

Développement d'une routine de reconnaissance rapide d'objets dans une image pour applications industrielles

Ing. J. SMEYERS
ECAM – Bruxelles

L'automatisation des chaînes de production se caractérise par une absence de contrôle visuel de la qualité du produit. Une caméra et un traitement informatique adapté permettent de reconnaître des objets et d'en assurer le contrôle. Les méthodes utilisées sont nombreuses et dépendent de la géométrie du produit et de sa prise de vue. Cette étude propose une méthode de reconnaissance d'objets à caractère industriel basée sur l'algorithme de Harris.

Mots-clefs : reconnaissance d'objet, apparier, point d'intérêt, point caractéristique, référence, apprentissage, méthode de Harris

The automation of production line is characterized by the lack of visual quality control of the product. A camera and an adapted software treatment enable to recognize objects and therefore to ensure the control. The methods used to this end are various and depend on the product geometry and it's angle's view. This study proposes an industrial object recognition method based on the Harris' algorithm.

Keywords: object recognition, to match, interest point, characteristic point, reference, learning, Harris method

1. Introduction

Très souvent, les industries modernes de production utilisent des machines pour automatiser au maximum les tâches humaines. Cela permet, non seulement, de diminuer les coûts de production, mais aussi, dans certains cas, d'assurer un volume de travail bien plus important.

Malheureusement, cette réduction de personnel enlève par la même occasion tout contrôle visuel concernant la qualité du produit.

Les technologies actuelles tentent de résoudre ce genre de problèmes en créant des environnements automatiques de plus en plus proches de l'homme.

Ainsi, il est possible de doter un ordinateur d'une caméra qui joue un rôle similaire à l'œil humain. Les images créées sont transmises à l'ordinateur qui les traite numériquement. Les résultats obtenus amènent la machine à prendre des décisions...

Mais comment amener un ordinateur à prendre la bonne décision ?

Il doit comprendre l'environnement dans lequel il se positionne et reconnaître les objets dont il doit effectuer le contrôle de qualité. C'est à ce moment qu'intervient l'ingénieur en imaginant les solutions les plus adéquates au problème posé.

Il existe un grand nombre de méthodes de reconnaissance d'objets.

Ce travail se base sur l'algorithme de Harris, qui est expliqué et approfondi de manière à reconnaître des objets au sein d'une image. Il est ensuite testé et adapté à un programme en particulier. Cette étude se termine par une batterie de tests sur des images en noir et blanc codées sur 8 bits à caractère quelconque (telles que des clés de portes et des clés USB) et à caractère industriel (telles que des lettres gravées au laser sur une surface plastique ou des chiffres en verre formés par un moule).

Remarque :

La reconnaissance d'objets pourrait paraître, pour une personne non avertie, une tâche des plus élémentaires ; mais en réalité, l'homme ne se rend pas compte du travail complexe qu'effectue à tout moment son œil en parallèle avec son cerveau. Derrière tout cela se cache une quantité innombrable d'opérations et de communications qui restent, à l'heure actuelle, sous certains aspects, encore très obscures.

2. Choix d'un algorithme comme « base de travail »

La reconnaissance d'objets n'est pas simple et nécessite certainement un développement approprié au type d'objet à reconnaître.

Une image est composée d'informations appelées **pixels**, chacun ayant une valeur bien déterminée.

Exemple : Une image noir et blanc codée sur 8 bits contient 256 valeurs par information.

Une analyse des informations invariantes telles que la **position**, la **rotation**, la **luminosité** et l'**occultation partielle** de l'objet permet de comparer les pixels d'une image à une autre.

Il est important de remarquer que le traitement doit être suffisamment rapide de manière à l'adapter en milieu industriel.

La référence [1] expose un développement appelé *la méthode de Harris*. Cette méthode semble être capable d'extraire ces informations invariantes quels que soient sa luminance, son bruit et sa rotation.

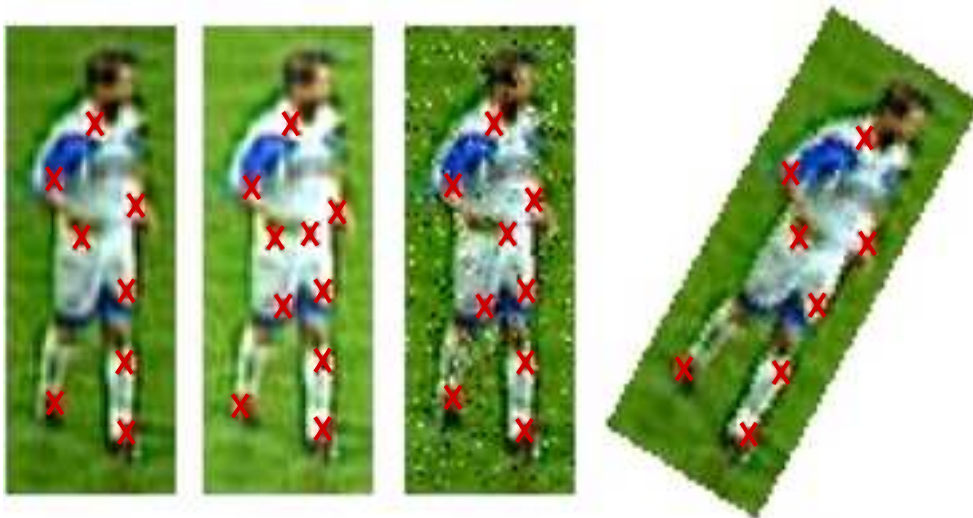


Figure 1 : Informations invariantes d'une image (provenance: [1])

Ces informations correspondent à des *variations bidirectionnelles de l'intensité de l'image*.

En d'autres termes : pour qu'une image possède une information utile, elle doit avoir une variation de son intensité de lumière suivant l'axe des X **et** suivant l'axe des Y.

Mathématiquement, une variation d'intensité de lumière peut être représentée par son gradient.

En prenant les cas de la fig. 2 : remarquons la différence entre la recherche d'informations suivant un axe ou suivant les deux axes en même temps.

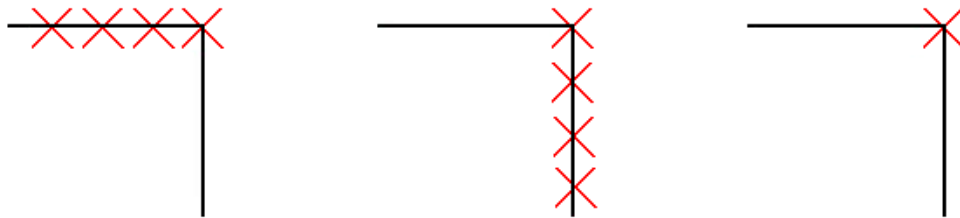


Figure 2 : Détection via le gradient en Y (image de gauche), en X (image du milieu) ou suivant Harris (image de droite).

Pour Monsieur Harris, le point représenté sur l'image à droite (fig. 2) s'appelle le *point d'intérêt* (PI). Ces PI ont, à eux seuls, suffisamment d'informations pour extrapoler un objet d'une image.

En résumé, un PI est un endroit précis d'une image correspondant à une information utile à la reconnaissance de l'objet.

De plus, une occultation partielle de l'image n'amène qu'une diminution de la découverte des PI sans pour autant influencer les autres points.

2.1 Explication plus détaillée de l'algorithme

Voici comment trouver mathématiquement ces PI.

Premièrement, définir une image comme étant "l'image" d'une **matrice I**. Ensuite effectuer, sur cette matrice, un **gradient horizontal** ainsi qu'un **vertical**. Cela donne comme réponse, des pentes représentatives des variations d'intensité pour deux pixels voisins du pixel sur lequel on travaille.

$$I_x := I \odot \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \qquad I_y := I \odot (-1 \ 0 \ 1)$$

Ensuite, appliquer le **carré** au gradient afin de ne garder que les valeurs positives qui seront **convoluées**¹ avec une **gaussienne W**. Cela permet de tenir compte de ses pentes dans une fenêtre circulaire autour de chaque pixel sans coupure brusque :

$$A := I_x^2 \odot W \qquad C := (I_x \cdot I_y) \odot W \qquad B := I_y^2 \odot W$$

Après, créer une **matrice M** pour chaque pixel qui reprend les matrices A, B et C :

$$M := \begin{pmatrix} A & C \\ C & B \end{pmatrix}$$

En calculant son **déterminant** sur chaque pixel ainsi que sa **trace** (qui équivaut à prendre la somme de la diagonale de la matrice), et en le soustrayant à un facteur K près, cela donne :

$$R := \text{Dét}(M) - K \cdot \text{Trace}_2(M)$$

On trouve ainsi une **réponse R** pour chaque pixel qui donne, en quelque sorte, le **degré** auquel celui-ci peut être un **point caractéristique**.

La réponse est positive lorsque les deux courbures sont élevées, ce qui est un critère nécessaire pour un point caractéristique.

R est négatif lorsqu'une seule des deux courbures est élevée, ce qui correspond plutôt à un point d'arête.

A cela, un seuil permet de filtrer les zones où les courbures sont trop faibles, ce qui enlève toute donnée parasite.

¹ Le produit de convolution sera représenté par \odot .

2.2 Portage de Harris

Le cahier des charges imposant la plateforme de développement (Visual C++ 6), quelques tests préliminaires sont néanmoins réalisés sous MatLab afin de vérifier la véracité de cet algorithme.

Malheureusement, les points trouvés sont trop nombreux.

La raison en est la suivante : si un pixel présente dans ses alentours une variation bidirectionnelle d'intensité suffisante, il trouve un PI. En prenant un pixel adjacent, il y a fort à parier que lui aussi trouve un PI correspondant à la même variation d'intensité. Ce PI n'est qu'une information redondante et peut donc être supprimée. Un masque est créé, assurant dans une zone bien définie, l'application du traitement sur le pixel le plus significatif.



Figure 3 : Algorithme de Harris employé sur un chiffre.

L'algorithme est amélioré en créant un seuil adaptatif de manière à toujours trouver un minimum de 10 PI dans l'image.

Ces informations peuvent s'avérer utiles, nécessaires et suffisantes dans la reconnaissance d'objets.

A ce stade, la méthode détecte encore un autre type de points parasites.
En effet, en examinant la fig. 3, l'algorithme détecte des PI en dehors du chiffre « 0 ».

En effectuant un zoom dans cette zone, il en ressort de petites variations d'intensité suffisantes amenant à la détection d'un point de Harris.

En réalité, ces erreurs ne proviennent pas de la qualité de la prise d'image mais du modèle lui-même.

Pour résoudre ce problème, il est nécessaire d'imaginer une méthode (fig. 4) permettant de « gommer » le bruit sans pour autant modifier de manière significative l'objet en question.

Il s'agit de chercher le pixel ayant la plus forte luminosité dans une matrice carrée délimitée à 3*3 pixels, et de le recopier sur le pixel en haut à gauche de la matrice. Il suffit ensuite d'effectuer ce balayage sur toute l'image.



Figure 4 : Harris détecte aussi les PI d'une image pauvre en informations.

Des tests sont menés dans les cas limites où l'image de base contient très peu d'informations au départ.

L'œil humain commence à avoir des difficultés à distinguer l'objet, Harris, lui, perçoit bien la différence (fig. 4).

Une fois l'algorithme de base réalisé et amélioré, il reste à développer une routine permettant, d'une manière ou d'une autre, de reconnaître les objets et par la même occasion, de donner un indice de corrélation représentant le pourcentage de reconnaissance.

3. Elaboration d'un système de reconnaissance

Trouver les points de Harris est une chose, inventer un procédé permettant de reconnaître les objets en est une autre. Les paragraphes suivants tendent à expliquer la méthode imaginée pour y parvenir.

3.1 Classement des informations

Les points de Harris véhiculent **deux informations** : la **position** exacte des points d'intérêts ainsi que leur **valeur** avec laquelle ils présentent une courbure importante.

Au plus une courbure est importante, au plus la réponse du détecteur de Harris donne une réponse élevée.

On peut dire que les points de Harris résultant des courbures les plus élevées, possèdent les plus grandes chances d'être les plus invariants suivant différents objets. Les points sont classés dans un tableau ordonné suivant leur poids.

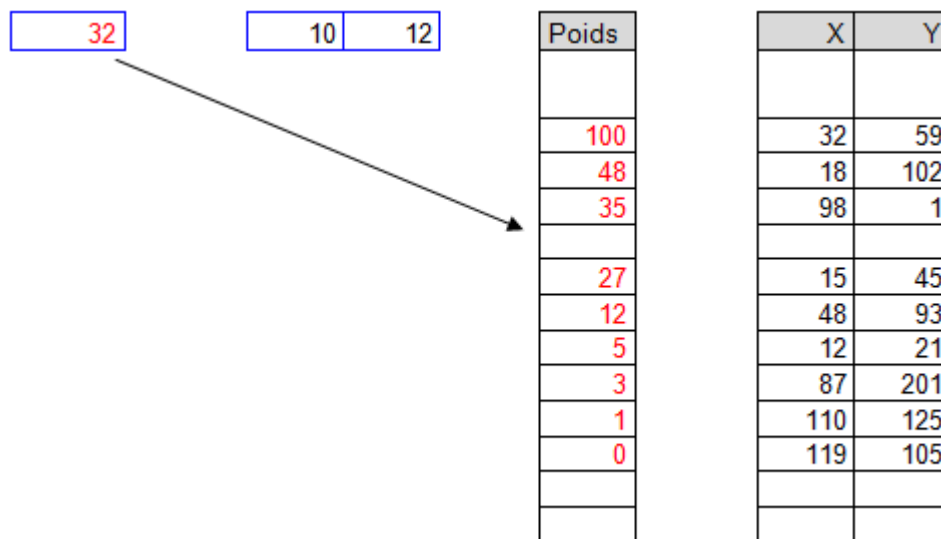


Figure 5 : Classement des PI suivant leur poids.

3.2 Mise en correspondance des points d'intérêts

Les points classés, il faut établir des liens entre chacun d'entre eux.

Le lien le plus direct entre les différents points caractéristiques est certainement la distance.

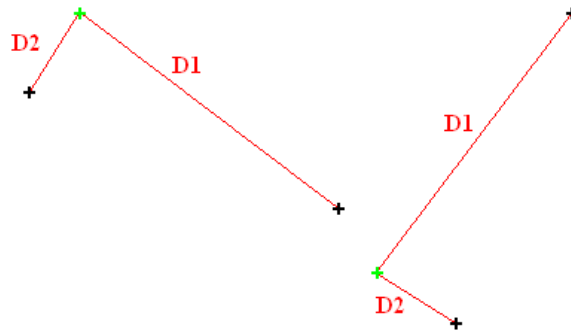


Figure 6 : Mise en correspondance des distances entre deux points.

Remarque : A la fig. 6, les distances séparant une référence choisie arbitrairement des autres points de Harris permettent d'établir rapidement le lien entre les deux images.

Cette méthode présente un avantage considérable. En effet, une rotation effectuée sur l'objet ne change en rien le calcul des distances.

Il suffit ensuite de considérer chaque point de Harris comme étant, à un moment donné, la référence. Cela permet d'assurer la reconnaissance si l'un d'entre eux disparaît.

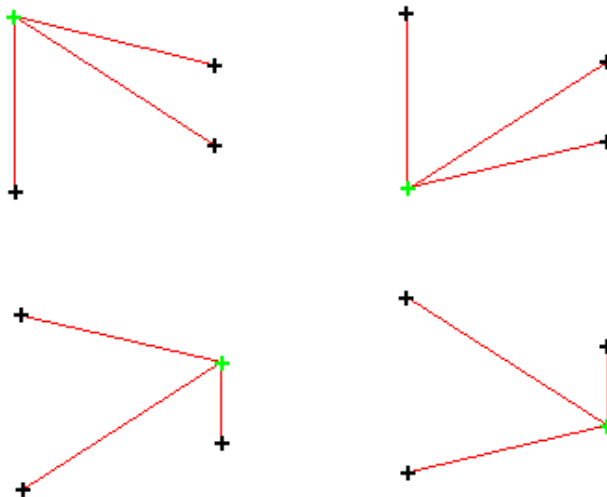


Figure 7 : Chaque point est considéré comme une référence.

Afin de ne pas confondre deux points se situant à égale distance d'une référence, il faut introduire un paramètre supplémentaire : la notion d'angle. Cet angle se mesure entre deux droites ayant pour sommet le point de référence.

Finalement, une information utile est définie par **deux distances** et leur **angle correspondant** pour un point de référence choisi.

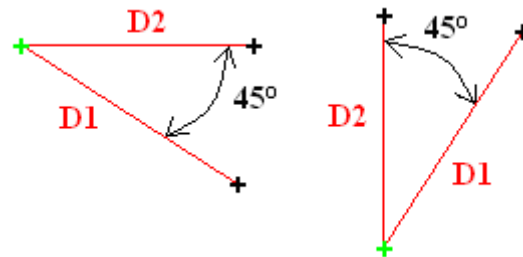


Figure 8 : Représentation d'une information utile.

Dans le cas de la fig. 7, il est alors possible de définir trois correspondances pour le premier point, ce qui fait un total de douze correspondances.

Ces correspondances se calculent en fonction du nombre de points par les formules de la statistique :

Avec N points, il y a $[N*(N-1)*(N-2)]/2$ possibilités.

3.3 Création d'une référence

Avant de pouvoir reconnaître un objet, il est important de définir à quoi il doit faire référence. Ainsi, un bouton est ajouté dans la routine afin de décider du choix de l'image de référence. Cela met en mémoire toutes les informations utiles définies plus haut.

En comparant les informations de la référence avec une nouvelle image, il est possible de donner avec une bonne précision son pourcentage de reconnaissance.

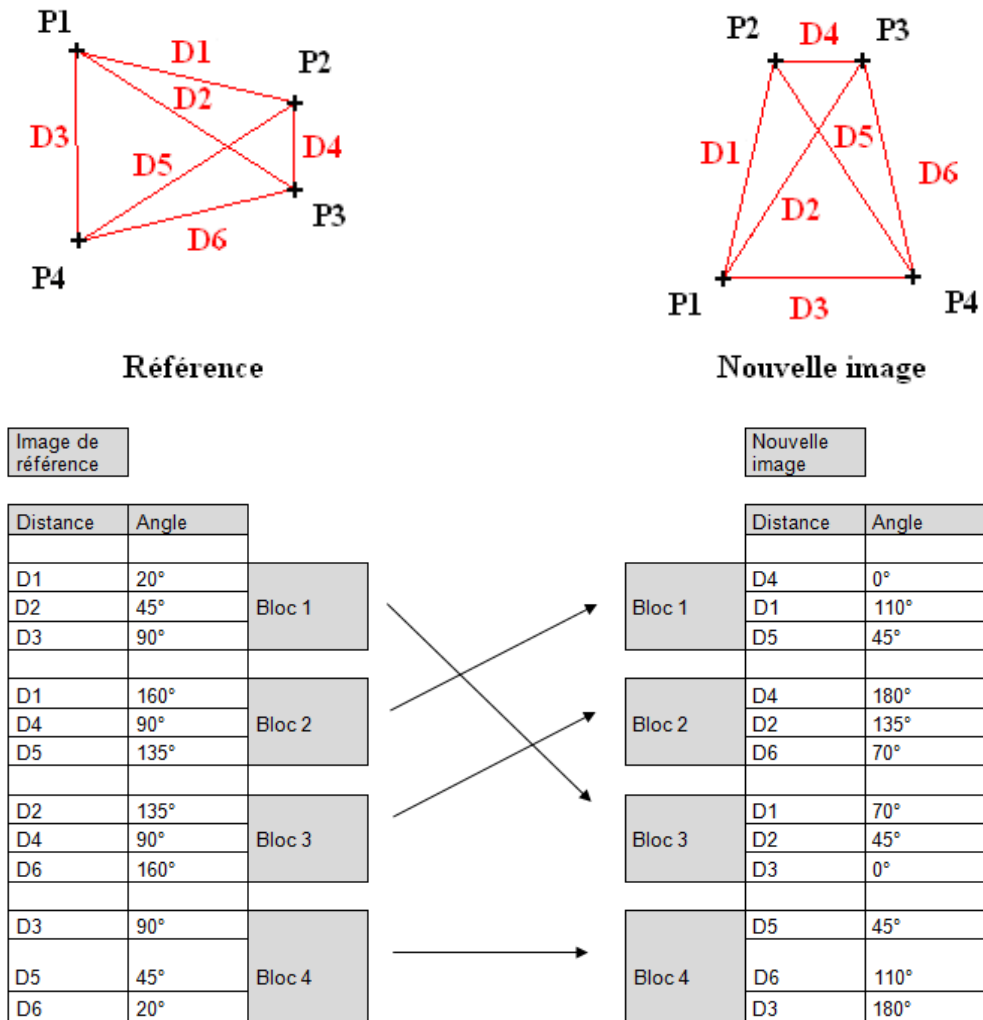


Figure 9 : Mise en correspondance de deux images.

Dans l'exemple de la fig. 9, l'image de gauche est la référence, et celle de droite, la nouvelle image à comparer.

Le programme recherche automatiquement toutes les correspondances entre ces deux tableaux.

L'organigramme de la fig. 10 montre son principe de fonctionnement.

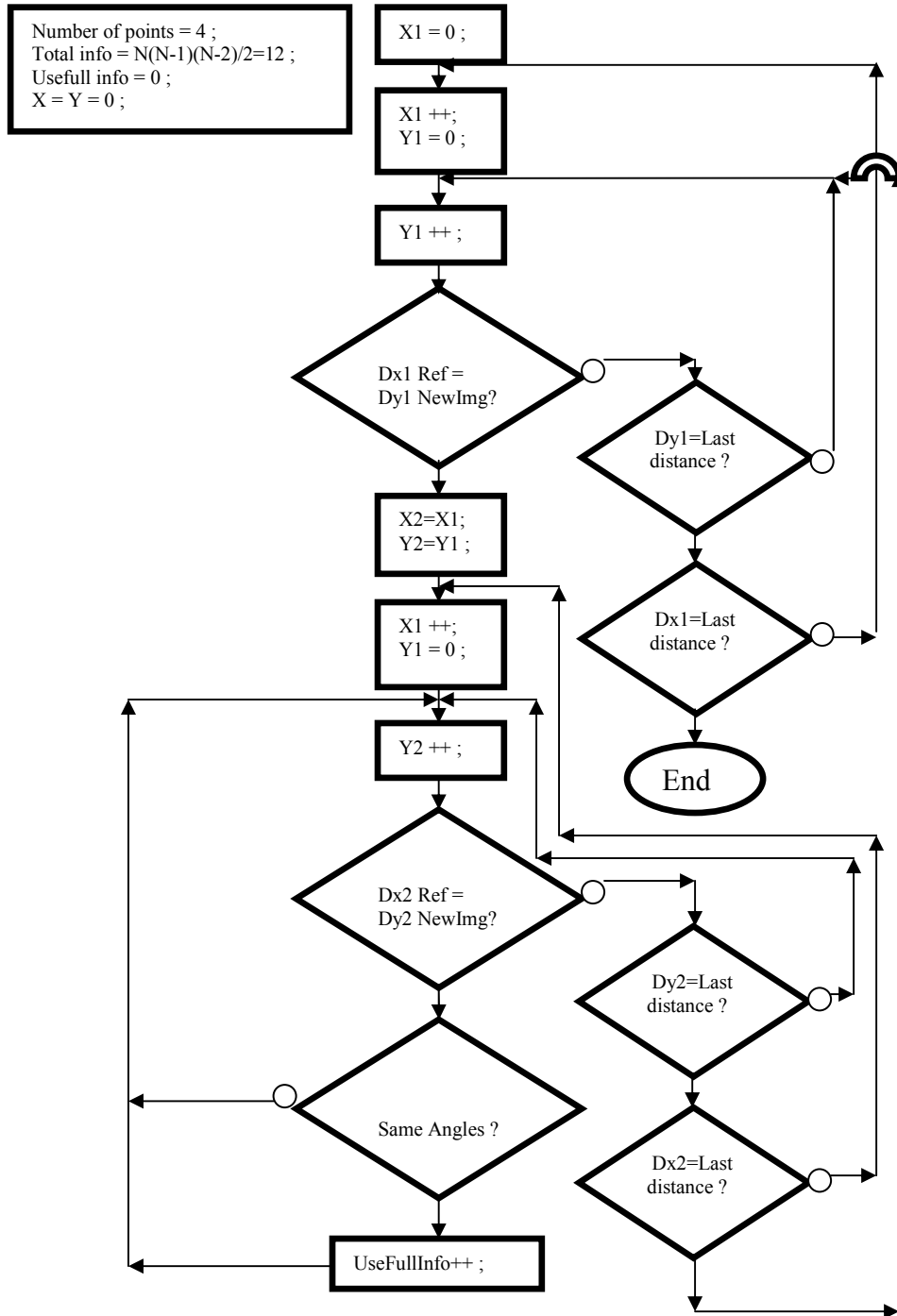


Figure 9 : Organigramme de mise en correspondance de deux images.

Au terme du procédé, il suffit de comparer la valeur de « Usefull info » avec « Total info ».

Pourcentage de reconnaissance = Usefull info / Total info = 12 / 12 = 100%.
Il est donc possible de conclure que la méthode est efficace si l'objet subit une variation de l'angle de rotation.

3.4 Ajout d'une tolérance

Dans la recherche d'objets en milieu industriel, chacun possède ses propres caractéristiques. Cela amène donc, au détecteur de Harris, à trouver des points d'intérêts variant légèrement suivant le modèle et suivant le bruit généré aléatoirement. Une tolérance est ainsi accordée sur chaque point de Harris. Une illustration à la fig. 11 représente cette tolérance accordée horizontalement, il en est de même dans toutes les directions de manière à créer un cercle de tolérance sur chaque PI.

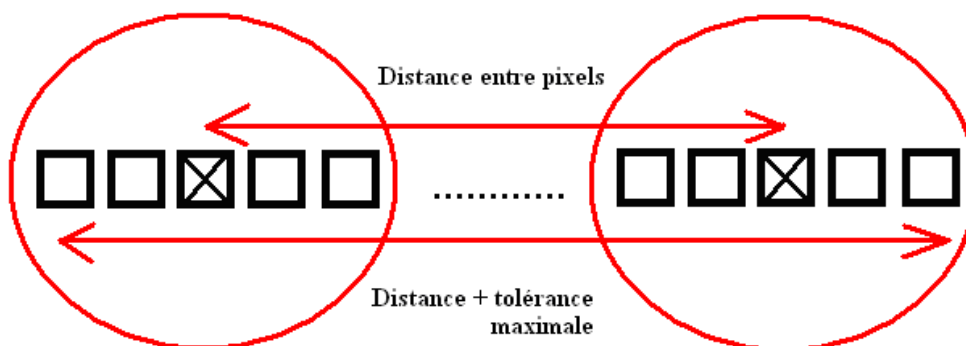


Figure 11 : Tolérance horizontale accordée à deux points de Harris.

3.5 Apprentissage de l'objet

De nombreux tests permettent de mettre en exergue l'importance de l'image de référence. En effet, si cette dernière ne présente pas beaucoup de points correspondant à la moyenne des images, le programme calcule des pourcentages de reconnaissance désastreux.

La méthode la plus simple pour contrer ce problème consiste à ne garder en mémoire que les points ayant la plus grande chance de se retrouver dans un nombre maximal d'images correspondantes.

Après avoir enregistré l'image de référence, il est possible, si le pourcentage de reconnaissance par rapport à la deuxième image est relativement élevé (par expérience un seuil de 60% semble être suffisant) de ne garder en mémoire que les points correspondant à la reconnaissance des deux images.

Cela amène donc à limiter le nombre de points à mettre en correspondance. Il est important de se rendre compte qu'une base de référence reposant sur un nombre très restreint de points ne peut pas être robuste.

Il faut donc impérativement effectuer des tests sur un nombre assez élevé de points. Dans la pratique, vingt à trente points semblent amplement suffisants pour s'appuyer sur des bases solides.

3.6 Mise en place de plusieurs modèles de référence

Une image de référence amène parfois à un pourcentage de reconnaissance trop limité. Le choix d'une deuxième image comme référence est alors nécessaire. La routine compare ainsi l'image aux deux références enregistrées et donne le pourcentage du meilleur candidat.

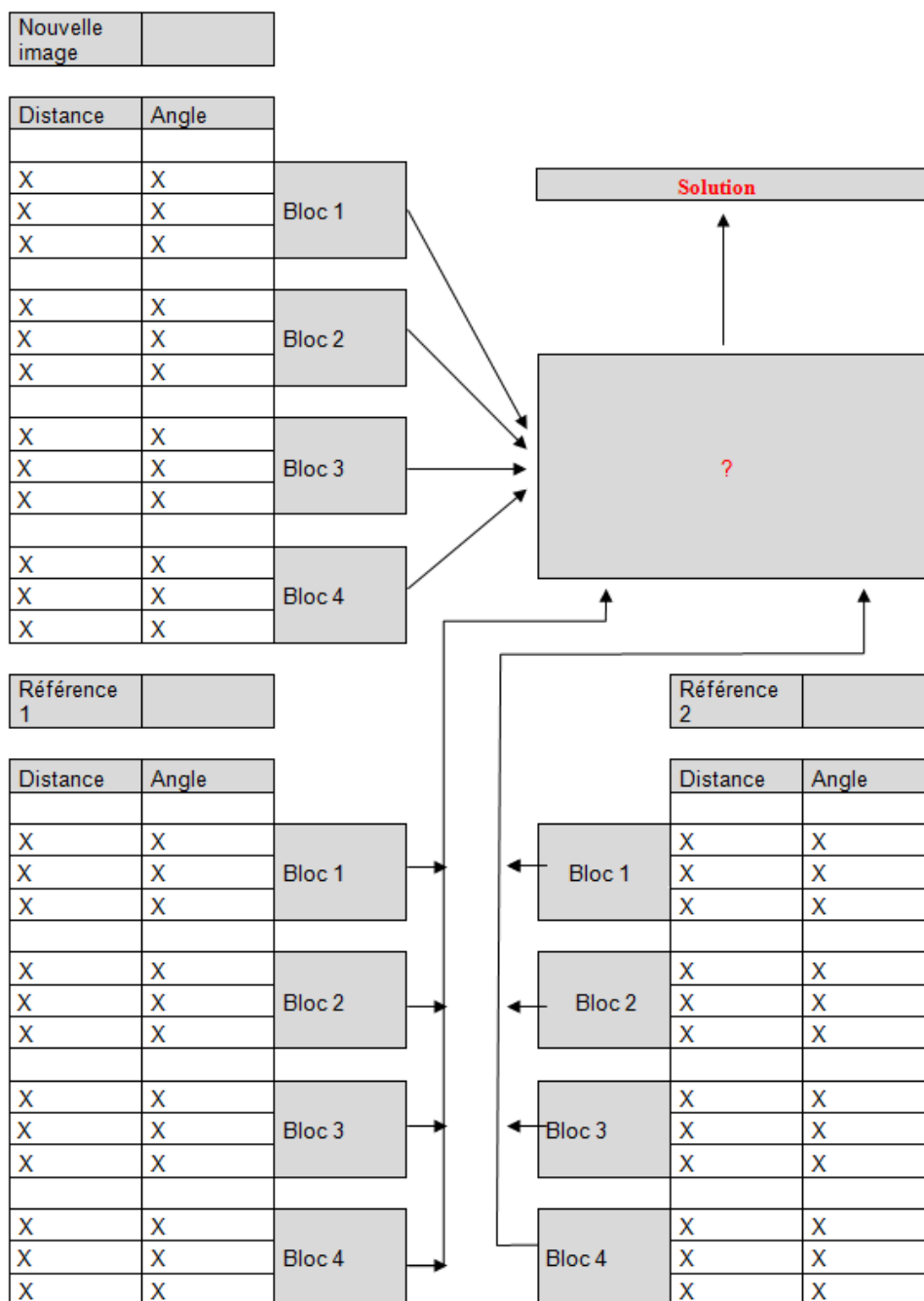


Figure 12 : Comparaison d'une image avec plusieurs références.

4. Résultats expérimentaux

Dans cette section, trois groupes d'images sont testés.

Le premier, sur des images dites « simples », les deux autres, sur des images à caractère industriel.

4.1 Images simples

Une image simple est une image présentant des contours très distincts et quasi identiques à l'image de référence.

Le test vérifie que l'algorithme peut distinguer une clé de porte d'une clé de type USB.

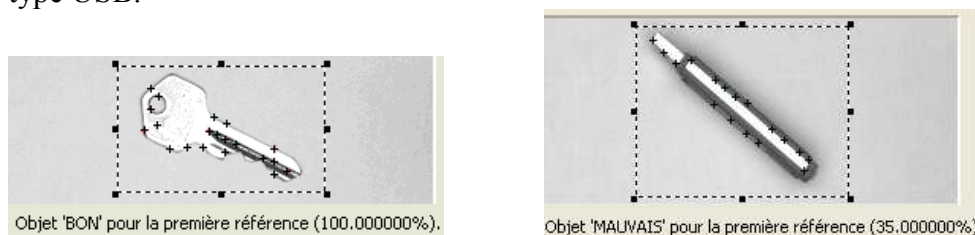


Figure 13 : Images simples utilisées comme premiers tests.

Les résultats sont très convaincants. Après avoir généré l'image de référence, une soixantaine d'images permettent d'arriver à la conclusion que l'algorithme retrouve en moyenne 93% de cette dernière, et que cette reconnaissance tombe aux alentours des 35% dans le cas d'une image différente. Un seuil décisionnel est mis en place affirmant qu'un objet est retenu comme étant l'objet de référence si son pourcentage de reconnaissance dépasse les 60%.

Cela donne lieu à 100% de bonne reconnaissance et 0% de mauvaise reconnaissance.

4.2 Images complexes

Chaque objet de la batterie d'images du deuxième type (image de gauche fig. 14) est photographié trois fois lors d'un passage dans une chaîne de production.

Mais peut-on effectuer trois traitements pour un même objet dans un process industriel ? Le temps ne joue-t-il pas un rôle important ?

En utilisant un processeur Intel® Core™2 Duo E6420 avec 4Mo de cache cadencé à 2400Mhz, le temps de calcul pour reconnaître l'objet ne dépasse pas les $3 \times 75\text{ms}$: soit 225ms.

En fonction des demandes d'une entreprise, ce traitement est applicable si et seulement si le temps de passage d'un produit est de 225ms au minimum.

Après calculs, en effectuant la moyenne des résultats obtenus par objet, le pourcentage de bonne reconnaissance est de 88.89% et la mauvaise reconnaissance de 0%.



Figure 14 : Images complexes utilisées en milieu industriel

Les résultats de la batterie de tests du deuxième type (image de droite fig. 14) sont insatisfaisants. En effet, les deux cent images analysées font ressortir un taux de bonne reconnaissance ne dépassant pas les 50% mais surtout un taux de mauvaise reconnaissance proche des 40% !

En analysant de plus près ces images, certains traits (fig. 15) beaucoup plus marqués que d'autres amènent à une détection beaucoup plus grande des points de Harris.

Dans ce cas, les imperfections sont liées à la non-linéarité de la vitesse de passage du marquage au laser et ainsi au modèle lui-même.

Les résultats sont faussés de manière significative.

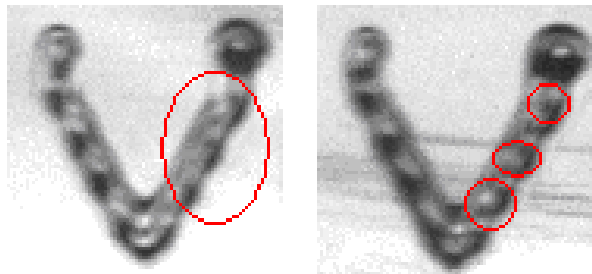


Figure 15 : Bruit perturbant l'algorithme de Harris.

Harris n'est donc pas satisfaisant dans ce cas de figure.

5. Conclusion

Ce travail débute par une recherche des différents ouvrages réalisés sur la reconnaissance d'objets. Ceux-ci démontrent qu'il faut développer un traitement propre au type d'objet à reconnaître.

Afin de traiter une image, il est nécessaire d'en extraire les informations utiles souvent appelées « invariants » ; celles-ci se positionnent à des endroits particuliers qui ne varient pas selon la position, la rotation et la luminosité de l'objet. Il est donc impératif de développer une routine permettant d'extraire ces informations avant de les apparier.

La méthode de Harris est adoptée car divers travaux prétendent qu'elle répond à ces critères. Après avoir effectué un « portage » de ce traitement en C++, ce dernier est amélioré pour le rendre plus robuste aux diverses interférences.

Dans un deuxième temps, une méthode complémentaire est mise au point afin d'apparier les objets en fonction des distances, des angles et des tolérances des multiples points de Harris.

Cette étude se clôture par une batterie de tests effectués sur divers types d'images.

Les résultats des expériences démontrent la robustesse de la routine sur la reconnaissance d'objets dits simples.

Sur des images plus complexes, ces résultats donnent un pourcentage de réussite de l'ordre de 88,89% si la donnée « temps » n'est pas un problème ou si la machine en question utilise des processeurs suffisamment performants.

D'autre part, une autre batterie de tests fait ressortir les faiblesses de la méthode de Harris qui empêchent une parfaite reconnaissance.

En somme, cette reconnaissance d'objets en milieu industriel est-elle possible ?

Oui, elle sera possible.

En effet, la méthode de Harris couplée à l'appariement est une méthode prometteuse mais qui nécessite certainement quelques travaux complémentaires.

6. Références bibliographiques

[1] GABRIEL P., HAYET J-B., PIATER J. et VERLY J., *Utilisation des points d'intérêts couleurs pour le suivi d'objets*, Université de Liège, Départements d'électricité, Electronique et Informatique Campus du Sart Tilman, Belgique.