

# NOTES SUR L'APPORT DU TRAITEMENT D'IMAGES À L'ÉTUDE DES SOLS

**Jean-Pierre LEFEBVRE**

*Centre de Recherches Orstom - Ile de France  
Laboratoire d'hydrophysique  
32 Avenue Varagnat, 93143 Bondy cedex*

## RÉSUMÉ

Un certain nombre de travaux récents dans le domaine des sciences du sol mentionnent le recours au traitement d'images. Bien que mal connues, ces techniques semblent *a priori* attrayantes parce qu'elles portent sur un objet proche d'une observation de terrain, à savoir l'image. Le but de cette note est de familiariser le lecteur aux différentes étapes constituant une chaîne complète de traitements, depuis le prélèvement de l'échantillon jusqu'à l'analyse proprement-dite des paramètres à étudier.

On insistera sur la nécessité de définir au préalable, les modalités de l'analyse à conduire afin de conserver l'information à extraire, sans dégradation excessive.

**Mots-clés:** conditionnement, digitalisation, traitement d'images

## INTRODUCTION

Le traitement d'images, même s'il ne constitue pas en soi une technique nouvelle, a bénéficié des progrès récents de l'opto-électronique et de l'informatique. De plus en plus de laboratoires se dotent à faible coût de systèmes complets d'acquisition et d'analyse d'images. Les sciences descriptives particulièrement, semblent plus à même de profiter de l'apport fourni par ces techniques. Bien que la globalisation de l'information offerte par le traitement d'images peut paraître *a priori* séduisante, on ne sait pas toujours précisément qu'attendre de ces systèmes. La présente note a pour but de définir les potentialités de ces techniques et de mettre en lumière les différents écueils à éviter lors de l'approche d'un problème par le biais du traitement d'images.

Trois étapes principales seront à distinguer dans le cadre d'un problème classique de traitement d'images : le conditionnement de l'échantillon, la digitalisation et le traitement proprement dit de l'image. On s'efforcera lors de chaque étape de préciser la distance entre l'objet réel et sa représentation obtenue. Quelques exemples ayant trait à l'étude de la porosité d'un sol illustreront cet exposé. Le problème du passage de l'observation *in situ* à la quantification au laboratoire n'étant pas spécifique à l'approche étudiée, il ne sera pas abordé au cours de cette note.

## CONDITIONNEMENT DE L'ÉCHANTILLON

Dans la majorité des cas, l'objet étudié est sensé être représentatif d'un ensemble plus vaste que l'on souhaite caractériser, c'est notamment le cas d'un échantillon extrait d'un sol particulier. En dehors de toute considération sur les perturbations liées au prélèvement proprement dit, la question de la représentativité de l'échantillon doit se poser suivant des critères logiques, (par exemple: la taille de l'échantillon permet-elle que les pores les plus longs soient représentés sans être tronqués ?), ce qui oblige à une définition préalable des termes précis de l'étude.

Il est assez rare de pouvoir exploiter l'échantillon tel quel. La plupart du temps, il doit être conditionné sous forme d'un bloc poli avec ou non inclusion de produit fluorescent, ou sous forme de lames minces. On passe alors d'une information tridimensionnelle : structure porale dans le volume, à une information bidimensionnelle : plan d'observation (fig. 1). De plus, le choix du plan d'observation, dans le cas de structures à orientations privilégiées, peut se révéler crucial (fig. 2). On notera que pratiquement, le plan d'observation n'est pas une coupe du matériau au sens strict, mais au mieux, une projection sur l'épaisseur de la section pour une lame mince voire une fonction plus complexe tenant compte de la profondeur des pores représentés, pour un bloc poli avec adjonction de produit fluorescent.

On peut se rendre compte des difficultés de proposer à l'analyse un objet physique représentatif du milieu dont il est issu, sans connaissance *a priori* d'un certain nombre d'informations telles que l'orientation et les tailles de pores, l'homogénéité du milieu etc. Or, c'est précisément ce type d'information que l'on souhaite extraire par analyse d'images. Le lien entre le milieu et le plan destiné à être digitalisé peut être délicat à définir. Mais c'est ce plan et non strictement l'échantillon qui fera l'objet d'une digitalisation.

## DIGITALISATION DE L'IMAGE

L'image informatique est en réalité une matrice, c'est-à-dire un tableau de valeurs. Le passage d'un horizon continu à un horizon discret est appelé digitalisation (fig. 3). Les lois régissant cette transformation sont bien connues (théorème de Shannon). Elles permettent de quantifier la perte d'informations induite (fréquence de coupure, overlapping...). Dans certains cas, on pourra envisager d'en limiter les effets (fenêtres de pondération).

L'échantillon conditionné, même dans l'approximation d'un plan théorique d'observation, reste un élément physique borné mais continu, pouvant être décrit suivant cinq dimensions : deux coordonnées d'espace et trois dimensions de couleur (par exemple dans le système HSI : teinte, saturation et intensité) (fig. 4).

## Codage de la couleur

On peut coder une couleur en une expression numérique, soit par discrétisation de trois composantes équivalentes (codage couleur type RVB), soit par discrétisation suivant une composante unique pouvant prendre des valeurs comprises sur une certaine plage (codage en niveaux de gris) ou soit, dans le cas extrême, par attribution de valeurs binaires à l'aide d'un seuillage (codage noir et blanc). Le codage couleur est rarement utilisé dans le cadre de petits systèmes, d'une part parce qu'il nécessite un matériel d'acquisition plus coûteux, et d'autre part, parce que les traitements ultérieurs doivent être effectués en parallèle sur les trois composantes de codage, ce qui conduit le plus souvent à une augmentation rédhibitoire de la lourdeur de l'algorithme. Dans le cas du codage en niveaux de gris, deux couleurs ne différant que par la teinte (i.e. même saturation et intensité), seront confondues par le codage (par exemple: un grain jaune paraissant blanc pourra être assimilé de ce fait à un pore). Le nombre de bits permettant de coder l'ensemble des valeurs entre le blanc et le noir est appelé dynamique de codage. Quant au codage binaire, il nécessite l'utilisation d'un seuil dont la détermination est le plus souvent arbitraire. La binarisation utilisée plus tard dans le traitement aura une toute autre fonction que celle de codage.

## Codage spatial

Le passage d'une représentation continue à une représentation discrète est fortement conditionné par des paramètres matériels. Une valeur discrète ne correspond en effet pas à un point théorique de l'horizon continu (fonction de Dirac), mais à une moyenne sur un élément de taille et géométrie définies par les caractéristiques de la matrice à CCD (Charge Coupled Device) et du dispositif optique associé à la prise d'image. Il en résulte une distorsion des formes appelée "erreur de pixélisation" (fig. 5). La connaissance de l'arrangement des cellules actives de la matrice (trame rectangulaire ou hexagonale) peut s'avérer essentielle puisque cette information disparaît avec le codage. En fixant arbitrairement une géométrie de trame, on peut introduire un décalage entre les lignes et risquer ainsi de biaiser fortement des estimations basées sur des mesures de périmètres.

Le dispositif optique peut bien entendu permettre d'augmenter ou de réduire la taille réelle du sub-horizon moyenné dans une cellule de la matrice de CCD, mais cette action a également pour effet de modifier la taille de l'horizon analysé. On retrouve la limitation déjà évoquée précédemment de la taille du champ investigué et donc de sa représentativité. Compte tenu du nombre de cellules de la matrice à CCD, on doit optimiser le grossissement pour sélectionner un champ de l'objet significatif en terme de représentation d'hétérogénéité (par exemple : distinction des plus petits pores, non-troncature des plus grands...).

Il est à noter que chaque dispositif optique introduit des déformations spécifiques (fonction du grossissement, de l'ouverture du taux de transmission...), qui participent à la fonction de transfert du système. D'autres sources de bruits, telles que la fonction de modulation de l'éclairage d'excitation ou la présence de lumière ambiante devront en toute rigueur, être prises en compte et dans la mesure du possible, réduites. Le choix d'une chaîne d'acquisition bénéficiant d'une bonne définition (i.e. nombre de cellules actives de la matrice à CCD) et d'une grande dynamique de codage est bien entendu souhaitable, néanmoins, on peut être amené à se limiter à des configurations sous-optimales pour des raisons de places en mémoire ou de temps de calcul.

On mesure ainsi la distance entre le matériau à étudier et sa représentation informatique sur laquelle porteront les traitements. On doit à ce stade s'interroger sur la conservation, éventuellement sous une forme peu dégradée, des informations que l'on souhaite extraire.

## TRAITEMENT DE L'IMAGE

Il n'existe pas de traitement d'images standard. On doit garder constamment à l'esprit la notion suivant laquelle plus le traitement effectué sera spécifique aux caractéristiques propres de l'étude, plus les probabilités d'obtenir une réponse fiable et pertinente seront grandes. Un certain nombre de traitements ont été élaborés, chacun dans les conditions particulières d'une étude. Ainsi, une transposition d'un algorithme à une problématique légèrement différente peut-elle conduire à des résultats aberrants. Le sujet est trop vaste pour qu'on puisse le développer ici. Retenons simplement qu'un traitement est constitué d'une suite d'opérations de morphologie mathématique (un opérateur morphologique modifiant la valeur d'un pixel par application d'une fonction sur un voisinage plus ou moins étendu du pixel considéré). La conception d'un algorithme peut donc être assimilée à la traduction d'une stratégie en une suite d'opérations élémentaires modifiant à chaque étape et d'une manière contrôlée, une image source en une image destination.

On se limitera donc à quelques considérations d'ordre générale sur ce que peut être le traitement d'une image en distinguant deux parties: identification des objets à analyser et caractérisation de ceux-ci à l'aide de critères variés.

### Identification des objets à analyser

On extrait une classe d'objet du reste de l'image par application d'un critère de sélection. Pour ce faire, on a souvent recours à une technique de seuillage. Dans le cas de l'étude de la porosité d'un sol à partir d'une lame mince, on peut utiliser un critère lié à l'intensité du signal codé par exemple en niveaux de gris, et extraire cette phase par application d'une valeur seuil de niveau de gris (dans l'exemple d'une lame mince, le niveau de luminosité transmise à travers un pore peut être déterminé avec assez de précision pour servir de référence, à condition que l'éclairage incident satisfasse à des critères notamment sur sa stabilité lors du temps, son homogénéité spatiale,...). Compte tenu de différents paramètres plus ou moins quantifiables (diffraction de la lumière, effets d'ombre, imperfection dans la réalisation de la lame mince, etc.), une variation faible sur le niveau de seuillage peut conduire à une grande variation sur le résultat obtenu (fig. 6). Il est par ailleurs assez hasardeux de fixer *a priori* une valeur de seuillage. D'autres alternatives peuvent être mises en œuvre pour éviter ce problème. Comme souligné précédemment, l'image binaire obtenue est le résultat d'une sélection et non d'un codage.

### Analyse et caractérisation de phase

Une fois la sélection de la phase à analyser effectuée, on cherche à caractériser ces éléments constitutifs soit globalement soit individuellement (labelisation). Le plus souvent, un traitement préalable est nécessaire pour rendre l'information exploitable. Sans entrer dans les détails, on peut plus ou moins aisément faire porter une analyse sur cinq types de caractéristiques : la proportion d'une phase par rapport à une autre, la taille d'objet, l'orientation privilégiée d'objet, leur dispersion dans l'horizon et leur forme (fig. 7). Le problème posé devra donc s'articuler autour de ces caractéristiques, et pourra déboucher sur des études relativement complexe (comme par exemple celle de la connexion de la porosité). A ce stade, des limitations spécifiques tels que les effets de bords pourront entacher l'analyse d'un biais éventuellement important.

## CONCLUSION

Il ressort de ce qui a été dit précédemment, que le traitement d'image n'a pas vocation de technique de mesure mais plutôt d'aide à l'interprétation. On jugera plus prudent d'attendre de tels traitements, des informations relatives (prédominance d'une classe sur une autre, orientation privilégiée...) plutôt qu'une quantification stricte (à titre d'exemple, quelle serait la

pertinence de l'estimation du diamètre d'un pore présent dans l'échantillon de la fig. 2, après mesure sur sa représentation fig. 6 ?). En cela, l'analyse d'images ne saurait remplacer l'analyse classique de laboratoire mais du moins, peut-elle la compléter.

Il est à peu près indispensable d'adapter le traitement effectué aux conditions propres de l'étude. On pourra bien entendu s'inspirer de tels ou tels traitements même pratiqués dans des domaines différents, mais le rendement d'un algorithme sera d'autant meilleur qu'il intégrera des conditions spécifiques de l'étude. Des logiciels type 'boîte à outils' permettent cette approche. Néanmoins, la relative facilité d'utilisation de ces techniques est à l'origine même d'attitudes peu critiques vis-à-vis de résultats obtenus. Pour cela, il est nécessaire d'être en possession d'une part, d'une bonne connaissance du problème et d'autre part, de l'action exacte des opérateurs utilisés. L'élaboration de l'algorithme devra être la plus théorique possible afin d'éliminer le plus grand nombre d'interventions humaines lors du traitement (choix de valeur d'après un rendu...). L'analyse devra idéalement être conçue par le pédologue intéressé, ou du moins être réalisée en collaboration avec celui-ci.

On a tenté de montrer dans ce qui précède, la richesse potentielle mais aussi les difficultés de mise en œuvre du traitement d'images. Le plus souvent, par manque de connaissances précises sur ces techniques, des études peu adaptées à ces approches conduisent à de résultats décevants. Afin d'augmenter les chances de succès, le problème posé devra être en adéquation avec ce que l'imagerie permet de faire, ce qui implique une certaine connaissance de ce domaine.

Les quelques clefs fournies au cours de cette note, ne prétendent pas constituer une étude exhaustive du sujet. Il existe malheureusement très peu d'ouvrages de vulgarisation en français sur ce domaine. On conseillera néanmoins le "précis d'analyse d'images" de M. COSTER et J.L. CHERMANT (presse du CNRS, 1989) pour de plus amples informations.

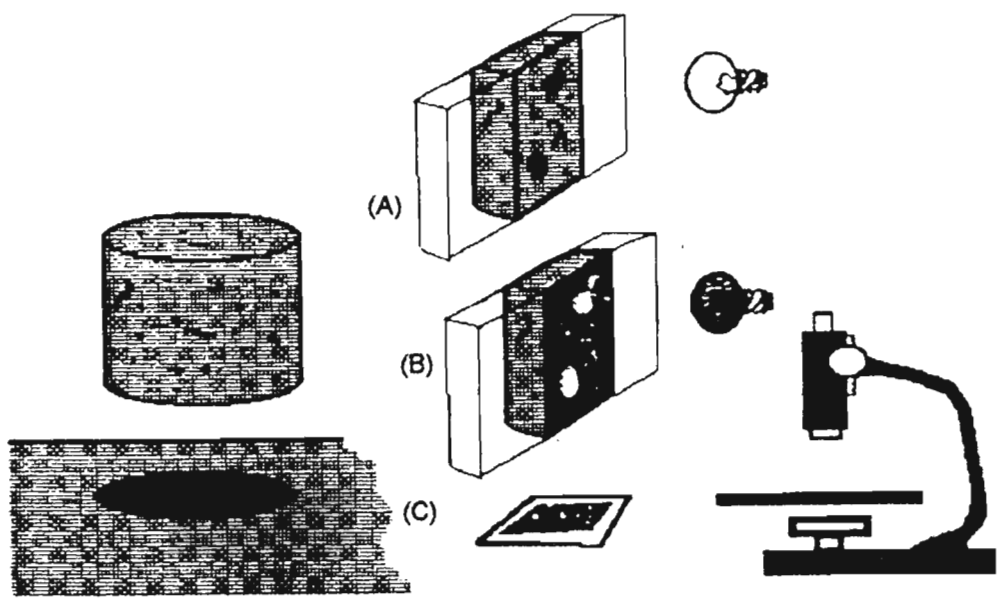


figure 1 : Passage d'un échantillon volumique de sol à un plan d'observation. Différents types de conditionnement (A) bloc poli, (B) bloc poli avec adjonction de produit fluorescent, (C) lame mince.

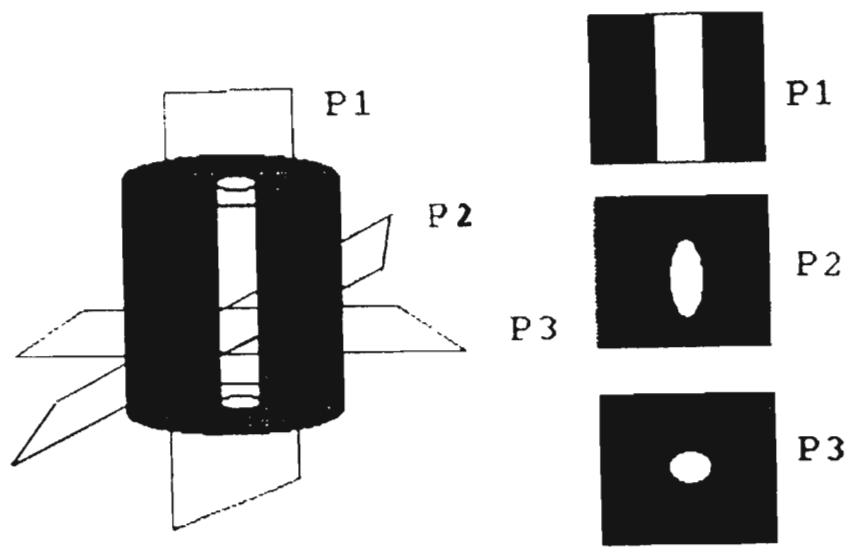


figure 2 : Influence du plan de coupe dans le cas de structures orientées.

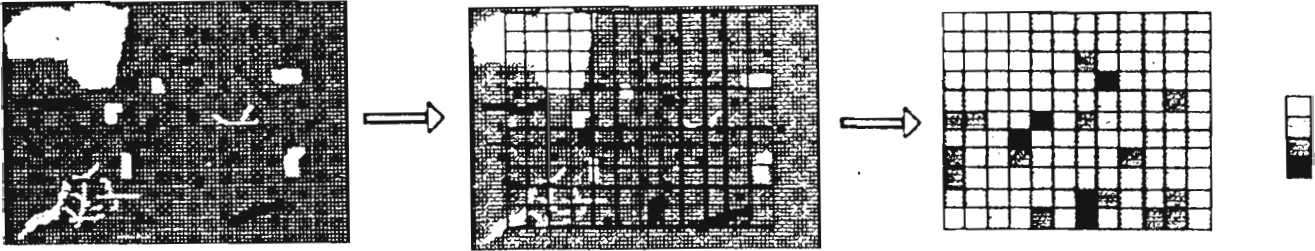


figure 3 : Digitalisation de l'image

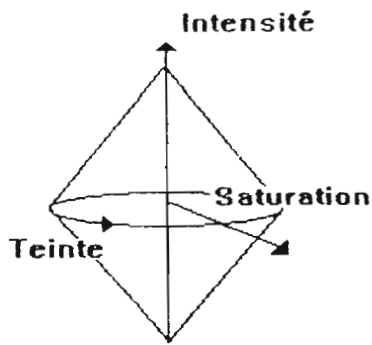


figure 4 : Représentation de la couleur dans le système HSI

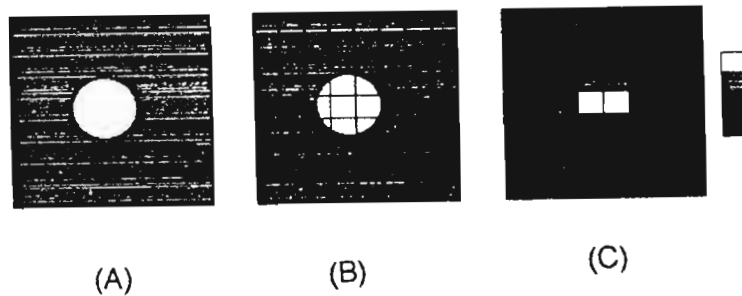


figure 5 : Erreur de pixelisation : (A) plan d'observation, (B) projection du plan d'observation sur la matrice à CCD, (C) image digitalisée résultante.

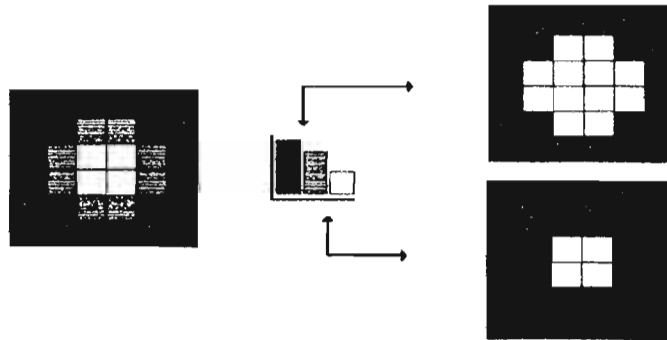


figure 6 : Erreur de seuillage : variation de l'image seuillée en fonction du choix de la valeur seuil

Proportion Taille Orientation dispersion forme

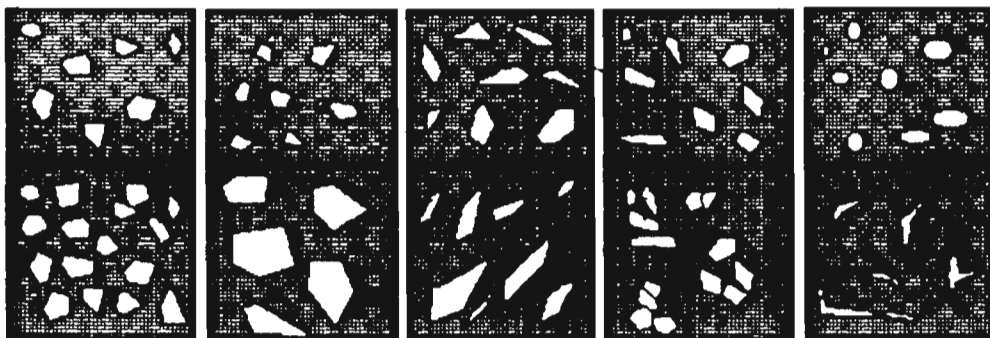


figure 7 : caractéristiques analysables par traitement d'images