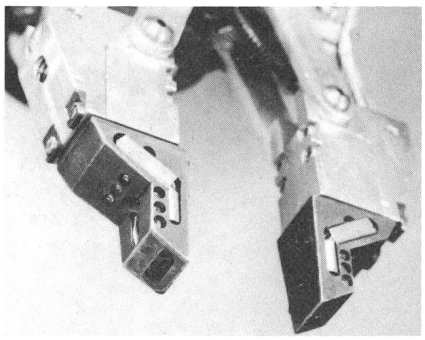

Fig. 11 Pince à ouverture programmable
Pour la phase terminale des mouvements, cette main possède un capteur de contact et neuf capteurs de proximité à infrarouge. Elle permet une correction fine (dx, dy, dz) au point (x, y, z) choisi par le sens visuel.

ce que la vision fait très difficilement (fig. 11).

Malheureusement, des difficultés nouvelles apparaissent aussi lorsqu'on couple la vision artificielle au contrôle d'un robot. En particulier, le problème qualifié de «cinématique inverse» est difficile à résoudre: la position et l'orientation d'un objet dans l'espace étant connues (par exemple, à l'aide du sens visuel), quelles consignes doit-on donner aux *n* articulations du robot pour qu'il y accède (typiquement, *n* vaut 6). Ce problème d'apparance anodine n'est pas toujours soluble en pratique, que ce soit à cause de la structure du robot, de difficultés de calibration ou du temps de calcul élevé; c'est pourquoi la majorité des robots industriels actuels se programment dans l'espace des variables d'articulations et non dans un référentiel cartésien (pour la position) avec, par exemple, les angles d'Euler pour l'orientation.

La vision artificielle permet d'adapter les robots industriels à certaines variations de leur environnement. Les temps de calcul liés à la vision et au contrôle, qui sont typiquement de l'ordre de la seconde, permettent de planifier des trajectoires adaptatives. Par contre, ces temps sont prohibitifs s'il s'agit de contrôler le robot le long de la trajectoire voulue. Ainsi, un vieux rêve de roboticien n'est toujours pas réalisable; Contrôler par la vision une structure mécanique peu rigide, pour en faire un robot de haute précision!

Bibliographie

[1] *R. Bajcsy* a.o.: What can we learn from one finger experiments? Philadelphia, University of Pennsylvania/Department for Computer and Informations Science, 1982.

[2] *J. Clot* et *J. Falipou:* Réalisations d'orthèses pneumatiques modulaires. Etude d'un détecteur de pressions plantaires. Publication LAAS No. 1852. Toulouse, Laboratoire d'Automatique et d'Analyse des Systèmes du Centre National de la Recherche Scientifique (CNRS), 1978.

[3] *P.W. Barth:* Silicon sensors meet integrated circuits. IEEE Spectrum 19(1981)9, p. 33...39.

[4] *P. Kleinschmidt* and *V. Mágori:* Ultrasonic remote sensors for noncontact object detection. Siemens Forsch. und Entwickl. Ber. 10(1981)2, p. 110...118.

[5] *J. Birk, J.-D. Dessimoz and R. Kelley:* General Methods to enable robots with vision to acquire, orient and transport workspieces. 9th National Science Foundation Grantees Conference on Production Research and Technology, Ann Arbor, Michigan, 1981, p. I1...I8.

[6] *J.-D. Boissonnat* and *F. Germain:* A new approach to the problem of acquiring randomly oriented workspieces out of a bin. Proceedings of the 7th International Joint Conference on Artificial Intelligence, Vancouver/Canada, 1981; p. 796...802.

[7] *R.O. Duda, D. Nitzan* and *P. Barrett:* Use of range and reflectance data to find planar surface regions. IEEE Trans. PAMI 1(1979)3, p. 259...271.

[8] *S. Parthasarathy, J. Birk* and *J.-D. Dessimoz:* Laser rangefinder for robot control and inspection. Technical Symposium East '82, International Society for Optical Engineering (SPIE) and IEEE Computer Soc., Proc. Vol. 336 (Robot Vision), Washington D.C., May 6, 1982.

[9] *R. Karg* and *O.E. Lanz:* Experimental results with a versatile optoelectronic sensor in industrial applications. Proceedings of the 9th International Symposium on Industrial Robots, Washington D.C., 1979; p. 247...264.

[10] *H. Yoda, S. Ikeda* and *M. Ejiri:* A new attempt of selecting objects using a hand–eye system. Hitachi Review 22(1973)9, p. 362...365.

[11] *P. Kammenos:* Performances of polar coding for visual localisation of planar objects. Proceedings of the 8th International Symposium on Industrial Robots, Stuttgart/Germany, 1978; p. 143...154.

[12] *S. Dudani, K.J. Breeding,* and *R. McGhee:* Aircraft identification by moment invariants. IEEE Trans. C-26(1977)1, p. 39...46.

[13] *F. de Coulon* e.a.: Systèmes de vision artificielle pour la localisation de pièces industrielles et l'analyse de texture. En: Le Point en Robotique. Vol. I. Paris, Ed. Lavoisier, 1983, p. 84...104.

[14] *J.-D. Dessimoz:* Traitement de contours en reconnaissance de formes visuelles: application à la robotique. Thêse de doctorat N° 387, EPF Lausanne, 1980.

[15] *J.-D. Dessimoz* a.o.: Between bins and goalsites: Workspiece pose estimation. Proceedings of the 3rd International Conference on Assembly Automation, Stuttgart/Germany, 1982; p. 603...614.

[16] *R. Kelley* a.o.: Forging: Feasible robotics techniques. Proceedings of the 12th International Symposium on Industrial Robots, Paris, 1982; p. 59...66.