

## **Alpage : réunion du groupe SIG à La Rochelle le 21/12/2006**

Présents : Anne-Laure Bethe, Romain Raveaux, Frédéric Pouget, Frédéric Rousseaux, Laure Paradis, Jean-Christophe Burie, Hélène Noizet.

### **1) Questions administratives**

Les notifications officielles de l'ANR ont bien été reçues par les 3 laboratoires financièrement reconnus par l'ANR comme entités dépendantes (Lamop, Otelo, L3i). La date de début du projet a été fixée par l'ANR au 6 novembre 2006. Comme le programme dure 3 ans, nous travaillerons ensemble jusqu'en novembre 2009.

Guillaume Champlong ayant décidé de travailler pour un autre projet, c'est Anne-Laure Bethe qui est finalement recrutée dans le cadre du programme pour la partie géographique (géo-référencement, montage du SIG, intégration des données des historiens). Nous serons amenés à raccourcir un peu la durée de son contrat afin qu'elle bénéficie d'un salaire équivalent à celui qu'elle percevait précédemment : elle sera employée de janvier 2007 à juillet 2009. C'est Laure Paradis qui prendra le relais pour le travail à effectuer au-delà, c'est-à-dire durant l'automne 2009, notamment pour l'intégration des données rendues par les historiens.

### **2) Géoréférencement des rasters Vasserot**

- Les géographes ont fait des essais de géo-référencement sur 4 images, notamment dans l'ancien 12<sup>e</sup> arrondissement (du côté de la rue Galande, actuellement le 5<sup>e</sup> arrondissement). Même si cela peut être très variable, selon que le rapport avec le parcellaire actuel se fait plus ou moins facilement, il faut compter environ une heure pour géo-référencer une image, en prenant une dizaine de points de calage. Pour trois de ces images géoréférencées (du côté de la rue Galande), le résultat était assez bon puisque le RMS (Root Mean Square), qui est l'indice de la déformation liée à la transformation de l'image pour la géo-référencer, est de 0,5 ou de 0,8, soit une déformation due au géo-référencement de 50 ou de 80 cm. Mais ailleurs, du côté des Champs Elysées, la déformation était beaucoup plus grande car il y avait très peu d'accroches possibles par rapport au parcellaire actuel, en raison de la profonde transformation de cet îlot entre 1830 et 2006 : dans ce cas, le RMS était environ de 20, ce qui est énorme. Cet indice est fourni par ArcGis comme une valeur moyenne par image raster recalée, mais on peut obtenir cet indice pour chacun des points de calage, ce qui permettra d'avoir une meilleure appréciation de la déformation, qui n'est pas répartie uniformément sur l'image. En revanche, cet indice ne permet pas de mesurer le décalage par rapport à la réalité, c'est-à-dire que les déformations entre la réalité et l'image raster de base, dues au travail de notre cher P. Vasserot, ne peuvent pas être prises en compte par le RMS. A priori, seuls des relevés précis de coordonnées à effectuer sur le terrain (avec un GPS), pour des bâtiments figurés sur les images et encore présents en élévation aujourd'hui, permettront de connaître ponctuellement la déformation de notre information parcellaire par rapport à la réalité. Cette information est capitale à connaître car elle permettra de fixer les échelles possibles d'analyse et les seuils en-dessous desquels on ne pourra pas utiliser cette source planimétrique que nous sommes en train de construire : il s'agit ainsi de connaître les limites de notre information parcellaire, limites inhérentes à la qualité des images raster de base. Il est envisageable de définir ces relevés comme des exercices à faire par des étudiants, notamment ceux de l'école d'architecture de Versailles, grâce au contact de L. Costa.

- Il faut noter que les noms des rues indiquées sur chaque plan reste lisible même après la déformation du géo-référencement.

- Afin de faciliter le géo-référencement, F. Pouget demande s'il est possible d'obtenir des photographies aériennes précises (avec un pixel pour 50 cm), notamment la BD Ortho. Il faudra voir s'il est possible de l'obtenir par convention car elle coûte très cher.
- De même, un élément fort utile serait de connaître la date de construction des immeubles du parcellaire actuel, pour savoir tout de suite si un édifice quelconque est ancien et peut être éventuellement utilisé comme point de calage ou s'il faut au contraire ne pas en tenir compte dans le géo-référencement. Le travail effectué par F. Cohen et M. Guelton, qui ont réalisé en 2004 une étude, comme chargées de mission à l'APUR, sur la cartographie historique de Paris comme indicateur de potentiel patrimonial, serait certainement très utile à cet égard.
- Dans la mesure où les rasters sont géoréférencés, et que l'information des noms de rues restent lisibles sur les rasters recalés, il est décidé de fournir aux historiens non pas les fichiers issus de la vectorisation automatique, ce qui prendrait trop de temps, mais ces rasters recalés et assemblés par arrondissement. En effet, les historiens n'ont pas besoin dans un premier temps de travailler sur les fichiers vecteurs eux-mêmes : au contraire, utiliser les images Vasserot sera plus commode que les fichiers vecteurs, grâce notamment au nom de rues qui permettront de localiser les lieux beaucoup plus vite. Dans un premier temps, le géo-référencement et la vectorisation se feront de front, puis les fichiers vecteurs seront récupérés par les géographes, pendant que les historiens travailleront sur leurs images Vasserot assemblées par arrondissement.
- Pour cela, il faut, préalablement au géo-référencement, découper les images pour ne garder que le plan de l'îlot renseigné (c'est-à-dire dont on a le parcellaire), sinon les bords chaque plans d'îlot se superposent les uns sur les autres et rendent illisibles les franges des îlots. A priori,
- Il est décidé que dès qu'un arrondissement est traité par les géographes, il sera donné aux historiens qui pourront ainsi commencer à travailler au fur et à mesure des livraisons, sans attendre d'avoir l'ensemble. Du point de vue du calendrier, il est souhaitable que l'ensemble des rasters soient géoréférencés d'ici le mois de juin, afin de fournir ces documents aux historiens pour qu'ils puissent travailler sur l'ensemble du plan dès l'été prochain.
- Etant donné la lourdeur d'une image d'un seul îlot, il ne sera pas possible de donner aux historiens une seule image par arrondissement, mais il faudrait arriver à limiter le nombre d'images pour un arrondissement à environ 4-5 pour que cela reste aisément manipulable et que les historiens n'aient pas besoin d'ouvrir de trop nombreux fichiers : par exemple, le 12<sup>ème</sup> ancien arrondissement comporte 120 plans par îlots ; une fois géoréférencés, il s'agit donc de fournir aux historiens un assemblage de ces plans d'îlots sous la forme de 4 ou 5 images. F. Pouget demande s'il est possible que les informaticiens diminuent la taille des images rasters, par exemple en passant de 36 millions de couleurs à 256, ce qui semble faisable. Ainsi allégés, ces fichiers seront plus aisément manipulables.
- F. Pouget demande si les informaticiens du L3i pourraient nous développer une application du type du logiciel PCI image développé pour la DGI, et qui permet de naviguer facilement parmi les rasters géoréférencés, ce qui semble faisable.
- Une fois les rasters géo-référencés et la vectorisation parcellaire terminée, les géographes devront récupérer les fichiers vecteurs : afin de ne pas recommencer le travail de géo-référencement, qui aura déjà été fait pour les rasters, il faudra trouver le moyen d'enregistrer la transformation appliquée à chaque raster pour la ré-appliquer ensuite aux fichiers vecteurs, ce qui devrait être possible avec ArcInfo selon F. Pouget.
- Le format de sortie des rasters recalés et assemblés destinés aux historiens peut être du jpeg ou de l'ecw, ce dernier format n'étant pas lisible par Illustrator. Le choix du format de sortie dépendra donc du logiciel utilisé par les historiens pour transcrire leurs couches historiques.
- Les géographes demandent s'il est possible de proposer aux historiens une petite journée de formation à des outils très simples de SIG et gratuits, du type OpenJump ou QGIS, afin qu'ils saisissent leurs objets spatiaux dans ces logiciels, ce qui permettrait ensuite de récupérer plus

facilement leurs données et de les intégrer plus rapidement dans le SIG. Les géographes se proposent de présenter ce type d'outil aux historiens et de travailler vraiment avec eux pour la saisie des premiers objets spatiaux afin de répondre à toutes les questions qui ne manqueraient pas de surgir. Cela dit, pour les historiens qui souhaiteraient malgré tout travailler sur Adobe Illustrator, cela restera possible (il faudrait simplement qu'ils donnent leurs fichiers illustrator en format DXF), de même que pour ceux qui souhaiteraient travailler non pas numériquement, mais directement sur le papier. H. Noizet fera le point sur ces besoins à la prochaine réunion des historiens le 18 janvier.

### 3) Vectorisation des rasters

- R. Raveaux présente le début de son travail concernant l'analyse des plans cadastraux dans un but de vectorisation du parcellaire. La vectorisation du Parcellaire s'inscrit dans une démarche complète de rétro conversion de document graphique, soit le passage d'un document papier à son interprétation sémantique. Dans notre contexte, nous possédons les plans déjà numérisés, des images brutes composées de pixels. Tout l'enjeu de la vectorisation est alors de donner du sens à ces images qui ne sont à ce stade que des amas de données inertes et de formaliser cette connaissance sémantique sous forme vectorielle. Dans cet objectif, il est nécessaire de mettre en exergue l'information contenue dans les images de plans cadastraux.

#### *Améliorer la qualité de l'image*

Nos premières expériences ont montré un affadissement des plans lié au temps. Les pigments de la peinture se sont altérés. Afin de pallier ce problème et de revenir vers la couleur originelle du document, une stratégie de restauration d'image a été mise en place.



Figure 1 : A gauche : avant restauration. A droite : après restauration.

#### *Segmentation couleur*

L'information couleur prend une place prépondérante dans les plans cadastraux de Vasserot car elle caractérise l'entité parcelle. En effet, à une parcelle est associée une couleur. Les techniques de segmentation couleur visent à localiser des ensembles chromatiques afin d'identifier les parcelles présentes sur un plan.

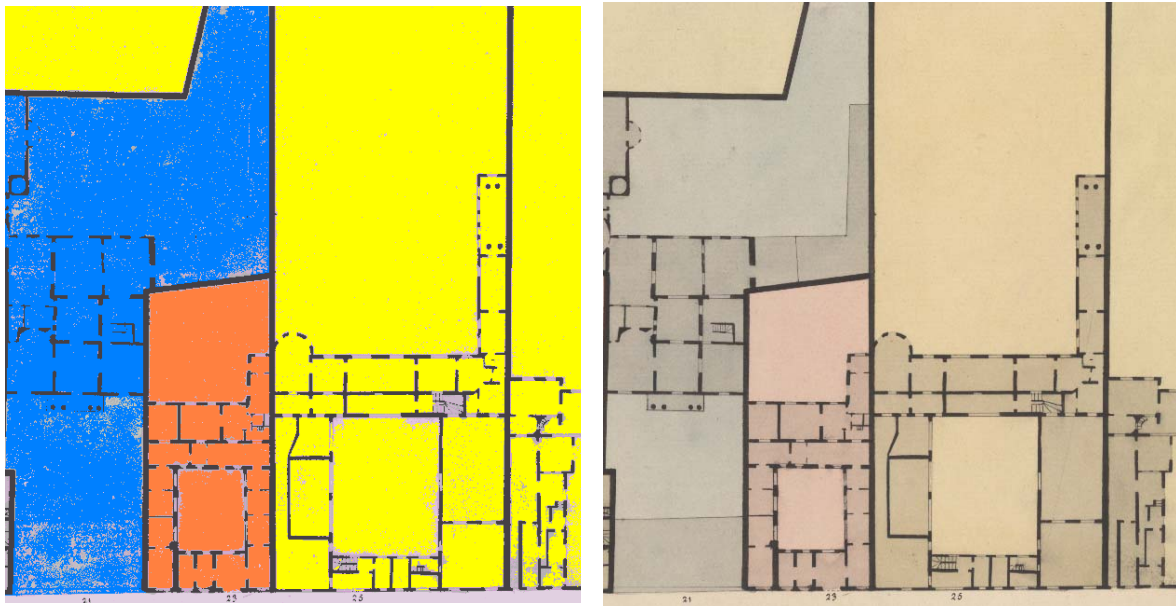


Figure 2 : A gauche un morceau de plan segmenté. A droite : le plan de départ.

Le travail concernant la segmentation couleur est loin d'être trivial, il existe une large panoplie de couleurs et une grande variabilité au sein d'un même ensemble chromatique. Ce domaine fait par conséquent l'objet de travaux importants.

- F. Pouget demande si les informaticiens du L3i peuvent appliquer un pré-traitement sur les images raster afin d'améliorer la netteté et le contraste des images, par exemple en utilisant une des procédures présentées (théorie du monde blanc, affadissement). Ce sont sur ces images pré-traitées que les géographes pourraient travailler afin d'améliorer notablement leur confort de travail, ce qui n'est pas superflu dans la mesure où il y a environ 930 images à géoréférencer.

#### 4) Couches historiques et modèle conceptuel de données

- Un compte-rendu de la réunion tenue avec les historiens et archéologues le 27 novembre dernier est fait par H. Noizet. Une diversité des utilisations possibles de l'outil émerge d'ores et déjà : un premier niveau relativement simple, pour lesquels les entités fonctionnelles sont suffisantes, et un second niveau plus poussé pour lesquels les entités spatiales sont nécessaires. Il faudra tenir compte de ces différents besoins pour établir le MCD du SIG. Pour cela, il est prévu de fixer une réunion à Paris, dans la dernière semaine de mars, afin de discuter spécifiquement du MCD du SIG, avec les géographes d'Otelo et les Parisiens compétents en la matière, notamment les géomaticiens d'ArScAn (mais la présence des informaticiens du L3i n'apparaît pas nécessaire).

- Pour le filaire des rues, il faudra tenir compte du travail déjà effectué par C. Bourlet sur l'identification des tronçons de rues, pour qu'on puisse faire le lien avec sa base de données, notamment en ce qui concerne les noms de rues anciens et actuels.

- Pour le plan de Paris du 14<sup>e</sup>, qui va en fait se traduire par deux plans successifs (un vers 1300 et un vers 1400) grâce au travail de C. Bourlet et N. Thomas, il doit être possible, selon les géographes, de récupérer le fichier vectorisé par les archéologues et de géoréférencer les îlots par aimantation de certains points (notamment les points aux angles des îlots) sur les îlots du parcellaire Vasserot. Evidemment, ces plans d'îlots du 14<sup>e</sup> n'auront pas du tout la même valeur géographique que le parcellaire Vasserot, dans la mesure où certains tracés sont déduits de sources strictement textuelles

*Alpage, réunion du groupe SIG à La Rochelle - 21/12/2006*

(et non planimétriques) et sont donc localisés un peu au pifomètre. Ce sera aux utilisateurs du SIG d'être conscient du fait qu'il s'agit de plans reconstitués à partir de sources textuelles et donc de ne pas surinterpréter ces plans.

Bonnes fêtes à tous !