

Les données vectorisées et maillées peuvent constituer des ensembles très complémentaires et de plus en plus de professionnels adoptent cette solution.

Les communes, grandes consommatrices de plans destinés à l'illustration de dossiers et de rapports, à l'information des administrés, pourront recourir sans hésitations aux plans scannés, éventuellement en complément d'un SIG.

Recommandations.

En tout état de cause, le scannage des cartes ou plans requiert quelques règles. Compte tenu du faible coût du scannage, il est conseillé de choisir une résolution élevée. Une résolution de 600 points par pouce soit 24 points par millimètre n'est pas excessive, surtout si des traitements doivent être effectués sur les documents.

Une résolution de 400 points par pouce soit 16 points par millimètre constitue un minimum. Cette résolution est suffisante pour obtenir des images de bonne qualité. En revanche, si des traitements sont nécessaires, l'image subira des altérations dès la première modification. Les planches cadastrales présentent, la plupart du temps, une orientation propre et doivent subir une rotation en vue d'un assemblage. Une rotation entraîne des approximations dans le calcul de l'image qui restent visible à l'œil nu, pour une résolution initiale de 400 points par pouce soit 16 points par millimètre.

Bibliographie :

Bernard ALLOUCHE "Les données maillées (ou données "raster") Signature n° 15, CERTU, Janvier 1999.

fiche établie par Michel Essevez-Roulet

avril 2000

Conseil National de l'Information Géographique
136 bis rue de Grenelle - 75700 PARIS 07SP

tél. 01 43 98 83 12 • télécopie 01 43 98 85 66 • Email cnig@cnig.fr • www.cnig.fr

FICHIERS RASTER ET DONNEES MAILLEES.

Les techniques de numérisation automatique progressent rapidement et modifient les options à disposition du technicien. La présente fiche rend compte de l'état de la question courant 1999. Une nouvelle fiche sera établie dans l'une des prochaines séries.

LES DEUX FAÇONS DE NUMÉRISER DES DONNEES

L'utilisateur peut numériser ses données de deux façons :

- La vectorisation, par la saisie de segments de droite (ou d'arc de cercle, de polygone ou de toute autre "primitive"), représentant le dessin initial, et appelés "vecteurs" dans le vocabulaire des mathématiques.

La vectorisation, manuelle ou semi-automatique, permet à l'opérateur de saisir de reconnaître les objets géographiques présentés par le document et de leur attribuer un code permettant à l'utilisateur de les retrouver, de les traiter, en fonction de ses besoins particuliers.

- Le scannage, ou la photographie numérique, qui produit un document électronique appelé fichier raster ou bit-map, contenant une information encore appelée "maillée". La norme française NF 52-000 (EDIGÉO) en décrit les règles de structuration sous la forme d'un modèle conceptuel de données matricielles.

LE SCANNAGE

Le terme raster vient du monde de la télévision et rappelle que l'image est traduite en électronique par balayage. Le terme bit-map est utilisé quand le fichier représentant l'image n'est pas compressé.

Les données recueillies n'incluent aucune reconnaissance de l'information. Elles ne font que reproduire une image, à l'aide des techniques de l'enregistrement électronique.

Le scannage ne se substitue pas à la numérisation vectorielle. Cette dernière est nécessaire pour reconnaître les objets qui constituent la carte, le plan, et dont la description est normalisée (cf. la nomenclature du CNIG).

LES DONNEES MAILLEES

Le scannage permet d'enregistrer sur disque magnétique ou sur CD-ROM un plan, une carte, un dessin ou une photographie qui sont déjà une représentation de la réalité.

L'appareil photographique numérique permet de constituer la représentation de la réalité



directement sur le média électronique. Par nature, la représentation de cette réalité est constituée par une grille de petits carrés, les "maillages", issue de la grille constituant le capteur de l'appareil.

Contrairement aux données en mode vecteur qui décrivent des objets géographiques à l'aide de points, de lignes ou de polygones, les données maillées décrivent la surface d'un territoire selon une grille régulière où chaque maille, ou encore pixel pour les anglo-saxons, présente une valeur codant l'information.

La structuration des données scannées ou maillées n'est pas impossible, mais sa mise en œuvre gomme l'intérêt de la technique.

INTERET DU SCANNAGE

Les cartes et plans scannés, ou les photographies numériques, présentent l'avantage de pouvoir être rapidement et facilement mis en œuvre. Leur production est économique et leur utilisation est facile. Les progiciels d'édition de fichiers raster facilitent l'assemblage des planches connexes, ainsi que la réutilisation des extraits de plans dans les rapports de présentation, les plans de situation dont les services font un usage fréquent.

La carte scannée : les photographies numériques peuvent ainsi servir de fond de plan, voire de référencement cartographique. Les progiciels récents savent afficher un document raster simultanément avec un graphisme vectoriel.

L'État a choisi cette solution pour numériser le plan cadastral de 1999 à 2001. La technique est plus simple et permet de mettre les

documents à la disposition du plus grand nombre dans un délai limité. Les services du cadastre n'ont pas pour autant abandonné la numérisation vectorielle du cadastre car les principaux utilisateurs en expriment le besoin (cas fréquent en agglomération urbaine).

Disponibilité

L'un des intérêts de la carte ou du plan scanné réside dans sa disponibilité. Scanner un document de grand format ne prend que quelques minutes et peut être effectué par de nombreux reprographes, pratiquement sans délai, pour un coût inférieur à 100 F.

Jusqu'à une date récente, les données maillées – qui existent depuis longtemps – étaient restées confinées à un usage relativement spécialisé (production de données par photo-interprétation, télédétection, vectorisation) faute de moyens techniques suffisants pour en tirer le meilleur parti. Cette situation évolue rapidement depuis quelques années au rythme de l'évolution technique.

L'I.G.N. a ainsi scanné l'ensemble de ses cartes du 1:25 000 au 1:1 000 000. Ces documents ont l'intérêt d'être en couleur – une couleur parmi 256 sur un pixel de 0,1 mm de côté. Les documents fichiers raster sont disponibles à la vente pour des prix variant en fonction des produits : (1 000 F pour la carte de France au 1:1 000 000 ou pour une dalle de 10 km de côté à l'échelle du 1:25 000).

Les planches cadastrales sont quant à elles en cours de scannage avec une précision identique. Toutes les communes devraient être couvertes à la fin de 2000 ou au début de 2001.

Les photographies aériennes et satellitaires bénéficient également des avancées techniques et permettent d'obtenir des données plus fines, dont la fabrication est fiabilisée et simplifiée.

Des images à très haute résolution (avec une maille correspondant à 10 cm sur le terrain) peuvent être obtenues directement par la caméra numérique, à des prix de revient en diminution constante.

L'I.G.N. s'apprête ainsi à commercialiser des orthophotoplans numériques départementaux avec une maille de 50 cm, et envisage pour un avenir plus lointain d'employer une caméra numérique pour proposer des produits d'une qualité encore meilleure avec des mailles très fines (de l'ordre de 20 cm) sur de plus faibles étendues (des villes par exemple).

La société Géosys propose également la couverture des principales villes avec des images numériques sur catalogue, mises à jour annuellement.

Cette liste d'offre n'est évidemment pas exhaustive, il faudrait citer d'autres sociétés : Eurosense, Ortea, InterAtlas, Géomage, etc.

INCONVENIENTS DU SCANNAGE

L'inconvénient de la carte ou du plan scanné, du document maillé, réside dans l'absence d'une reconnaissance des objets constituant le document. Cette caractéristique interdit certaines utilisations : liaison avec une banque de données techniques, définition d'attributs, analyse spatiale, etc.

Un autre inconvénient de la carte ou du plan scanné se rencontre dans les changements d'échelle. Alors que la représentation vecto-

rielle autorise une grande souplesse, les plans scannés acceptent difficilement les agrandissements ou les réductions d'un facteur supérieur à 2. C'est particulièrement vrai pour le plan cadastral dont l'assemblage regroupe déjà des feuilles d'échelles différentes.

Une réduction au 1:5 000, essentiellement utilisée comme fond de plan pour le P.O.S. constitue une limite, tandis que l'agrandissement au 1:500 d'un assemblage comprenant une planche initialement dessinée au 1:2 000 montrera cette partie avec des graphismes trop épais pour certaines utilisations.

Les difficultés d'exploitation

Bien que d'apparence souvent plus lisible pour l'œil, les données maillées présentent quelques difficultés d'exploitation dans un SIG.

En effet, sauf à représenter le terrain tel qu'il est (cas des images numériques du territoire utilisées comme simple "fond de carte"), ces données nécessitent souvent un travail d'interprétation des informations portées par la maille.

En outre, le principe du maillage est peu compatible avec une notion de base de données attributaires. Il est encore difficile de mettre en œuvre des informations attributaires complexes associées à chaque maille.

Enfin, les logiciels permettant une intégration de données maillées dans un SIG n'offrent pas de souplesse réelle de manipulation, en particulier pour le croisement des couches raster entre elles. Il est, par exemple, difficile de gérer des images dont les mailles ont des tailles différentes. Par ailleurs, en terme de res-

titution visuelle ou cartographique, une seule "couche" est visible. Certains logiciels offrent la possibilité de rendre transparente une couleur, mais cette solution n'est applicable qu'à des cartes rasters peu chargées, en général monochromes. L'exploitation croisée avec des données vecteur n'est généralement pas disponible sur les logiciels actuels, malgré l'intérêt énorme que peuvent représenter de tels croisements.

DIFFICULTES LIEES A LA NATURE DE L'INFORMATION

A ces difficultés générales d'exploitation, dues essentiellement aux outils logiciels, s'ajoutent d'autres difficultés liées à la nature de l'information elle-même. On peut ainsi distinguer quatre catégories de données maillées, comportant chacune leur limite méthodologique :

Les données raster traduisant une grandeur naturellement continue

On doit entendre ici des informations qui ont un sens en chaque point du territoire : une donnée d'altitude, de pente, d'orientation, des mesures de pression atmosphérique, de bruit, des temps d'accès, des cartes de distance, des valeurs de contrainte ...

Le mode maillé est pour ce type d'information le moyen le plus pertinent de représentation. La valeur portée par la maille est une valeur directement liée au phénomène représenté (par exemple, l'altitude pour le relief, figure 2). Il n'y a pas d'ambiguïté pour exploiter ces données.

Les difficultés viendront souvent des logiciels SIG qui ne permettent pas d'intégrer ces données dans toute leur richesse, ou ne permettent pas d'en tirer l'information utile.

Les données nécessitant une interprétation.

On classera ici les photographies aériennes ou satellites. Ce sont des images "naturelles" du territoire sous forme numérique.

Dans ce cas l'information portée par la maille correspond à une intensité lumineuse renvoyée par le sol. Toute la difficulté consiste à déterminer à quoi correspond cette intensité lumineuse : s'agit-il de forêt, de bâti, d'eau, de roche, de surfaces labourées ?

En fait, le codage de l'information n'est pas univoque et nécessite une interprétation faisant souvent appel à d'autres sources d'information (texture de l'image, connaissance des lieux, ...). Pour cela, trois voies sont couramment utilisées :

- soit l'interprétation totalement automatique, utilisant au mieux les techniques du traitement d'image : cette solution est ancienne, dispose de nombreuses références, mais souffre d'une évolution technique permanente et d'une absence de stabilisation des procédés ;
- soit une interprétation totalement réalisée par un opérateur : il faut alors disposer d'une réelle compétence, et la régularité de l'interprétation est toujours un problème ;
- soit une forme d'interprétation utilisant les potentiels des deux solutions précédentes, dans ce qui est appelé l'interprétation assistée par ordinateur : on prépare différentes vues de l'image à partir des techniques automatiques, et elles sont interprétées par un opérateur.

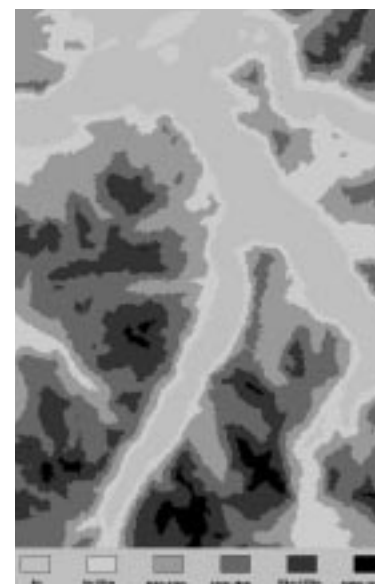
Ces données sont donc des sources d'information très riches et facilement actualisables,

mais elles nécessitent en contrepartie un haut niveau de savoir-faire pour être exploitées.

Les données portant une information déjà interprétée :

Chaque maille porte un code faisant référence à une information de façon univoque. Le code fait référence à une nomenclature (figure 5) bien précise (classification statistique du territoire selon la nomenclature CLUSTER proposée par EUROSTAT, par exemple).

Afin de représenter des objets très proches les uns des autres, on est obligé de diminuer globalement la taille de la maille, ce qui augmente de manière importante la taille des fichiers. Une autre solution consiste à utiliser



es techniques plus sophistiquées, telles que les quad-trees, en coupant une maille en quatre mailles plus petites, et ceci uniquement là où cela est utile. Malheureusement, elles sont peu souvent implémentées dans les logiciels SIG. Le principal frein à leur utilisation est le peu de données qui existent sous cette forme aujourd'hui.

Les données reproduisant une information déjà interprétée, existant sur papier :

Il s'agit des cartes scannées, telles que les cartes routières, le cadastre, ou les cartes géologiques.

Extraire des informations géographiques d'une carte scannée en s'appuyant sur la légende graphique présente beaucoup de difficultés : en effet, une couleur correspond rarement à un poste de la légende.

Des techniques de reconnaissance de caractères, de vectorisation automatique, alliées à des classifications de couleur, peuvent dans certains cas aider à extraire une partie de l'information contenue dans la carte scannée, afin d'en produire une information vectorielle. Leur emploi reste cependant difficile, sauf par des professionnels.

Les cartes scannées peuvent donc être utilisées comme référentiel visuel, comme fond de plan (elles présentent un aspect familier aux utilisateurs), comme source d'information (numérisation sous forme vecteur d'éléments graphiques contenus dans la carte), mais ne peuvent que rarement être utilisées comme une information pouvant être croisée directement avec une autre.

LES DIFFICULTÉS PRATIQUES

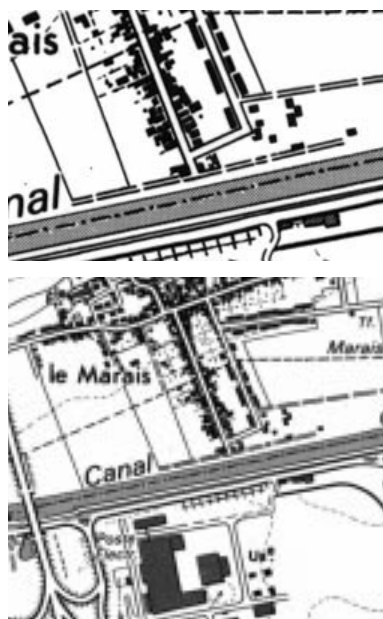
En faisant abstraction des raisons qui motivent (démontrent) l'intérêt d'utiliser des données maillées et des difficultés méthodologiques qui freinent leur usage réel, il demeure un certain nombre de réalités pratiques qui expliquent aussi une grande part de ce décalage tardif de l'usage des données raster dans les SIG.

La constitution

En raison de la rareté de l'offre "sur étagère" d'images numériques, géométriquement correctes sur des territoires étendus, les besoins des utilisateurs en données du type image, ne peuvent être satisfaits que par des prestations ponctuelles coûteuses ou par l'emploi d'outils et de méthodes réclamant des savoir-faire peu répandus.

Réalisée à partir de clichés, une couverture orthophotographique numérique d'un territoire nécessite au moins trois étapes importantes et délicates :

- numériser les clichés ; les scanners actuels permettent une numérisation facile, avec des gammes de qualité allant d'une qualité bureautique (trop faible pour une utilisation géographique) à des qualités métrologiques, avec du matériel très coûteux.
- les corriger géométriquement individuellement, en tenant compte des déformations introduites par l'appareil de prise de vue, celles inhérentes au relief (ortho-rectification), celles introduites lors de la numérisation ;
- assembler (mosaïquer) les différentes images en gommant les bords de feuille, les superpositions, et en corrigeant les diffé-



rences de saturation de couleur, de contraste, pour ne faire qu'une seule image continue et homogène.

Autant dire que peu d'utilisateurs de SIG se sont lancés dans ces productions en dehors d'organismes dont c'est le métier (IGN, producteurs privés...).

Depuis quelques temps, la mise à disposition d'outils d'ortho-rectification laisse la possibilité d'effectuer cette opération pour des utilisateurs moins chevronnés. Cependant ceci restera limité à de petites productions.

Le stockage

La principale difficulté technique soulevée par l'utilisation des données raster était la

consommation importante d'espace mémoire, en particulier pour en assurer la conservation (stockage).

L'évolution des techniques de stockage a pratiquement effacé la difficulté. Une planche de cadastre, scannée à 400 points par pouce soit 16 points par millimètre et compressée dans un format classique (CCITT ou TIFF) mesure en moyenne moins de 1 000 Ko. Les planches d'une commune (de 10 à 30 planches dans la plupart des cas) occuperont moins du dixième d'un CD-R (enregistrable), dont le prix d'achat se situe autour de 10 F. Et les 10 à 30 Mo que mesure l'assemblage du plan cadastral d'une commune tiendront sans difficulté sur un disque dur dont les prix continuent à baisser de manière spectaculaire. Un disque de 10 Go dont le prix s'approche des 1 000 F pourrait contenir les planches de toutes les communes d'un département.

MAILLE, PRECISION, RESOLUTION

La taille de la maille est une information importante pour caractériser les données raster. C'est même souvent la seule information caractérisant la finesse de la donnée.

Or, la taille de la maille ne donne d'indication que sur l'agrandissement maximum possible avant apparition des mailles (voir figure 3). En aucun cas la taille de la maille ne permet d'estimer une précision ou une résolution de la donnée.

Ainsi, une carte au 1 :25 000 peut être scannée à 300 points par pouce soit 12 points par millimètre ou à 1 200 soit 50 points par millimètre, sans modifier la précision du résultat,

qui est celle de la carte d'origine, parfois dégradée par le scannage. De même, une orthophotographie peut avoir une maille de 0,50 m au sol, mais n'avoir une précision annoncée que de quelques mètres ; cette dernière découle en effet de la rectification géographique indispensable pour passer de la photo à une donnée géographique.

Par ailleurs, la résolution, entendue comme la possibilité de différencier des objets sur l'image, sera très variable selon le type de donnée maillée. Sur une carte scannée avec un pas de numérisation adapté, la résolution sera celle de la carte et pas plus. Sur une photo aérienne, elle sera dépendante du contraste entre deux objets limitrophes, et de leur forme : on peut voir sur une image à maille au sol de 0,50 m la signalisation horizontale sur une route, alors qu'elle a une largeur bien inférieure à la maille. A l'inverse, la frontière entre deux objets ne présentant qu'un faible contraste, et aux limites non régulières, ne pourra être déterminée qu'avec au moins deux mailles.

Par rapport aux données vecteur, on peut donc limiter l'échelle d'affichage des données maillées en jouant sur la taille de la maille. Cependant, les difficultés dues à l'utilisation de données de précision ou de résolution trop différentes restent les mêmes que pour des données vecteurs.

CONCLUSION

Toutes les difficultés qui viennent d'être évoquées ne doivent pas faire oublier les avantages majeurs que représente la complémentarité des informations raster par rapport à des informations sous forme vecteur.