

Analyse des semences par vision artificielle

F. Anouar, D. Bertrand, Didier Demilly, M.R. Mannino

► **To cite this version:**

F. Anouar, D. Bertrand, Didier Demilly, M.R. Mannino. Analyse des semences par vision artificielle. Colloque. Restitution des recherches sur les semences. Contrat de plan état-région. Région des Pays de la Loire 1994-98, Dec 1998, Angers, France. hal-02771672

HAL Id: hal-02771672

<https://hal.inrae.fr/hal-02771672>

Submitted on 4 Jun 2020

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.

Copyright

COLLOQUE BIOLOGIE ET QUALITE DES SEMENCES

RESTITUTION DES RECHERCHES SUR LES SEMENCES
CONTRAT DE PLAN ETAT - REGION DES PAYS DE LA LOIRE 1994-98

INRA
UNITE REGIONALE
DE DOCUMENTATION
Rue de la Géraudière - BP 71627
44316 NANTES CEDEX 03
Tél. : 02 40 67 51 90
Fax : 02 40 67 50 05

Inventaire	040499
N°	98. 2099

Angers
1-2 Décembre
1998

Inventaire	RUB
N°	2478

Organisateurs

Université d'Angers
Institut National de la Recherche Agronomique (INRA)
Institut National d'Horticulture (INH)
Groupe d'Etude et de contrôle des Variétés et des Semences (GEVES)
Fédération Nationale des Agriculteurs Multiplicateurs de Semences (FNAMS)
Pôle de Recherche et d'Innovation d'Angers - Secteur Semences (PRIA)

Pierre-Louis LEFORT
Coordinateur

ANALYSE DES SEMENCES PAR VISION ARTIFICIELLE

Fatiha ANOUAR¹, Dominique BERTRAND², Didier DEMILLY¹, Maria-Rosaria MANNINO¹

¹GEVES - SNES, BP 24, 49071 BEAUCOUZE Cedex

²INRA - PHYTEC, BP 71627, 44316 NANTES Cedex 3

RESUME

Nous présentons un outil d'analyse de semences par vision artificielle qui est composé d'un module physique concernant le matériel mis en place et d'un module logiciel relatif au développement et la méthodologie de reconnaissance des semences. Le premier module comporte une caméra couleur pour acquérir des images de semences, une carte d'acquisition pour digitaliser l'image et un ordinateur, la partie logiciel a donné lieu au développement d'un outil spécifique pour analyser les images de semences. Le logiciel, qui tourne sous un environnement convivial sous WINDOWS, inclut un module de traitement d'image, de traitement de données et un module de classification. Plusieurs méthodes de classification ont été développées et comparées en les appliquant à des cas d'analyses de traités en routine à la SNES (Station Nationale d'Essais de Semences). Ces méthodes de classification sont linéaires comme la méthode des proches voisins ou non linéaires comme les réseaux de neurones et permettent à l'utilisateur de choisir selon la difficulté du problème. Une optimisation est nécessaire pour trouver le meilleur modèle pour un problème donné. Différentes applications sont traitées : la reconnaissance d'adventices dans des lots de luzerne (*Medicago sativa* L.) et l'identification de luzerne, mélilot (*Melilotus sp*), trèfle violet (*Trifolium pratense* L.) et minette (*Medicago lupulina* L.).

INTRODUCTION

La vision artificielle par analyse d'image est en pleine expansion, elle permet de rendre compte des caractéristiques difficilement détectable à l'œil et permet d'automatiser des traitements de routine. Dans le domaine de la reconnaissance des semences, cette technique a montré ses preuves en particulier pour analyser un seul type de semences, comme par exemple le développement d'un système de vision dans le but d'identifier trois variétés de blé (UTKU, 1998). Des études sur des semences de petite taille sont inexistantes à cause de la difficulté technique de manipulation et de la variabilité que l'on peut rencontrer. La reconnaissance des semences, qui se fait actuellement manuellement, nécessite un savoir d'expert et est très coûteuse en temps. La SNES s'est dotée d'un système de vision pour cet effet et a mené une première étude de faisabilité du projet de la classification des semences par vision artificielle (CHTIOUI, 1997). Cette étude a en effet montré sur plusieurs types de semences que les méthodes de classifications développées ont des performances comparables aux analyses manuelles.

Dans la suite, nous présenterons le matériel utilisé pour l'acquisition des images, le logiciel d'analyse de semence et les méthodes de classification étudiées. La dernière section montre les performances du système pour des problèmes d'analyse courants en routine.

MATERIEL DU SYSTEME DE VISION

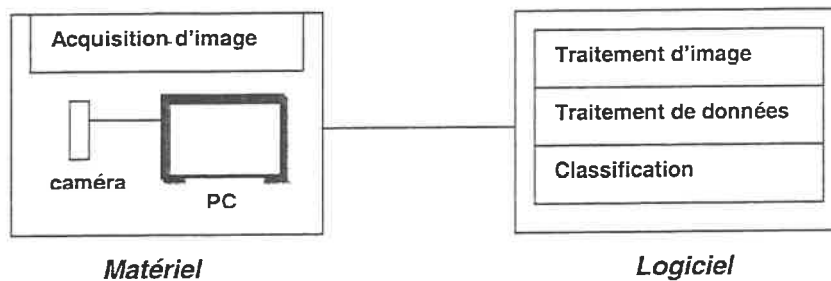
Le matériel utilisé est composé d'une caméra couleur haute résolution Sony. La caméra repose sur un bras qui peut se déplacer verticalement pour régler la hauteur et donc le champ de vision.

De part et d'autre de la caméra, un système d'éclairage avec deux lampes néon à puissance fixe permet un éclairage stable et uniforme. L'ensemble caméra/éclairage est isolé par une couverture noire pour garder un éclairage constant et stable sans interférence avec la lumière du jour. La caméra est reliée à un PC muni d'une carte d'acquisition Imascan pour digitaliser les images.

LOGICIEL ARTIVISU

Le logiciel ARTIVISU est développé en C++ sous WINDOWS et est composé de trois parties principales : traitement d'images et extraction de paramètres, traitement de données, classification. Le schéma de la figure 1 montre les deux modules du système de vision matériel et logiciel ainsi que les différents modules du développement logiciel.

Figure 1 : Schéma du système de vision



Traitement d'images

Après l'acquisition, un prétraitement est nécessaire pour améliorer la qualité de l'image (GONZALEZ, 1992). L'image est filtrée pour éliminer le bruit comme par exemple tout objet de l'image qui n'est pas une graine. Les objets touchant le bord sont supprimés. Après l'étape de prétraitement, les paramètres sont extraits de l'image couleur. Ils sont en nombre de 103 paramètres maximum répartis en trois classes : paramètres de taille, paramètres de forme et paramètres de texture. Le tableau suivant en montre quelques exemples.

Tableau 1 : Trois familles de paramètres extraits

Taille	Forme	Texture
Surface	Moments invariants	Mesures statistiques
Périmètre	Descripteurs de Fourier	Niveau de gris
Largeur	Longueur des plages
.....	

Les variables de texture représentent les variations de couleur et sont extraits dans les trois couleurs de base rouge, vert et bleu. Ces paramètres permettent de rendre compte de caractéristiques invisibles à l'œil. Les paramètres extraits peuvent être exploités pour résoudre différents problèmes. Nous traiterons dans les applications des problèmes de classification de semences et un problème d'évaluation du calibre d'un lot de semences.

Traitement de données

Les paramètres extraits ne sont pas tous utilisés pour tous les problèmes, un choix judicieux des paramètres utiles permet de diminuer la charge des calculs et de s'affranchir des corrélations qui existent entre les variables.

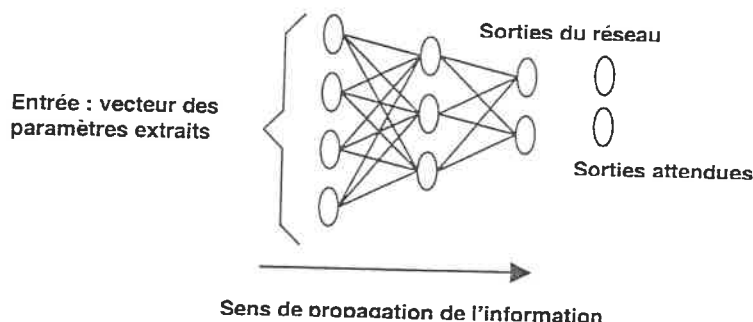
Pour un problème donné, la sélection des variables permet de ne garder que les variables qui permettent de discriminer au mieux les semences étudiées. Cela a l'avantage de bien représenter les caractéristiques des graines en gardant les variables les plus pertinentes.

Méthodes de classification

Une méthode de classification a pour but de trouver la meilleure fonction de séparation entre les classes selon un critère bien défini (ANOUAR, 1997). Plusieurs algorithmes de classification ont été étudiés, des algorithmes classiques comme la classification par proches voisins ou l'analyse discriminante. Ces deux méthodes génèrent des fonctions séparatrices linéaires. Les modèles de réseaux de neurones, quant à eux, trouvent une séparation non linéaire.

Les méthodes, citées plus haut, sont supervisées dans la mesure où on procède à une étape d'apprentissage supervisée dans laquelle est exploitée la connaissance *a priori* du problème. L'apprentissage supervisé utilise le type de semence pour adapter le système. Nous disposons d'un ensemble d'observations pour lesquelles la classe est connue, cet ensemble, appelé ensemble d'apprentissage, servira à adapter le système. La performance du système est évaluée sur un autre ensemble d'observations nommé ensemble de test. Le schéma de la figure 2 illustre l'apprentissage des réseaux de neurones. Dans le réseau de neurones, au cours de l'apprentissage, la sortie du réseau est comparée à la sortie attendue et l'erreur entre les deux mesures est rétropropagée pour adapter le système. La sortie attendue permet donc de superviser l'apprentissage.

Figure 2 : Schéma d'un réseau de neurones multicouches à apprentissage supervisé



APPLICATIONS

Pour chaque problème, des échantillons représentatifs des espèces connues sont utilisés. Les semences sont placées sur des étiquettes adhésives pour des raisons de commodité. Dans le cas d'un problème de classification, pour chaque application, plusieurs tests sont effectués pour pouvoir choisir le meilleur nombre de variables à sélectionner ainsi que la meilleure méthode de classification. Ce choix est effectué selon le taux de bonne classification. Nous présentons deux types d'applications d'analyse de semences dans un but de classification. La troisième application concerne l'analyse des semences de carottes dans le but d'évaluation du calibre du lot de semences.

Application 1 : Détection et identification d'adventices dans des lots de luzerne

Quatre adventices importants à détecter et à identifier pour la certification de lot de semences de luzerne (*Medicago sativa*) ont été choisis : le rumex (*Rumex sp*), la cuscute (*Cuscuta sp*), la folle avoine (*Avena fatua*) et l'avoine cultivée (*Avena sativa*). Ces adventices sont relativement faciles à détecter

par les opérateurs. Pour cette étude, environ trois cents semences de chaque espèce ont été utilisées pour l'apprentissage et deux cents semences pour le test. Plusieurs méthodes de classification ont été testées. Nous avons retenu la meilleure qui est la méthode des réseaux de neurones. Le nombre de variables retenues est de l'ordre de 10 variables. Les résultats de cette étude sont présentés sur le tableau 2.

Tableau 2 : Pourcentages de classifications correctes obtenus par la méthode des réseaux de neurones pour identifier des adventices dans un lot de luzerne. Les résultats sont donnés pour l'ensemble d'étalonnage et l'ensemble de test.

Espèce	Apprentissage	Test
Luzerne	100 %	100 %
Rumex	100 %	100 %
Avoine cultivée	98.23 %	98.36 %
Folle avoine	92.98 %	88.76 %
Cuscute	100 %	100 %
Pourcentage total	98.24 %	97.42 %

Pour ce problème, les erreurs du système concerne l'avoine cultivée et la folle avoine. Les semences de folle avoine mal classées sont toutes classées comme avoine cultivée. La difficulté de distinguer entre ces deux types de semences s'explique par le fait que les paramètres extraits n'intègrent pas la structure des semences qui permet de distinguer les deux types d'avoine. Il faut aussi noter que les semences d'avoine qui sont asymétriques sont placées sur les étiquettes de manière aléatoire sur une face ou sur l'autre. Les structures des deux faces sont complètement différentes. Il serait donc intéressant d'étudier l'effet de ce choix sur la performance de classification.

Ces résultats montrent l'intérêt de ces travaux pour séparer les différentes espèces, cependant une vérification des erreurs de l'identification des semences d'*Avena* est nécessaire pour obtenir 100 % de bonne classification. L'introduction d'une classe supplémentaire appelée *classe de rejet* permet de ressortir ces erreurs de classification. Toute observation qui n'est pas suffisamment homogène à l'une des classes existantes est attribuée à la classe rejet. Ainsi la performance sur les graines non rejetées est de 100 %. L'opérateur pourra ensuite analyser manuellement les graines rejetées.

Application 2 : Identification de luzerne (*Medicago sativa*), mélilot (*Melilotus* sp), trèfle violet (*Trifolium pratense*) et minette (*Medicago lupulina*)

Le problème de reconnaissance des semences de ces quatre espèces est un des problèmes difficiles traités par les analystes. En effet les graines présentent une grande ressemblance tant sur la forme que sur la couleur, et à l'intérieur de chaque espèce on rencontre une grande variabilité.

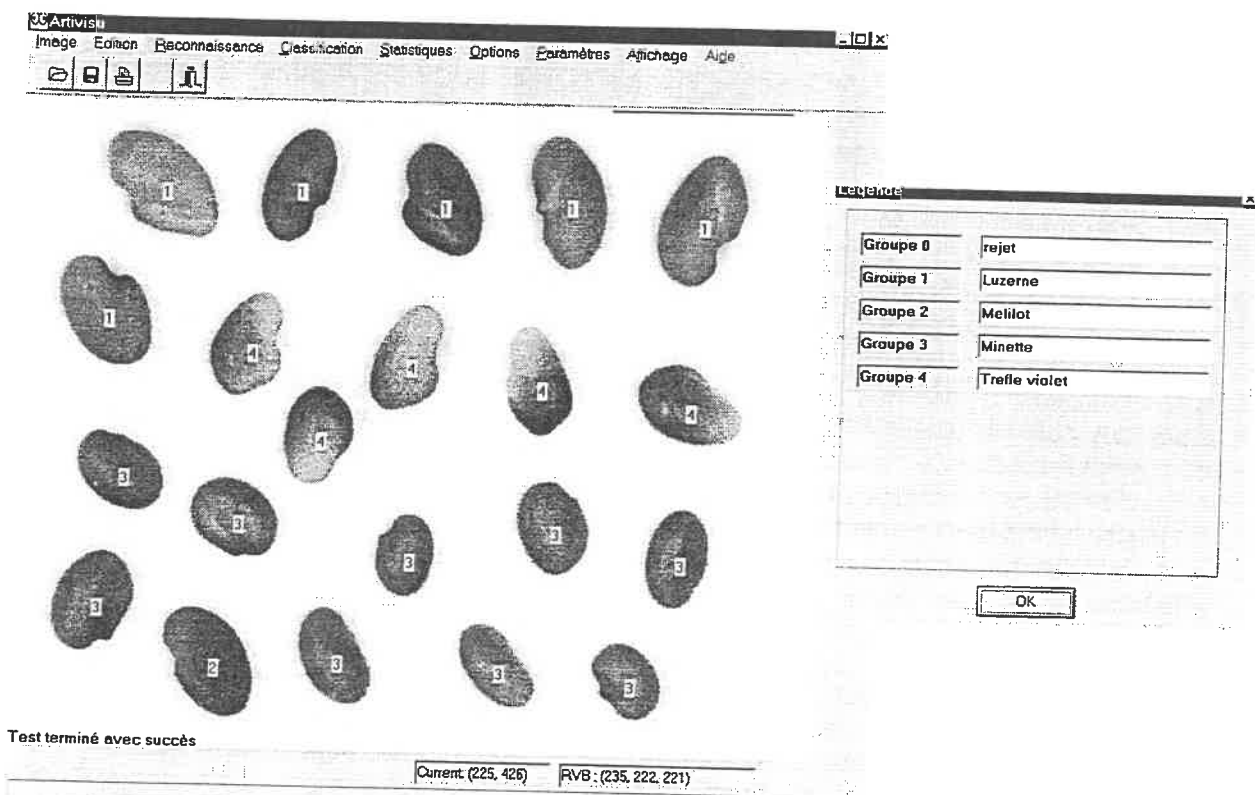
L'ensemble d'apprentissage et l'ensemble de test contiennent respectivement trois cents et deux cents exemples. Là aussi, plusieurs tests ont été effectués pour trouver le meilleur nombre de paramètres ainsi que la meilleure méthode de classification. Les performances de classification sont données dans le tableau 3.

Tableau 3 : Pourcentages de classifications correctes obtenus par la méthode des réseaux de neurones, sans classe de rejet, pour identifier luzerne, mélilot, minette et trèfle violet

Espèce	Apprentissage	Test
Luzerne	99.66 %	78.28 %
Mélilot	100 %	88.67 %
Minette	100 %	91 %
Trèfle violet	100 %	96.92 %
Pourcentage total	99.91 %	88.26 %

La figure 3 présente une image test sur laquelle nous avons disposé des semences des quatre groupes. Après traitement, le système affecte un numéro de groupe à chaque semence. L'image est accompagnée d'une légende. Sur cette image toutes les semences ont été bien reconnues.

Figure 3 : Identification de luzerne, mélilot, minette et trèfle violet



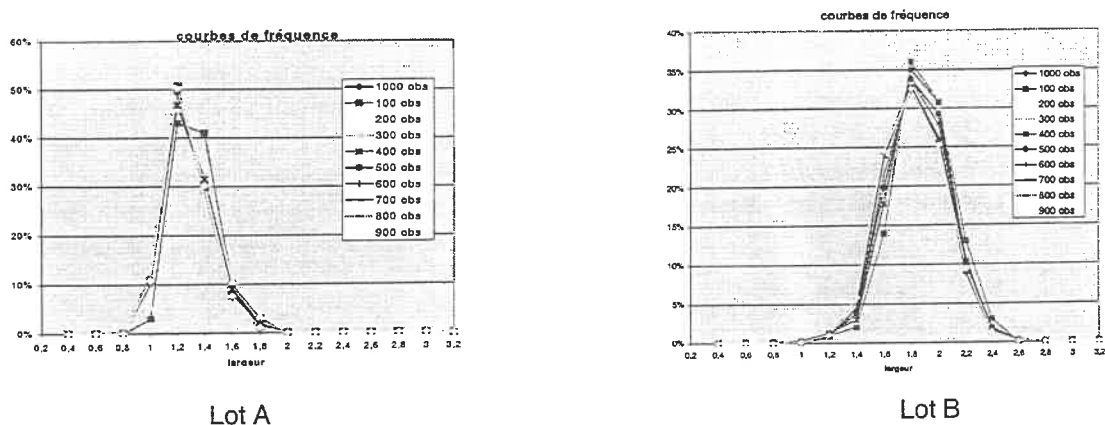
Pour ce problème, la difficulté rencontrée par le système est du même ordre que celle rencontrée par les analystes. Le système confond la luzerne et la minette : environ 20 % de semences de luzerne sont considérées comme des semences de minette. Par contre, comme pour les analystes, le trèfle violet est reconnu sans difficulté. Des améliorations peuvent être apportées à cette étude comme le placement des semences et l'utilisation de la classe de rejet.

Application 3 : Evaluation des calibres de lots de semences de carottes

Cette étude s'intègre dans le cadre d'un contrat de branche qui implique la SNES et dont la FNAMS (Fédération Nationale des Agriculteurs Multiplicateurs de Semences) est responsable. Ce programme porte sur l'élaboration de la qualité des semences de carottes, en particulier le lien entre le calibre et le pouvoir germinatif. A l'aide du système de vision, il est possible d'évaluer le calibre graine à graine et avec précision. Nous consacrons cette première étude à la détermination de la taille de l'échantillon pour pouvoir conclure sur le calibre d'un lot. Seul le paramètre largeur est utilisé pour évaluer le calibre des lots de carotte. Deux populations de carotte ont été étudiées. Deux échantillons de mille semences chacun ont permis d'étudier la taille de l'échantillon à retenir pour évaluer le calibre du lot. Sur la figure suivante, nous avons visualisé les courbes de calibre pour différentes tailles de l'échantillon de cent jusqu'à mille semences. Des tests statistiques ont permis de conclure qu'un échantillon de trois cents semences est suffisant pour évaluer le calibre du lot. L'erreur sur le calibre est dans ce cas de l'ordre de 0,02 mm.

L'évaluation du calibre à l'aide du système de vision se fait de manière automatique et donne des mesures très précises. D'autres caractéristiques de taille, comme la longueur ou la circularité, peuvent être exploitées pour étudier le calibre.

Figure 4 : Evaluation du calibre de deux lots de semences de carotte : Lot A = variété type nantaise, population à petite graine ; Lot B = variété type nantaise, population à grosse graine



CONCLUSION

La vision artificielle a montré sur plusieurs exemples son apport dans le domaine d'analyse de semences. Elle peut être utilisée en routine pour des problèmes simples dans le but de minimiser le coût d'analyse en terme de temps. Pour les problèmes plus compliqués, le système de vision peut être utilisé pour l'expertise. Toutes les études ont montré la validité de cette technique. Néanmoins, pour une utilisation en routine, des améliorations doivent être apportées comme le traitement d'images en 3D ou traitement des images de semences sur les deux faces. Par ailleurs, l'automatisation de toutes les tâches nécessite le développement d'un système mécanique qui permet d'acheminer les semences sous la caméra et les trier après traitements informatiques.

BIBLIOGRAPHIE

ANOUAR F., BADRAN F., THIRIA S., 1997 .

Probabilistic Self Organizing Map, application to classification. European Symposium on Artificial Neural Networks, ESANN, Bruges (B).

CHTIOUI Y., BERTRAND D., AND DEVAUX MF., 1997A.

Comparison of multi-layer perceptron and probabilistic neural networks in artificial vision. Application to the discrimination of seeds. Journal of chemometrics, vol.1 1, n° 2, 111-129.

CHTIOUI Y., BERTRAND D., DATTÉE Y., 1996A.

Identification of seeds by color imaging: comparison of discriminant analysis and artificial neural network. Journal of the Science of Food and Agriculture, 71, 433-441.

UTKU H., KOKSET H., KAYHAN S., 1998.

Classification of wheat grains by digital image analysis using statistical filter. Euphytica, 100, 171-178.

GONZALEZ R. C., WOODS R.E., 1992.

Digital image processing. Reading, Mass, Addison-Wesley Publishing Co.