

PROBLÉMATIQUES DE RECHERCHE AUTOUR DE LA NOTION DE CLUTTER EN CARTOGRAPHIE AUTOMATIQUE

par Guillaume Touya, Sidonie Christophe, Charlotte Hoarau

Institut national de l'information géographique et forestière – Laboratoire COGIT

73 avenue Pasteur 94165 Saint-Mandé

prénom.nom@ign.fr

L'effet de clutter intervient lorsqu'une carte comporte une quantité excessive d'information, quand cette information est désorganisée dans la carte. Il est indispensable de savoir mieux mesurer cette complexité afin de contrôler la qualité des cartes produites par des processus automatiques. Cet article rend compte de mesures existantes du clutter ou de la complexité, issues de différents domaines de recherche, et pointe le manque de méthodes adaptées à la conception automatique de cartes. Trois cas d'étude sont présentés pour illustrer le type de mesures requis pour améliorer l'automatisation de la généralisation cartographique, et de la spécification du style cartographique. Des verrous scientifiques sont déduits des problèmes pour traiter les cas d'étude avec les connaissances actuelles sur la complexité. Il en ressort notamment qu'une seule mesure ne peut pas saisir tous les aspects de la complexité des cartes et que la combinaison de mesures devrait être étudiée.

Introduction

Les données géographiques et les cartes sont aujourd'hui de plus en plus utilisées par le grand public ou des professionnels pour des tâches diverses, qui sont souvent limitées par la complexité de ces cartes. La recherche en cartographie s'est très tôt intéressée à la question de la complexité des cartes (McEachren 1982, Knopfli 1983). Mais il est toujours très difficile de mesurer la corrélation entre la complexité et la lisibilité d'une carte et son efficacité pour la réalisation d'une tâche donnée : par exemple, quand est-ce qu'une carte devient trop complexe pour qu'un utilisateur lambda puisse calculer un itinéraire ? Notre objectif de recherche à long terme est de maîtriser des indicateurs pour évaluer la qualité d'une carte : les techniques de cartographie automatique nécessitent d'utiliser des mesures pour s'autoévaluer (Stoter *et al* 2014), et la complexité visuelle est une piste privilégiée pour qualifier des cartes comme adaptées aux capacités cognitives ou de perception d'un utilisateur. La complexité des cartes est un des facteurs qui permettent d'estimer si une carte est lisible, efficace et adaptée à un besoin donné. Il s'agirait en fait d'un indicateur de qualité pour permettre de maintenir un niveau de complexité visuelle acceptable, mais aussi d'un indicateur de bruit ou de surcharge cognitive dans une carte rendant difficile l'utilisation de la carte. Par conséquent, il est intéressant de savoir si les mesures existantes de la complexité peuvent nous aider à mieux adapter les processus de cartographie automatique aux besoins des utilisateurs. La question de la complexité des cartes a déjà été abordée par différentes communautés de recherche, avec souvent des termes différents pour la désigner, mais il manque des recherches qui tentent

d'unifier toutes ces propositions. Par exemple, les chercheurs en vision par ordinateur utilisent le mot *clutter* pour parler de la quantité excessive d'information dans une image (Rosenholtz *et al* 2007). Il est ainsi intéressant d'étudier la profusion de méthodes de mesures du *clutter* en vision par ordinateur (Moacdieh et Sarter 2014) pour identifier des mesures transposables à l'estimation de la complexité d'une carte : en quoi ces méthodes dédiées aux images en général sont-elles pertinentes pour les cartes vues comme des images (et non une superposition de couches vectorielles) ? Quelle particularité des images-cartes nécessiterait une adaptation des méthodes génériques ? L'objectif des recherches présentées dans cet article est de définir des verrous scientifiques, un agenda de recherche, sur l'utilisation des méthodes de mesure de la complexité pour améliorer la qualité des processus de cartographie automatique comme la généralisation cartographique ou l'adaptation automatique du style des cartes.

Pour atteindre cet objectif, l'article fait un état de l'art de propositions passées pour mesurer la complexité dans les cartes ou les images. Nous analysons ce qui est utile pour traiter des scénarios de cartographie automatique nécessitant une mesure de la complexité, et nous pointons les manques actuels. La deuxième partie de l'article présente des définitions et l'état de l'art sur les méthodes existantes. La troisième partie détaille nos trois cas d'étude : la généralisation cartographique, la visualisation de données géographiques hétérogènes, et la spécification du style d'une carte. Enfin, la dernière partie propose des verrous scientifiques dérivés de nos cas d'étude.

Méthodes existantes pour mesurer la complexité

Complexité, entropie, lisibilité et clutter

De nombreux chercheurs ont étudié la complexité des cartes, utilisant un vocabulaire souvent différent, dont les définitions se recoupent en partie : complexité, entropie, *clutter* ou simplement lisibilité.

Tout d'abord, la **complexité** d'une carte est un terme couramment utilisé dans la communauté des chercheurs en cartographie. La plupart des définitions proposées de la complexité s'accordent sur la distinction entre complexité visuelle ou graphique, et la complexité intellectuelle (McEachren 1982, Fairbairn 2006, Jégou et Deblonde 2012, Ciolkosz-Styk et Styk 2013), même si chacune de ces deux notions n'est pas parfaitement définie (fig. 1). La complexité visuelle est liée à la difficulté pour le lecteur de la carte à percevoir l'ensemble des symboles de la carte, alors que la complexité intellectuelle est liée à la difficulté pour le lecteur de la carte à comprendre la signification des symboles de la carte (Fairbairn 2006). Fairbairn (2006) explique que la complexité intellectuelle a été moins étudiée en cartographie car elle est intimement liée au lecteur de la carte et à la tâche accomplie par le lecteur avec la carte, facteurs éminemment divers. Jégou et Deblonde (2012) ajoutent un troisième niveau de complexité, appelé complexité sémiologique, à l'interface des complexités visuelle et intellectuelle. Selon eux, les trois niveaux de complexité sont liés à des fonctions cognitives : la complexité visuelle est liée à la perception pré-attentive des formes, textures et couleurs, la complexité sémiologique est liée à la perception attentive et consciente des signaux graphiques, et la complexité intellectuelle est liée à la compréhension des structures spatiales de l'espace géographique.

Il existe une façon de combiner complexité visuelle et intellectuelle, en mesurant l'**entropie** d'une carte (Knopfli 1983, Bjorke 1996, 2012, Li et Huang 2002). L'entropie fait référence à la théorie de l'information de Shannon et Weaver (1949), et caractérise l'incertitude d'une source d'information, ici la carte. Plus la carte tente de véhiculer un grand nombre d'informations, plus l'entropie est grande. En d'autres termes, l'entropie indique la quantité et l'organisation de l'information nécessaires pour transmettre l'information de la carte sans ambiguïté.

Dans le domaine de la vision par ordinateur, le concept de **clutter**, qui n'a pas d'équivalent en français, fait référence à une quantité excessive d'informations dans une image, ou la désorganisation de l'information dans une image (Rosenholtz *et al* 2005, 2007). D'après Rosenholtz, l'entropie est liée au *clutter* : quand l'entropie est faible, plus d'information peut être interprétée par le cerveau humain. Dans la plupart des travaux de recherche, le *clutter* est mesuré pour une tâche donnée,

par exemple l'identification d'un objet dans l'image (Moacdieh et Sarter 2014), ce qui explique la profusion de mesures différentes dans le domaine.

D'autres travaux de recherches s'intéressent à la **lisibilité** des cartes (Chesneau 2007, Stigmar et Harrie 2011, Harrie *et al* 2011, Olsson *et al* 2011, Touya 2012, parmi d'autres). Ce concept est assez proche de la complexité visuelle, car il se fonde sur les limites de perception de l'œil humain pour définir des légendes cartographiques lisibles. En cartographie, ces règles proviennent principalement des règles de conception décrites par Bertin (1967) à propos de l'utilisation des variables visuelles. Par exemple, deux symboles identiques devraient être séparés de 0,1 mm sur la carte pour qu'on puisse les distinguer l'un de l'autre. D'autres règles portent sur l'utilisation des palettes de couleurs pour préserver les relations (association, ordre, différence) entre les données (Chesneau 2007, Christophe 2011).

Parmi tous ces concepts, nous choisissons celui de *clutter*, car nous considérons que sa définition couvre les autres termes. Mais pour l'instant, il est difficile de discriminer quels sont les aspects de la complexité couverts par une mesure du *clutter* donnée. Des études doivent être menées pour déterminer des seuils de clutter dans divers contextes, pour déterminer les conditions de validité d'une mesure donnée, et pour contrôler comment le *clutter* est perçu, comment il influence la performance dans des tâches cartographiques, comment il peut causer des effets d'encombrement et d'occlusion, comment il peut pénaliser les tâches de recherche visuelle dans une carte, ou comment il agit sur les limites de la mémoire à court terme (Miller 1956).

Propriétés principales des mesures existantes du clutter

Les mesures existantes du clutter proviennent essentiellement du traitement d'image, où la carte est vue comme une image et non comme un empilement de couches vectorielles symbolisées (voir Ciolkosz-Styk et Styk 2013, ou Moacdieh et Sarter 2014 pour un état de l'art exhaustif), et elles sont souvent fondées sur la détection de contours ou sur des techniques de compression d'image (Da Silva *et al* 2011). Fairbairn (2006) par exemple liste plusieurs techniques de compression pour établir un lien entre quantité d'information et complexité, dans le contexte de la conception de cartes. Ces méthodes posent un problème majeur : elles sont très sensibles à la résolution de l'image alors que la résolution influence peu la perception de la complexité (Jégou et Deblonde 2012). Par opposition, il existe assez peu de méthodes utilisant des données vectorielles cartographiques (MacEachren 1982, Bjorke 1996, Olsson *et al* 2011), mais elles paraissent plus à même de mesurer les occlusions et superpositions qui affectent la lisibilité des cartes.

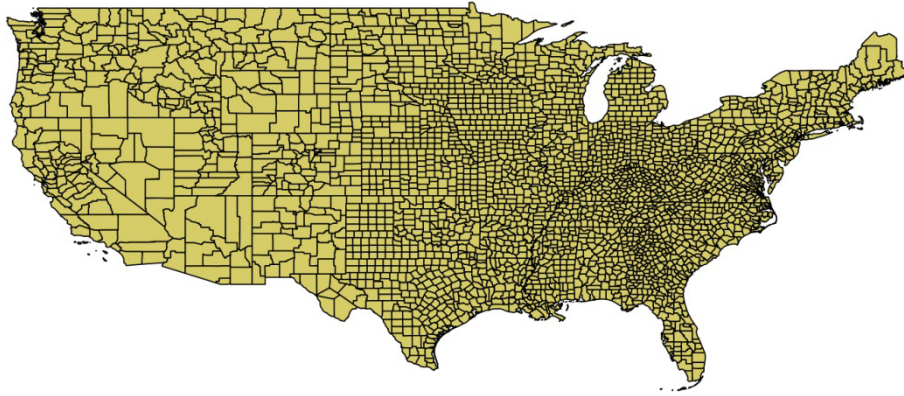


Figure 1 : Cette carte des limites de comtés aux Etats-Unis a une forte complexité visuelle mais une faible complexité intellectuelle.



Figure 2 : (a) faible entropie topologique pour les symboles de points d'intérêt (par exemple les édifices religieux sont proches les uns des autres). (b) forte entropie topologique pour les symboles de points d'intérêt qui sont distribués partout dans la carte.

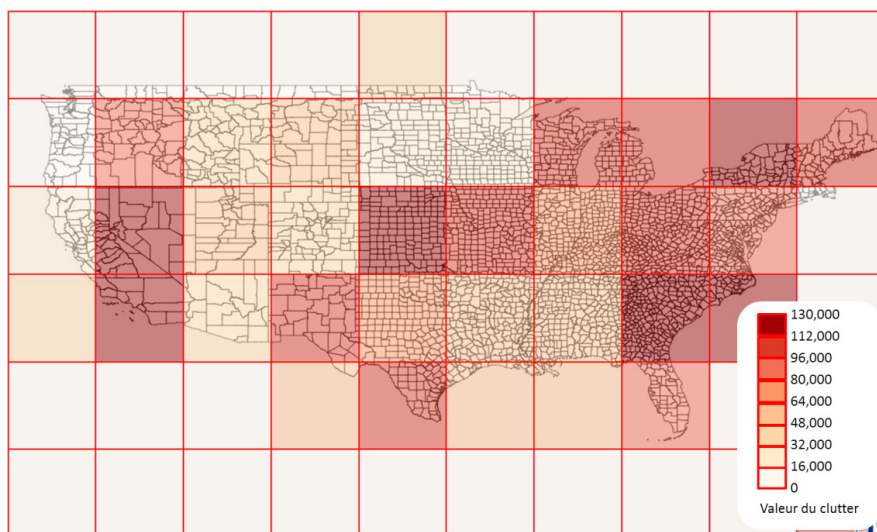


Figure 3 : Visualisation locale du clutter grâce à une grille sur la carte des comtés américains.

Il existe également des mesures globales et des mesures locales du *clutter*. Les mesures globales donnent une valeur numérique globale du *clutter* pour toute l'image ou toute la carte, et sont les plus fréquentes dans la littérature. Par exemple, les mesures fondées sur l'entropie (Bjorke 1996, 2012, Li et Huang 2002) évaluent de manière globale le désordre dans l'information véhiculée par la carte. Bjorke (1996) évalue ce qu'il appelle l'entropie topologique en analysant les régularités dans le voisinage des symboles de la carte. Par exemple, dans la carte de la fig. 2a, les symboles peuvent être groupés par types : les symboles d'édifices religieux sont tous situés dans la même zone, et l'entropie topologique est faible. Dans la carte de la fig. 2b, tous les symboles sont distribués aléatoirement. Dans ces cas, l'entropie est induite par la géographie plutôt que par des choix de symbolisation.

A l'inverse, les mesures locales du *clutter* évaluent les parties de la carte les plus et les moins complexes. Par exemple, Jégou et Deblonde (2012) utilisent une structure de type *quad tree* pour identifier des zones plus ou moins grandes où les couleurs sont uniformes. La mesure de congestion de Rosenholtz *et al* (2007) utilise la métaphore du bureau pour évaluer le *clutter* dans une image : un bureau est bien rangé tant que l'on peut ajouter un objet sur le bureau et qu'il reste facilement identifiable. La congestion analyse la saillance locale (Itti *et al* 1998) des couleurs, du contraste de luminance et de l'orientation, et combine des mesures pour des voisinages de différentes tailles. Ces mesures globales ou locales sont complémentaires, car les mesures locales permettent de repérer les parties d'une carte qui ne sont pas lisibles, alors que les mesures globales donnent une idée générale de la lisibilité de la carte. Mais quand une mesure globale est nécessaire, des mesures locales peuvent être agrégées, et à l'inverse, des mesures globales peuvent « localisées ». Par exemple Jégou et Deblonde (2012) proposent simplement de caler une grille sur la carte et d'appliquer une méthode globale à chaque maille de la grille. La fig. 3 montre comment nous avons découpé la carte des comtés américains pour appliquer localement la mesure globale de densité de contours (Rosenholtz *et al* 2007). Mais ces solutions sont très sensibles à la taille de la grille employée.

Problèmes persistants avec les mesures du clutter

Il existe déjà en cartographie des critères et méthodes pour mesurer la lisibilité mais ils sont souvent très spécifiques à un processus donné (généralisation cartographique, spécification de légende, placement de toponymes par exemple), à des données particulières, à des structures spatiales données ou à des objectifs bien précis. Il y a un manque de méthodes de portée un peu plus générique. Par ailleurs, il y a un manque de méthodes locales et la lisibilité globale d'une carte est souvent complexe à interpréter : les origines de la faible lisibilité

d'une carte sont difficiles à discriminer. Par conséquent, il reste difficile d'automatiser le recours à des méthodes globales dans des processus de cartographie automatique visant à optimiser la lisibilité de la carte. Afin d'améliorer les processus de cartographie automatique, des méthodes locales de mesure de la complexité seraient très utiles :

- pour identifier les zones denses en information, afin de mieux généraliser ces zones, ou d'optimiser les choix de symboles dans ces zones ;
- pour identifier les zones moins denses en information, afin de les utiliser comme espace libre pour ajouter de l'information dans la carte (comme des toponymes par exemple).

Des méthodes de mesures du *clutter* seraient très utiles si elles étaient applicables non seulement sur l'image de la carte, mais aussi sur des couches précises de la carte, sur le premier plan ou l'arrière-plan, quelles que soient les données, ou entre du texte et le fond de carte (Raposo et Brewer 2013, 2014). Enfin, nous avons besoin de mieux contrôler le niveau de complexité d'un espace géographique donné et de ses structures spatiales, pour mieux préserver ces structures et contrôler le niveau de détail dans les processus de cartographie automatique (*e.g.* généralisation, stylisation en aquarelle, conception de carte hybride, voir dans la partie suivante).

Cas d'étude en cartographie automatique

Dans cette partie, nous présentons plusieurs exemples de processus de cartographie automatique qui nécessitent d'une manière ou d'une autre l'utilisation de mesures du *clutter*. Ces besoins en mesure du *clutter* sont illustrés en calculant le *clutter* avec la mesure très simple de la densité de contours, issue de Rosenholtz *et al* (2007). Cette mesure est globale et prend une image comme source de données. Elle détecte les contours dans l'image, c'est-à-dire les limites entre deux pixels présentant une forte variation d'intensité, et en calcule la densité dans toute l'image. Rosenholtz *et al* (2007) affirment que cette mesure est simple mais qu'elle donne une valeur assez correcte du *clutter* dans une image.

Complexité et généralisation cartographique

Les processus automatiques de généralisation cartographique reposent souvent sur une évaluation de la lisibilité de la carte (avant et après généralisation), pour optimiser la qualité de la carte produite (Bard 2004). Les techniques actuelles d'évaluation utilisées dans les processus de généralisation reposent sur une grande quantité de mesures locales et ciblées (par exemple vérifier si l'aire d'un bâtiment est assez grande pour que le bâtiment soit lisible), pour déduire la lisibilité globale de la carte (Stigmar et Harrie 2011,

Touya 2012). Stigmar et Harrie (2011) affirment que des mesures du *clutter* aideraient à identifier les zones trop denses en information, mais aideraient également à vérifier que la quantité globale d'information dans la carte a été préservée, ce qui est un objectif de la généralisation. Afin d'illustrer ce propos, nous avons appliqué la mesure par densité de contours sur les cartes utilisées par Touya (2012), et ces tests montrent que la généralisation, même quand elle est bien faite, augmente toujours le *clutter* ainsi mesuré (fig. 4). Cette augmentation est principalement due à l'élargissement des bâtiments pour les rendre lisibles, ce qui réduit mécaniquement les zones contenant des pixels uniformes blancs ou verts. La généralisation nécessite clairement une mesure locale pour différencier les zones peu complexes des zones très complexes où la généralisation a échoué, mais aussi une mesure prenant bien en compte la réduction des occlusions et des superpositions par le processus de généralisation.

Complexité pour l'intégration et la visualisation de données hétérogènes

Une grande quantité de données géographiques variées et hétérogènes sont maintenant disponibles pour des utilisateurs souhaitant les analyser ou les visualiser conjointement. L'hétérogénéité de ces données, en termes de type (images, vecteurs, modèles 3D, etc.), de schémas de données, de niveau de détail, ou de propriétés visuelles (contrastes de couleur, textures, etc.), nécessite l'utilisation de processus cartographiques adaptables. Nous illustrons ici ce problème avec deux exemples, la production de cartes à partir de données collaboratives, et la visualisation hybride d'ortho-imagerie et de cartes.

Conception de cartes à partir de données collaboratives

Cartographier des données collaboratives comme celles issues d'OpenStreetMap peut s'avérer très complexe car les données ont été saisies par des utilisateurs différents et disposant de sources de données différentes, ce qui a engendré une forte hétérogénéité du niveau de détail de ces données (Touya et Reimer 2015). Connaître le niveau de détail des objets à cartographier est indispensable pour déterminer à quelles échelles ces objets peuvent être représentés de manière lisible. Touya et Brando (2013) font l'hypothèse qu'une mesure du *clutter* permettrait de mieux estimer le niveau de détail de certains de ces objets issus d'une saisie collaborative hétérogène, pour lesquels la géométrie et la sémantique ne suffisent pas. Par exemple, dans la fig. 5, des chemins de randonnée issus d'OpenStreetMap sont symbolisés au 1 : 20 000 et au 1 : 50 000. L'estimation du niveau de détail à partir de la géométrie et de la sémantique nous indique que ces chemins peuvent être représentés au 1 : 50 000 mais la figure nous montre qu'une représentation à cette

échelle n'est pas lisible du fait de la trop forte densité de chemins dans la zone. La mesure de densité de contours nous indique que le *clutter* est deux fois plus important au 1 : 50 000 qu'au 1 : 20 000. La différence pourrait être plus importante si la mesure ne tenait pas compte du fond de carte qui reste lui inchangé aux deux échelles. Dans ce cas, nous aurions besoin d'une mesure locale et seulement appliquée au thème de la carte qui nous intéresse.

Visualisations hybrides entre images et données vectorielles

Mixer des données géographiques hétérogènes, comme des données vectorielles symbolisées et des ortho-images pour les visualiser conjointement pose des problèmes de lisibilité spécifiques : la quantité d'information contenue dans l'image avec en plus la superposition de symboles rend ces visualisations difficiles à lire. Hoarau et al (2013) et Hoarau et Christophe (2015) proposent des mécanismes pour visualiser de telles données via un continuum de représentations plus ou moins réalistes allant de la carte seule à l'image seule, en jouant sur les couleurs, les textures et la transparence. La fig. 6 montre plusieurs possibilités pour optimiser la co-visualisation des routes sur une ortho-image, en jouant sur les contrastes entre les couleurs dans l'image, la couleur de la route, et la couleur du bord de son symbole. La solution sans bord en haut à gauche est celle privilégiée par les géoportails existants, mais nous préférons les solutions avec un bord gris ou un bord aux couleurs adaptatives. Nous souhaitons évaluer si une mesure du *clutter* pourrait appuyer ou non ces hypothèses.

Une mesure locale du *clutter* serait utile dans ce contexte pour deux raisons :

- Une mesure du *clutter* dans l'ortho-image avant co-visualisation pourrait permettre d'identifier les zones complexes de l'image, comme les zones boisées. Cette identification permettrait soit de réduire cette complexité locale, par exemple par des outils de traitement d'image, soit de copier cette complexité avec des textures pour des représentations intermédiaires peu différentes de l'image en termes de complexité. Nous pourrions également utiliser une adaptation des couleurs (comme dans la fig. 6) spécifiques à ces zones complexes.
- Une mesure du *clutter* sur la visualisation hybride pourrait nous permettre d'évaluer les méthodes qui ont permis la construction de cette représentation hybride, comme l'adaptation locale des couleurs présentée en fig. 6.

En testant la mesure de densité de contours sur les images présentées en fig. 6, on remarque qu'elle est très sensible à la texture dans l'image, notamment les textures de la végétation, la rendant peu utilisable pour les opérations décrites au-dessus.



Figure 4 : Carte topographique au 1 : 50 000 non généralisée (à gauche) et généralisée (à droite). Les valeurs de clutter mesurées par densité de contours sont de 429 et 1459.

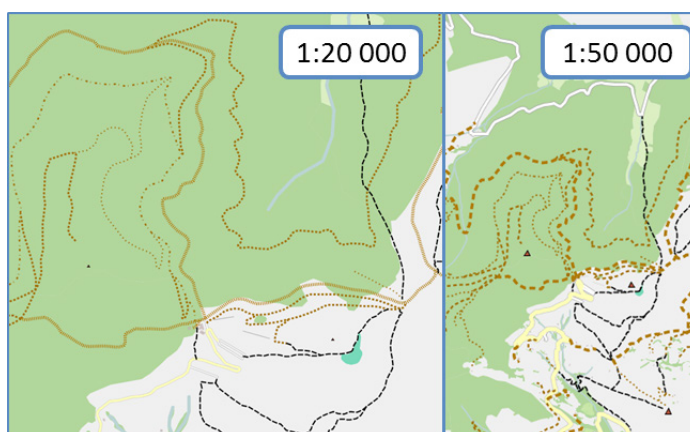


Figure 5 : Les chemins sur la carte de gauche sont géométriquement adaptés à l'échelle du 1 : 50 000 mais leur densité les rend plutôt à des échelles supérieures au 1 : 20 000.



Figure 6 : Différentes méthodes pour adapter le symbole des routes à la photo en optimisant le contraste local.

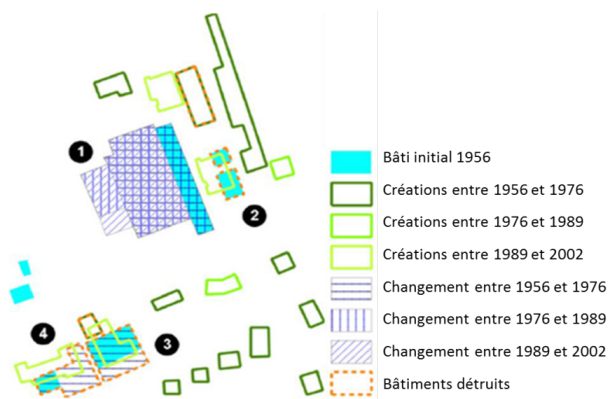


Figure 7 : Légende optimisant la quantité de superposition de symboles pour représenter l'évolution du bâti (Rémy 2014).

Complexité et style topographique

Formaliser, concevoir et interpoler des styles topographiques dans le but d'améliorer le photoréalisme, le niveau de détail, ou d'imiter un style artistique repose sur le respect de contraintes cartographiques préservant la lisibilité des données et des relations spatiales (Christophe 2012, Hoarau *et al* 2013, Hoarau et Christophe 2013, Brasebin *et al* 2015, Ory *et al* 2015, parmi d'autres). Que le contexte soit la conception de cartes personnalisées, ou la production de cartes topographiques en série comme le fait l'IGN, appliquer des styles nouveaux à des cartes en ajoutant des symboles ou effets visuels inhabituels, et on court le risque de rendre le message cartographique confus. Par conséquent, il est important de mesurer la complexité visuelle, sémiologique et cognitive induite par les processus de stylisation. Nous identifions deux objectifs principaux :

- Nous souhaitons préserver un certain niveau de complexité visuelle lors d'un changement de style, considérant que leur niveau de complexité visuelle est lié à leur qualité. Par exemple, basculer d'un style français à un style suisse sur une carte de la frontière (Ory *et al* 2015) nécessite de manipuler la complexité pour maintenir un niveau de complexité cohérent (le style français est plus complexe que le style suisse, ce qui est confirmé par la mesure de densité de contours).
- Nous souhaitons également diminuer le *clutter* induit par l'application d'un style artistique comme l'aquarelle : simplifier les données est indispensable pour appliquer un tel style, mais il augmente tout de même la complexité du fait d'un effet de dilution, et l'aquarelle provoque aussi des phénomènes d'occlusion et de superposition de symboles (Christophe *et al* 2015).

Les mesures de *clutter* pourraient être considérées comme des indicateurs guidant le processus de stylisation dans ses choix, afin de maintenir le niveau de complexité ou de le baisser dans certaines zones. Il faut donc des mesures locales, et un point important ici est de pouvoir déterminer les seuils par des tests auprès d'utilisateurs, car le style personnalisé doit être adapté à la perception par les utilisateurs. De plus, si les mesures locales de *clutter* peuvent identifier des zones peu complexes, ces zones peuvent être utilisées pour y accentuer la stylisation sans rendre la carte trop complexe. Par exemple, dans les cartes topographiques classiques comme celles de l'IGN, les toponymes sont souvent placés dans la mer car la mer est représentée par un aplat simple bleu clair. Mais si la mer est représentée avec des textures imitant les vagues pour la rendre plus réaliste (comme dans les cartes de Cassini), ou avec en transparence une ortho-image pour faire apparaître les rochers immergés, les toponymes doivent être placés ailleurs, dans des zones moins complexes. Si le

clutter est trop fort dans une zone, il peut être ajusté en optimisant les choix de symbolisation. Par exemple, on peut jouer sur les contrastes de couleur entre le fond et les couches de premier plan (fig. 6), ou entre les objets qui se superposent (fig. 7). Cela suppose que l'on puisse adapter localement la légende de la carte (les objets n'ont pas exactement la même couleur partout dans la carte) pour améliorer la lisibilité globale.

La fig. 7 montre des résultats d'une optimisation automatique de choix de symbolisation (ici les types de hachures) pour représenter la dynamique historique de construction et destruction de bâtiments, en minimisant la superposition de symboles (Remy 2014). Là aussi, une détection locale du *clutter* permettrait de mieux adapter cette optimisation de symboles en la concentrant sur les zones les plus complexes en termes de superpositions.

En considérant qu'un haut niveau de complexité n'est pas toujours synonyme de problème de lecture de la carte, et que la complexité n'est pas toujours un facteur à niveaux, mais parfois à préserver, les mesures de la complexité pourraient être utilisées comme des indicateurs de qualité. Les cartes topographiques classiques peuvent être considérées comme complexes, du fait de la quantité d'information qu'elles peuvent contenir et transmettre, mais on ne va pas chercher à diminuer cette complexité mais plutôt à la préserver : si nous essayons de transformer cette carte topographique pour répondre à un besoin donné, comme la stylisation par aquarelle, préserver la quantité d'information transmise est un véritable défi. Par exemple, si un utilisateur cherche à appliquer sur ses données un style topographique comme le style des cartes IGN ou des cartes suisses (Ory *et al* 2013), ou appliquer un style en aquarelle, une mesure pourrait indiquer si le niveau de *clutter* est le même avant et après application du style. L'hypothèse dans ce cas est que le niveau de *clutter* est représentatif de la complexité que nous souhaitons préserver.

Verrous scientifiques liés à la complexité des cartes

Les cas d'étude présentés dans la partie 3 ont montré plusieurs problèmes liés à la mesure de la complexité dans les cartes pour la conception automatique de cartes. Ces problèmes sont reformulés ici sous la forme de verrous scientifiques que les chercheurs en cartographie devraient aborder.

Tout d'abord, l'état de l'art résumé dans la partie 2 montre une forte prédominance des méthodes utilisant des cartes raster pour évaluer la complexité, que les méthodes proviennent de chercheurs en géomatique ou en vision par ordinateur. Les cas d'étude montrent pourtant l'importance fondamentale de l'évaluation de la lisibilité à partir des couches de symboles, à cause

notamment des risques d'occlusion. Est-il alors utile d'utiliser ces méthodes « basées-image » pour évaluer la complexité ? Notre hypothèse est que ces méthodes seront utiles, comme le montre leur utilisation pour le placement automatique de textes sur une carte par Rylov et Reimer (2015).

Nous avons vu qu'il existe une très grande quantité de mesures différentes, il est donc nécessaire d'identifier quelles sont les meilleures pour un problème de cartographie automatique donné. Par exemple, quelles sont les meilleures mesures pour estimer si une carte généralisée contient un niveau de complexité similaire à celui de la carte source dont elle est dérivée ?

Les différentes définitions du *clutter* ou de la complexité présentées dans cet article couvrent différentes propriétés des cartes, et la plupart des méthodes proposées ne mesurent qu'une partie de ces propriétés. La suite logique est donc de se demander comment il est possible de combiner ces mesures pour évaluer de manière plus globale la lisibilité des cartes. Cette combinaison est-elle possible en utilisant des techniques de décision multicritères (Figueira *et al* 2007) ?

Enfin, Moacdieh et Sarter (2014) ont recensé plusieurs méthodes de mesure de la complexité directement liées à la bonne réalisation d'une tâche donnée. Par exemple, ils présentent des mesures pour la densité d'information sur l'écran de contrôle d'un pilote d'avion. Il existe aussi des tâches spécifiques liées à l'utilisation de carte. Existe-t-il des méthodes spécifiques pour ces tâches plus efficaces que des méthodes génériques ? Doit-on concevoir uniquement des méthodes adaptées à une tâche cartographique donnée ?

Plusieurs initiatives peuvent être prises afin d'avancer sur ces questions de recherche. Tout d'abord, nous devrions tester un échantillon représentatif des mesures de complexité existantes sur des cartes issues des cas d'études présentés dans la partie précédente. Ensuite, il est nécessaire de comparer les métriques obtenues par les différentes mesures avec la perception d'un utilisateur de carte, ou avec sa difficulté à réaliser une tâche donnée

à cause de la complexité de la carte. L'objectif est ici de vérifier dans quelle mesure il existe des corrélations entre la difficulté de lisibilité de la carte et les valeurs chiffrées de complexité. C'est ce que Stigmar et Harrie (2011) ont réalisé pour trouver une corrélation entre les métriques d'une carte (par exemple le nombre de symboles, le nombre de thèmes, etc.) et l'évaluation de lisibilité de la carte par un utilisateur. Enfin, il est nécessaire d'explorer comment les méthodes de mesure de la complexité peuvent être combinées, par exemple en utilisant des techniques d'apprentissage artificiel comme le *Support vector machine* dans (Harrie, Stigmar et Djordjevic 2015), ou en utilisant des techniques de décision multicritères (Figueira, Greco, et Ehrogott 2007) comme celle utilisée par Touya et Brando (2013).

Conclusion et perspectives

En conclusion, cet article montre que la question de la complexité des cartes a été abordée par différents points de vue et domaines de recherche, mais des verrous scientifiques existent encore pour expliquer efficacement ces recherches dans des processus de cartographie automatique. Différents cas d'études ont été présentés, et l'on y voit à chaque fois qu'une évaluation basique de la complexité ne permet pas de bien mesurer la quantité d'information ou sa désorganisation. Nous avons déduit plusieurs verrous de l'étude de ces exemples, comme l'utilisabilité de méthodes de traitement d'image, ou la combinaison possible de plusieurs mesures par des techniques de décision multicritères.

Pour aller plus loin et travailler sur ces verrous, une première étape serait de tester un nombre relativement important des mesures existantes sur des cartes issues de nos cas d'étude, ou d'autres liés à la cartographie automatique (placement de texte dans les cartes par exemple). Ensuite, afin de relier les métriques de complexité à l'efficacité ou la perception de la carte, des tests devraient être réalisés auprès d'utilisateurs. Cela permettrait d'identifier les meilleures mesures pour un cas d'étude, ou une tâche donnée. Enfin, nous souhaitons étudier la faisabilité de la combinaison de plusieurs mesures par des techniques de décision multicritères comme celle utilisée dans (Touya et Brando 2013).

Bibliographie

- Bard S.**, 2004, « Quality assessment of cartographic generalisation », *Transactions in GIS*, n°8 (1), p.63-81.
- Bjørke J.T.**, 1996, « Framework for entropy-based map evaluation », *Cartography and Geographic Information Science*, n°23 (2), p.78-95.
- Bjørke J.T.**, 2012, « Exploration of information theoretic arguments for the limited amount of information in a map », *Cartography and Geographic Information Science*, n°39 (2), p.88-97.
- Ciolkosz-Styk A., Styk A.**, 2013, « Advanced image processing for maps graphical complexity estimation », dans *Proceedings of 26th International Cartographic Conference (ICC 2013)*, Dresden, Germany.
- Chesneau E.**, 2007, « Improvement of Colour Contrasts in Maps: Application to Risk Maps », dans *Proceedings of 10th AGILE International Conference on Geographic Information Science*, Aalborg, Denmark.
- Christophe S.**, 2012, « Cartographic Styles between traditional and original (towards a cartographic style model) », dans *Proceedings of AutoCarto Conference 2012*, 16-18 Septembre, Columbus, Ohio, USA.
- Christophe S., Hoarau C.**, 2013, « Expressive map design based on Pop Art: Revisit of Semiology of Graphics? », *Cartographic Perspectives*, n°73.
- Da Silva M.P., Courboulay V., Estrailier P.**, 2011, « Image complexity measure based on visual attention », dans *Proceedings of 18th IEEE International Conference on Image Processing (ICIP)*, p.3281-3284.
- Fairbairn D.**, 2006, « Measuring map complexity », *The Cartographic Journal*, n°43 (3), p.224-238.
- Figueira J., Greco S., Ehrogott M., dir.**, 2005, *Multiple Criteria Decision Analysis: State of the Art Surveys*, vol. 78 of International Series in Operations Research & Management Science, Springer.
- Harrie L.E., Mustière S., Stigmar H.**, 2011, « Cartographic quality issues for view services in geoportals », *Cartographica*, n°46 (2), p.92-100.
- Harrie L., Stigmar H., Djordjevic M.**, 2015, « Analytical estimation of map readability », *ISPRS International Journal of Geo-Information*, n°4 (2), p.418-446.
- Hoarau C., Christophe S., Mustière S.**, 2013, « Mixing, blending, merging or scrambling topographic maps and orthoimagery in geovisualizations ? », dans *Proceedings of 26th International Cartographic Conference*, Dresden, Germany.
- Itti L., Koch C., Niebur E.**, 1998, « A model of saliency-based visual attention for rapid scene analysis », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, n°20 (11), p.1254-1259.
- Jégou L., Deblonde J.P.**, 2012, « Vers une visualisation de la complexité de l'image cartographique », *Cybergeo: European Journal of Geography*.
- Knöpfli R.**, 1983, « Communication theory and generalization », dans *Graphic Communication and Design in Contemporary Cartography*, Taylor D.R.F., New-York, John Wiley & Sons, p.177-218.
- Li Z., Huang P.**, 2002, « Quantitative measures for spatial information of maps », *International Journal of Geographical Information Science*, n°16 (7), p.699-709.
- MacEachren A.M.**, 1982, « Map complexity: comparison and measurement », *American Cartographer*, n°9, p.31-46.
- Miller G.A.**, 1956, « The magical number seven, plus or minus two: some limits on our capacity for processing information », *Psychological review*, n°63 (2) p.81-97.
- Moacdieh N., Sarter N.**, 2014, « Display clutter: A review of definitions and measurement techniques », *Human Factors: The Journal of the Human Factors and Ergonomics Society*, n°3.
- Olsson P.O., Pippig K., Harrie L.E., Stigmar H.**, 2011, « Identifying map regions that are difficult to read », *Mapping and Image Science*, n°3, p.22-29.
- Ory J., Christophe S., Fabrikant S.**, 2013, « Identification of styles in topographic maps », dans *Poster Proceedings, 26th International Cartographic Conference (ICC 2013)*. Dresden, Germany.
- Raposo P., Brewer C.A.**, 2013, « Guidelines for consistently readable topographic vectors and labels with toggling backgrounds », dans *Proceedings of 26th International Cartographic Conference (ICC 2013)*, Dresden, Germany.
- Raposo P., Brewer C.A.**, 2014, « Landscape preference and map readability in design evaluation of topographic maps with an orthoimage background », *The Cartographic Journal*, n°51 (1), p.25-37.

- Remy, C.**, 2014, Sémiologie graphique pour la représentation des dynamiques urbaines, rapport de master, Université Paris-Est, France.
- Rosenholtz R.**, Li Y., Mansfield J., Jin Z., 2005, « Feature congestion: A measure of display clutter », *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '05*, New York, NY, USA, p.761-770.
- Rosenholtz R., Li Y., Nakano L.**, 2007, « Measuring visual clutter », *Journal of Vision*, n°7 (2).
- Rylov M. A., Reimer A., 2015, « Improving label placement quality by considering basemap detail with a raster-based approach », *GeoInformatica*, n°19 (3), p.463-486.
- Shannon C.E., Weaver W., 1949, *The Mathematical Theory of Communication*, University of Illinois Press.
- Stigmar, H., Harrie L.E.**, 2011, « Evaluation of analytical measures of map legibility », *The Cartographic Journal*, n°48 (1), p.41-53.
- Stoter J., Zhang X., Stigmar H., Harrie L.**, 2014, « Evaluation in generalisation », dans *Abstracting Geographic Information in a Data Rich World*, Burghardt D., Duchêne C., Mackaness W. dir., Springer International Publishing, p.259-297.
- Touya G.**, 2012, « Social welfare to assess the global legibility of a generalized map », dans *Geographic Information Science 7th International Conference, GIScience 2012*, Xiao N., Kwan M.-P., Goodchild M. F., Shekhar S., dir., Berlin, Heidelberg, Springer, p.198-211.
- Touya G., Brando C.**, 2013, « Détection d'incohérences de niveau de détail dans des données collaboratives », *Cartes et géomatique*, n°217, p.59-72.
- Touya G., Reimer A.**, 2015, « Inferring the scale of OpenStreetMap features », dans *OpenStreetMap in GIScience*, Jokar Arsanjani J., Zipf A., Mooney P., Helbich M., dir., Springer International Publishing, p.81-99.