

---

# Proposition d'une famille de courbes remplissant l'espace de niveau de localité comparable à la courbe de Hilbert

Giap Nguyen<sup>1</sup>, Patrick Franco<sup>1</sup>, Rémy Mullot<sup>1</sup>,  
Jean-Marc Ogier<sup>1</sup>

Laboratoire L3i, Université de La Rochelle  
Avenue M. Crépeau  
F-17042 La Rochelle cedex 01

{giap.nguyen,patrick.franco,remy.mullot,jean-marc.ogier}@univ-lr.fr

---

*RÉSUMÉ. Les courbes remplissant l'espace sont largement utilisées dans plusieurs domaines de l'informatique où la conservation de la localité est souvent considérée comme le critère pour choisir le type de courbe opérant dans une application. C'est pourquoi, la courbe de Hilbert, en qualité de courbe préservant au mieux la localité, reste majoritairement employée. Cependant, existe-t-il d'autres courbes vérifiant un niveau de localité comparable à la courbe de Hilbert ? Dans cet article, nous proposons une méthode flexible permettant de construire une famille de courbes - remplissant un espace multidimensionnel - de niveau de localité comparable à la courbe de Hilbert et parfois meilleur. Des tests expérimentaux comparatifs tendent à confirmer ces résultats. Notons que des éléments de preuve sur la validité de notre proposition sont avancés, sa mise en œuvre est illustrée à travers de nombreux exemples.*

*ABSTRACT. Space-filling curves are widely used in many fields of computer science in which the locality preserving is often seen as the criterion to select the curve type working in an application. This is why the Hilbert curve, as the best locality preserving curve, is mainly used. However, are there other curves who preserve the locality as well as to the Hilbert curve ? In this paper, we propose a flexible method for constructing a family of multidimensional space-filling curves whose locality preserving level is comparable or better than that of the Hilbert curve. Comparative experimental tests tend to confirm these results. Note that the evidence on the validity of our proposal is advanced, its implementation is illustrated through numerous examples.*

*MOTS-CLÉS : courbe remplissant l'espace, fractal, courbe de Hilbert, localité, motif primitif, généralisation.*

*KEYWORDS: space-filling curve, fractal, Hilbert curve, locality, primitive pattern, generalization.*

---

DOI:10.3166/TS.29.553-574 © 2012 Lavoisier

### Extended abstract

The space-filling curves have shown their utilities in many fields as databases, signal and image processing. The most famous curves include that was proposed by Hilbert, Peano and Lebesgue.

These curves perform a mapping from a multidimensional space to a 1-D space. In other words, this mapping lets us know the order (1-D value) of a multidimensional point along the considered curve. This can be considered as a linearization of a multidimensional space by ordering its points.

A part of the topological relationships between multidimensional points in the original space is preserved in the 1-D space, this is called locality-preserving mapping. Visually, each point is ordered near its neighbors. This property is useful because it conditions that any interpretation of original data by using the analysis on the 1-D space is still acceptable.

Among the space-filling curves, the Hilbert curve is known to be one that maximizes the locality according to many criteria. This is the reason why it is used in many applications where the decision process is often taken on the 1-D space.

However, in the multidimensional case, the generation of the Hilbert curve remains an open problem. There are some algorithms which generate the Hilbert curve in the multidimensional case but there is no explicit definition. If we refer to Hilbert's original paper, we have only the construction rule that focuses on two dimensions while some applications require the higher-dimensional curves.

In this paper, at first, we propose a definition of the multidimensional Hilbert curve by putting forward 3 conditions that the only 2-D curves satisfying them are the original 2-D Hilbert curves. So this is a generalization of the definition formulated by Hilbert in his original paper explicit enough for a computer implementation.

From 3 dimensions, this new formulation gives birth to a family of curves and not a single curve as in the case of 2-D Hilbert curve. These curves preserve differently the locality and their locality preserving is comparable to that of Hilbert curve. This track opens the possibility to choose a curve according to the expected degree of locality preserving, which is the interest of many applications.

We also offer a flexible method for constructing the defined space-filling curves. By determining the geometric transformations of the sub-curves, they are articulated to satisfy the conditions of the curve. The geometric transformation of each sub-curve is controlled by its beginning, ending points and the pattern, which connects these points. Various choices of these elements fill conditions of proposed curve definition, that creates different curves. Particularly, there are many patterns other than the familiar RBG pattern can be used in the multidimensional Hilbert curve.

## 1. Introduction

Les courbes fractales déterministes et en particulier les courbes ayant la propriété de remplir un espace (*space-filling curves*) ont montré leur utilité dans des domaines aussi variés que les bases de donnée (Chen, Chang, 2011 ; Angiulli, Pizzuti, 2005 ; Faloutsos, Roseman, 1989 ; Lawder, 2000 ; Asano *et al.*, 1997), le traitement du signal et des images (Ebrahim *et al.*, 2009 ; Liang *et al.*, 2008 ; Yiu *et al.*, 2008) et, les problèmes d'optimisation (Oliveira, Petraglia, 2011 ; Lera, Sergeyev, 2010 ; Butz, 1969 ; Muelder, Ma, 2008). Parmi les plus célèbres, on peut citer les courbes de Hilbert (Hilbert, 1891), Peano (Peano, 1890) et Lebesgue (Sagan, 1994) (cf. figure 1).

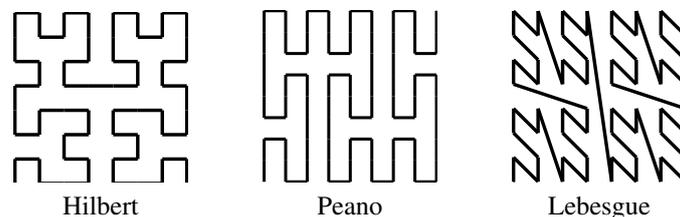


Figure 1. Exemples de courbes remplissant un espace de dimension  $d = 2$

Ces courbes permettent le plongement d'un espace multidimensionnel vers un espace monodimensionnel (*mapping*). Plus précisément, le plongement d'un point  $d$ -dimensionnel sur une courbe fractale consiste à connaître son rang (index, valeur monodimensionnelle) le long de la courbe fractale considérée. Cette opération revient à linéariser un espace multidimensionnel en ordonnant de manière déterministe les hypercubes traversés par ces courbes.

Une partie des relations topologiques existantes (voisinages) entre les points multidimensionnels évoluant dans l'espace original est ainsi conservée après plongement dans l'espace des rangs, on parle de *localité* (*preserving mapping*) (Gotsman, Lindenbaum, 1996 ; Moon *et al.*, 2001). Cette propriété est précieuse car elle conditionne que l'interprétation - concernant le paysage des données évoluant dans l'espace original - menée *a posteriori* sur une analyse de l'espace des rangs reste acceptable.

La courbe de Hilbert est connue pour être celle qui maximise la localité et cela quel que soit le critère utilisé pour son estimation (Faloutsos, Roseman, 1989 ; Gotsman, Lindenbaum, 1996 ; Moon *et al.*, 2001). C'est l'une des raisons pour laquelle elle reste majoritairement utilisée dans de nombreuses applications (Lawder, King, 2000 ; Liang *et al.*, 2008 ; Muelder, Ma, 2008).

Cependant, dans le cas multidimensionnel la génération d'une courbe de Hilbert reste un problème ouvert. Si l'on se réfère à l'article originel de (Hilbert, 1891), on ne dispose que de règles de construction axées sur la dimension  $2^1$  alors que certaines applications nécessiteraient la construction de courbes de dimension supérieure (Lera, Sergeyev, 2010).

1. Aucun élément n'est donné sur les dimensions supérieures.

Dans cet article, nous proposons une extension de la définition d'une courbe de Hilbert dépassant le cas de la dimension 2. Il s'agit donc d'une généralisation de la définition formulée par Hilbert - dans son article originel (Hilbert, 1891) - suffisamment explicite pour envisager une implémentation informatique.

Cette nouvelle formulation<sup>2</sup> donne naissance à une famille de courbes solutions et non plus une courbe unique comme dans le cas de Hilbert (pour  $d = 2$ ). Ces courbes solutions ont un impact sur la propriété de localité. Cette piste ouvre la possibilité de choisir une courbe suivant le degré de vérification de la localité attendu, ce qui présente un intérêt certain en termes applicatifs.

L'article est organisé de la façon suivante : la définition de Hilbert et notre positionnement par rapport aux travaux existants sont présentés dans la section 2. Notre proposition étayée par des éléments de preuve est abordée dans les sections 3 et 4. La section 5 est consacrée à l'estimation de la localité pour les différentes courbes résultant des nouvelles règles de construction.

## 2. Courbe de Hilbert et quelques extensions

En réponse à la proposition de Peano<sup>3</sup> (Peano, 1890), (Hilbert, 1891) propose une manière alternative de faire correspondre les points d'un carré avec ceux d'un segment préservant mieux la localité.

Initialement, le carré est découpé en 4 parties par 2 droites orthogonales. Chacune des parties est ensuite ordonnée de façon à ce que deux parties consécutives aient un côté commun (cf. figure 2 (a)). Les mêmes opérations (découpage et ordonnancement) sont appliquées sur chacune des parties (résultant du premier découpage), il en résulte des sous-parties, dont l'ordre est choisi de sorte que deux sous-parties consécutives ont un côté commun (cf. figure 2 (b)). Le processus est réitéré sur chacune des sous-parties (cf. figure 2 (c)).

Après application de ce processus de construction (cf. figure 2) on constate que la courbe de Hilbert à l'ordre 2 (b) est constituée de copies transformées de la courbe de Hilbert d'ordre 1 (b). Dans la suite de l'article, la courbe remplissant l'espace à l'ordre unité  $n = 1$  est appelée motif primitif.

En généralisant à travers les ordres, à dimension d'évolution  $d$  fixée, une courbe de Hilbert d'ordre  $n$  est constituée de  $2^d$  sous-courbes. Chaque sous-courbe est une version transformée de la courbe de Hilbert d'ordre antérieur  $n - 1$  (cf. figure 2 (b)-(c)). De plus, on remarque que pour  $n$  donné, les transformations géométriques (rotation, réflexion) diffèrent suivant la localisation même des sous-courbes (exemple pour  $n = 2$ , cf. figure 2 (b) coins inférieurs gauche et droit). L'enchaînement de ces transformations est un facteur-clé pour construire une courbe qui préserve la localité. Or, lorsque les dimensions augmentent, la détermination de ces transformations et de leur

2. Des règles de construction.

3. Notons que la courbe proposée par Peano est considérée comme la première courbe fractale déterministe remplissant un espace.

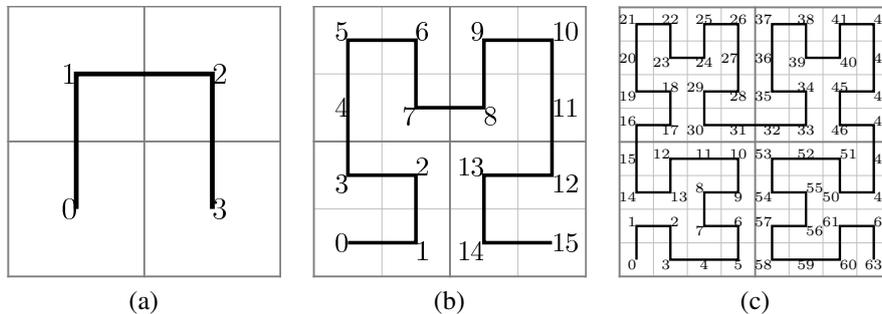


Figure 2. Exemple de courbes de Hilbert en dimension  $d = 2$  et d'ordre variable respectivement de gauche à droite  $n = 1, 2, 3$

enchaînement n'est plus triviale y compris en faisant référence aux règles de Hilbert (établies en dimension 2) ou *via* le modèle abstrait proposé dans (Sagan, 1994). La dimensionnalité est source de difficulté de construction, cela se confirme dans la littérature où finalement peu d'algorithmes (à notre connaissance) hormis les célèbres schémas de (Bially, 1969) et (Butz, 1971) opèrent au-delà de la dimension 3.

(Bially, 1969) fait référence à des machines états-transitions (cf. figure 3) pour le calcul du plongement d'un point (à partir de ses coordonnées binaires) sur une courbe de dimension arbitraire ( $d \geq 2$ )<sup>4</sup>. Les matrices de transformations - qui participent à la construction même de ces machines - sont étroitement liées à des discussions portant essentiellement sur un seul type de courbe multidimensionnelle, la courbe RBG (*Reflected Binary Gray code*) (Bitner *et al.*, 1976). C'est une approche claire, interprétable, servant de référence à divers articles (Faloutsos, Roseman, 1989; Lawder, King, 2001). Néanmoins, ces machines doivent être régénérées lors de tout changement de dimension d'entrée  $d$  (modification du nombre d'états, de transitions etc.). La génération de ces machines, dans sa version automatique (Faloutsos, Roseman, 1989), se heurte à la difficulté d'arbitrer parmi un ensemble de solutions qui apparaissent à diverses étapes de la conception, ce qui en pratique rend son implémentation (dans sa version automatique) délicate au-delà de la dimension 5.

Le schéma de Butz dans (Butz, 1971) s'appuie sur un modèle de transition valide quel que soit la dimension, codé par une série d'opérations binaires. Il permet de retrouver un point de la courbe à partir de son rang et non l'inverse. Or, de nombreuses applications nécessitent également le calcul du plongement (réversibilité).

Notons que ces propositions reposent essentiellement sur la courbe multidimensionnelle RBG, *i.e.* finalement sur un nombre limité de manière de remplir l'espace. Est-il possible de construire d'autres courbes pour remplir un espace multidimensionnel qui vérifieraient un niveau de localité comparable à la courbe de Hilbert ?

4. communément appelée courbe de Hilbert

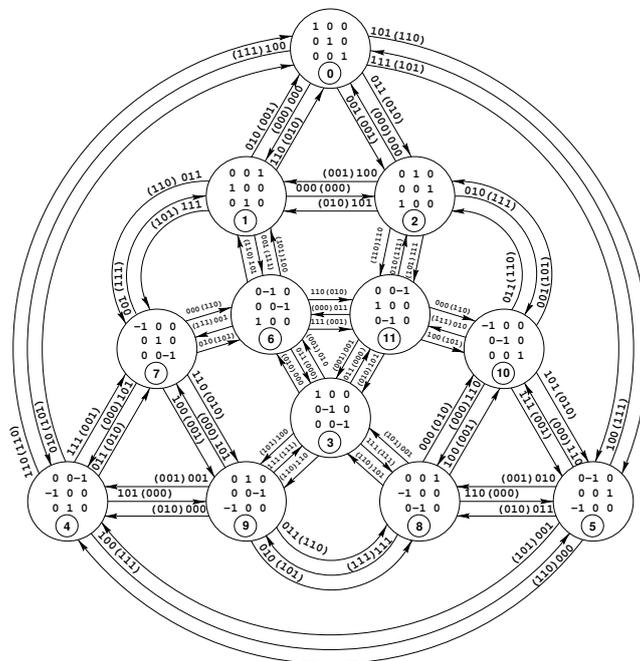


Figure 3. Exemple de machine états-transitions (Lawder, King, 2001) pour le calcul du plongement d'un point sur la courbe de Hilbert en  $d = 3$  quel que soit  $n$ . Cette machine fait référence à un modèle (interne) de construction de courbe multidimensionnelle s'appuyant sur un motif primitif de type RBG. On reconnaît les matrices de transformation (cercles) traduisant l'enchaînement des sous-courbes à travers les divers ordres  $n$

### 3. Proposition : courbe de Hilbert généralisée

#### 3.1. Motivations

Nous proposons une méthode (cf. section 3.2) permettant de construire une famille de courbes remplissant un espace multidimensionnel de niveau de localité comparable à la courbe de Hilbert (cf. section 5). Nous avons souligné en section 2 que l'identification des transformations géométriques (à l'ordre  $n$  donné) joue un rôle dans l'articulation des  $2^d$  sous-courbes, contribuant ainsi à construire globalement une courbe conservant la localité. Nous pensons que le contrôle des points de départ et d'arrivée (extrémités) au niveau de chaque sous-courbe est un facteur déterminant dans le cadre d'une extension multidimensionnelle. En imposant *via* une contrainte (cf. définition 3.2/condition C2) que la distance entre les points d'arrivée et de départ de deux sous-courbes consécutives soit unitaire (adjacence), on garantit la continuité inter-sous-courbes ce qui permet :

- de contourner la difficulté<sup>5</sup> d'identifier explicitement un ensemble de matrices de transformations (cf. section 2),
- d'éviter toute rupture de topologie (saut dans l'espace original) nuisible pour la localité.

De plus, en formulant des règles guidant la construction des sous-courbes (cf. définition 3.2/conditions *C1* et *C3*), nous montrons qu'à extrémités fixées il existe plusieurs courbes remplissant un espace. Ceci est également vrai à l'ordre  $n = 1$  ce qui nous permet de sortir du seul recours à la courbe RBG comme motif primitif (cf. section 4.1). Ces hypothèses sont confirmées par les expérimentations menées dans la section 5.

### 3.2. Définition et exemples d'application

Nous postulons qu'il existe plusieurs courbes au-delà de la dimension 2, autres que celles générées par l'algorithme de Butz (Butz, 1971) ou Lawder (Lawder, King, 2001). Notre hypothèse est qu'en proposant une extension au cas multidimensionnel (cf. définition 3.2) de la définition de Hilbert établie en dimension 2 (cf. section 2), nous accédons à l'ensemble de ces courbes. Nous proposons une méthode qui consiste à discrétiser un hypercube unité<sup>6</sup> de dimension  $d$ , en  $2^{nd}$  hypercubes de largeur  $1/2^n$  dont l'ordre de parcours satisfait un niveau de localité (au sens de (Faloutsos, Roseman, 1989 ; Gotsman, Lindenbaum, 1996 ; Moon *et al.*, 2001)) comparable à celui de la courbe de Hilbert. Notons que c'est l'ordre de parcours des hypercubes qui forme une courbe.

**DÉFINITION (Courbe de Hilbert généralisée).** — *Une courbe remplissant un hypercube de dimension  $d$  est construite de façon récursive. À chaque itération  $n$ , chaque hypercube est découpé en  $2^d$  sous-hypercubes égaux par  $d$  hyperplans deux à deux orthogonaux, et ensuite, tous les sous-hypercubes sont ordonnés de la manière suivante :*

- C1. Les sous-hypercubes d'un même hypercube sont ordonnés consécutivement,*
- C2. Le dernier sous-hypercube d'un hypercube précède le premier sous-hypercube du suivant,*
- C3. Les deux sous-hypercubes consécutifs partagent une hyperface commune.*

Les sous-hypercubes sont donc ordonnés suivant une règle exprimée par les conditions *C1*, *C2* et *C3*. Notons que c'est la simple liaison par des segments des sous-hypercubes ainsi ordonnés qui forme une courbe. Pour mieux cerner le rôle de chaque condition dans le processus de construction des courbes, la Définition est appliquée en dimension  $d = 2$ , jusqu'à l'ordre  $n = 2$ . Diverses configurations sont testées et analysées allant du respect individuel de chaque condition jusqu'au respect total de l'ensemble.

5. croissante avec les dimensions.

6. les arêtes sont égales à 1.

La condition  $C1$  régit l'ordonnement des sous-hypercubes situés à l'intérieur d'un hypercube donné (cf. figure 4). Aucun sous-hypercube appartenant à un hypercube voisin n'est inclus dans cet ordonnancement. C'est donc une contrainte locale qui crée la localité à l'intérieur de chacun des hypercubes. Elle s'applique dès le premier découpage (pour  $n = 1, 2 \dots$ ).

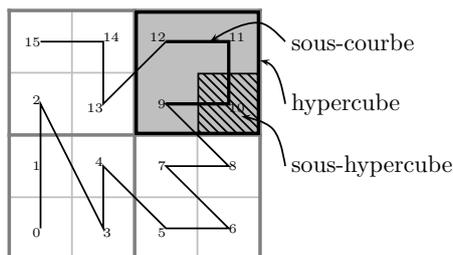


Figure 4. Influence de la condition  $C1$  dans le processus de construction des courbes. Cas de sous-courbes synthétisant le respect de la condition  $C1$  (parties droite haut et bas) ou non (parties gauche haut et bas) dans l'ordonnement des sous-hypercubes

La localité ainsi créée à l'intérieur de chaque hypercube, doit être conservée entre les hypercubes voisins. C'est le rôle de la condition  $C2$  qui assure l'enchaînement des ordonnancements locaux lors du voyage entre les hypercubes dans le sens imposé par le motif primitif (cf. figure 5). Elle s'applique dès le deuxième découpage (pour  $n = 2, 3 \dots$ ) et contribue à préserver la localité lors de tout changement d'ordre.

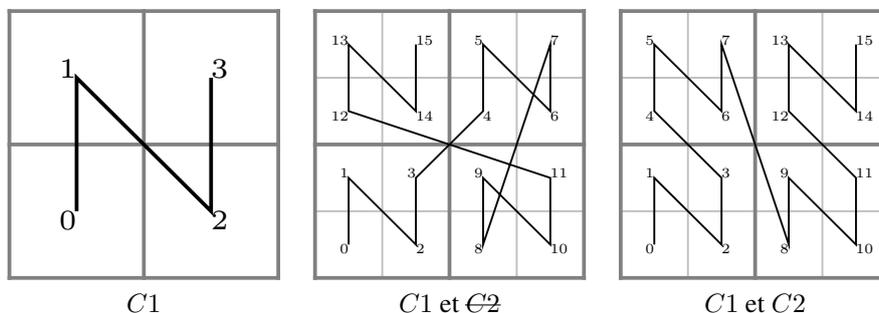


Figure 5. Influence de la condition  $C2$  dans le processus de construction des courbes. Cas de sous-courbes synthétisant le respect de la condition  $C2$  (à droite) ou non (milieu) dans l'ordonnement des sous-hypercubes (sachant le respect de la condition  $C1$ ). On reconnaît la célèbre courbe de Lebesgue à droite

On remarque que le respect conjoint des conditions  $C1$  et  $C2$  donne naissance à d'autres types de courbes remplissant l'espace telles que, par exemple, celle de Lebesgue dans le cas de la figure 5.

L'ajout de la condition  $C3$  au respect conjoint des conditions  $C1$  et  $C2$  permet de construire une courbe dont le niveau de localité est comparable à la courbe de Hilbert (cf. section 5).

La condition  $C3$  opère localement (dès l'ordre 1) dans l'ordonnement des sous-hypercubes, venant compléter la condition  $C1$  (cf. illustrations à  $n = 1$ , figure 6) ; mais également (dès l'ordre 2) dans l'enchaînement des ordonnancements locaux complétant ainsi la condition  $C2$  (cf. illustrations à  $n = 2$ , figure 7). Ainsi tous les sous-hypercubes appartenant à un hypercube donné doivent certes être connectés consécutivement avant d'être enchaînés aux sous-hypercubes voisins mais de plus, ces connections tout comme ces enchaînements doivent vérifier l'adjacence (cf. figure 7, milieu et droit). L'adjacence de 2 sous-hypercubes consécutifs est définie en annexe (cf. corollaire 10).

En imposant une contrainte d'adjacence aux sous-hypercubes consécutifs, la condition  $C3$  crée une partition, en termes de localité, entre les courbes obéissant à la définition et d'autres types de courbes comme celle de Lebesgue par exemple.

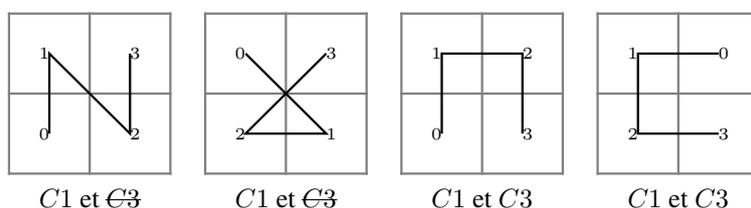


Figure 6. Influence de la condition  $C3$  à l'ordre 1 dans le processus de construction des courbes. Cas de sous-courbes synthétisant le respect de la condition  $C3$  (2 courbes à droite) ou non (2 courbes à gauche) sachant le respect de la condition  $C1$

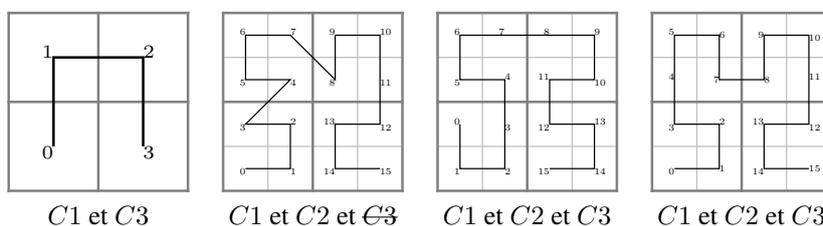


Figure 7. Vers le respect de l'ensemble des conditions : influence de la condition  $C3$  à l'ordre 2 dans la construction des courbes. Sachant le respect de  $C1$  et  $C2$ , cas de courbes synthétisant le respect ou non de  $C3$ . Les 2 courbes à droites satisfont toutes les conditions, on reconnaît la courbe de Hilbert à l'extrême droite

Finalement, la Définition proposée est suffisamment large pour que plusieurs courbes (une famille) puissent en satisfaire toutes les conditions (cf. les 2 courbes à droite de la figure 7). La section 4 est consacrée à l'apport d'éléments de preuve concernant sa validité. Notamment, nous répondrons à trois questions-clés :

- existe-t-il, en dimension  $d$  donnée, des courbes vérifiant toutes les conditions de la Définition ?
- si oui, combien de courbes ?
- ont-elles le même niveau de localité ?

Ces éléments présentent l'intérêt de mieux identifier les degrés de liberté apparaissant dans le processus de construction ce qui s'avère utile dans la perspective d'une implémentation informatique.

#### 4. Processus de construction et preuves

Nous rappelons que ce processus consiste à discrétiser un espace de dimension  $d$ , à l'ordre  $n$ , en  $2^{nd}$  hypercubes dont l'ordre de parcours satisfait un niveau de localité comparable à celui issu de la courbe de Hilbert. Lors du passage de l'ordre  $n$  à  $n + 1$ , chacun des  $2^{nd}$  hypercubes est discrétisé localement en  $2^d$  hypercubes. Deux types de contraintes doivent être vérifiées :

- une contrainte locale portant sur l'ordre de parcours des  $2^d$  hypercubes (conditions  $C1$  et  $C3$ );
- une contrainte globale s'exerçant sur l'articulation (l'enchaînement) des divers parcours établis localement sur chacun des  $2^{nd}$  hypercubes (conditions  $C2$  et  $C3$ ).

La contrainte globale est plus restrictive que la contrainte locale. Autrement dit, nous devons d'abord identifier les points de départ et d'arrivée (dans chaque hypercube) garantissant les conditions  $C2$  et  $C3$  avant de remplir l'espace ainsi délimité par des sous-courbes vérifiant la contrainte locale. Ceci est synthétisé par le processus de construction suivant :

- **Entrée** : un espace de dimension  $d$
- **Répéter à l'ordre  $n$  désiré** :
  - a) découper les hypercubes,
  - b) choisir un couple d'extrémités dans chaque hypercubes vérifiant les conditions  $C2$  &  $C3$ ,
  - c) choisir un parcours entre les couples d'extrémités vérifiant les conditions  $C1$  &  $C3$ .

Par souci de clarté, nous abordons dans un premier temps (cf. section 4.1) la construction de la courbe vérifiant les contraintes locales (le motif primitif). Nous traitons ensuite (cf. section 4.2) la construction de la courbe satisfaisant la contrainte globale.

##### 4.1. Construction d'une courbe à l'ordre 1 : le motif primitif

La courbe obtenue après application de la Définition (proposée dans la section 3.2) à l'ordre  $n = 1$  est appelée motif primitif. L'hypercube unité est donc découpé pour la première fois en  $2^d$  sous-hypercubes dont l'ordre de parcours est guidé par les conditions  $C1$  et  $C3$ . Notons qu'à l'ordre 1, la condition  $C2$  dédiée à la conservation de la localité lors de tout changement d'ordre ( $n = 1 \rightarrow 2$ ) n'agit pas (premier découpage).

En d'autres termes, à l'ordre  $n = 1$ , seules les contraintes internes (locales) opèrent. Nous sommes donc *a priori* libre de choisir des sous-hypercubes de départ et d'arrivée autres que ceux imposés par la courbe RBG<sup>7</sup>. Ces éléments sont développés par la suite.

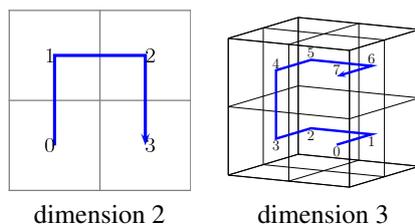


Figure 8. Remplir un espace de dimension 2 et 3 : recours classique à la courbe RBG

Notons que le recours classique à la courbe RBG comme courbe de Hilbert  $d$ -dimensionnelle d'ordre 1, satisfait les conditions  $C1$  et  $C3$  de notre proposition quel que soit  $d$ , elle est donc un motif primitif (cf. figure 8). Le code RBG permet de générer des mots binaires de longueur finie tels que deux mots consécutifs ne diffèrent que d'un seul bit, par conséquent :

THÉORÈME 1 (Existence du motif primitif). — *Quel que soit la dimension  $d$ , il existe au moins un motif primitif.*

En Annexe B, nous fournissons l'algorithme 1 permettant d'identifier tous les motifs primitifs à dimension  $d$  donnée. Il est fondé sur la démonstration du théorème 2. La figure 9 illustre la diversité des motifs primitifs solutions en dimension 2.

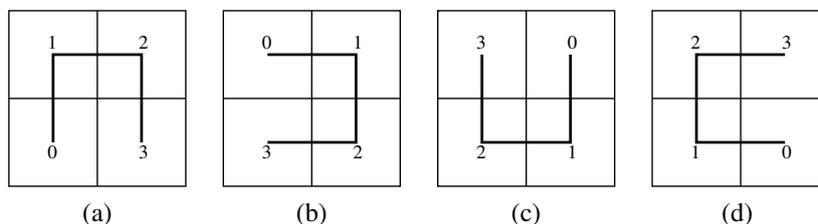


Figure 9. Plusieurs façons de remplir un espace de dimension 2 : exemples de motifs primitifs résultant de l'application de la définition 3.2 (on reconnaît la courbe RBG (a) utilisée dans la littérature)

La condition  $C3$ , qui porte sur l'adjacence des sous-hypercubes consécutifs, est une contrainte souple au gré de l'augmentation des dimensions. En effet, elle peut être satisfaite dans de multiples directions. Cette souplesse autorise l'existence de plusieurs motifs primitifs solutions, ouvrant le choix à d'autres courbes que RBG.

7. Chacun des mots générés suivant le code binaire réfléchi définit les coordonnées d'un hypercube évoluant dans un espace  $d$ -dimensionnel. Une suite de mots consécutifs délivre un ordre de lecture (parcours) des hypercubes qui est synthétisé par une courbe dite RBG.

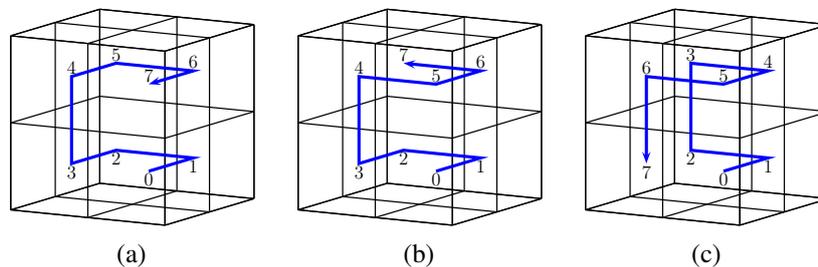


Figure 10. Plusieurs façons de remplir un espace de dimension 3 : exemples de motifs primitifs résultant de l'application de la définition 3.2 (seule la courbe RBG (a) est utilisée dans la littérature)

THÉORÈME 2 (Distance impaire). — Entre deux extrémités séparées par une distance impaire, il existe toujours au moins un motif primitif permettant de les connecter.

PREUVE 3. — Nous démontrons le théorème 2 par induction.

Quel que soit  $d$ , il existe toujours au moins un motif primitif entre deux extrémités dont la distance est 1, par exemple, la courbe RBG (cf. section 4.1).

Supposons que nous pouvons toujours construire un motif primitif entre deux extrémités dont la distance est impaire et vaut  $2m-1$  ( $\forall m \in \mathbb{N}, 2m-1 \geq 1, 2m+1 \leq d$ ).

Soit deux points  $a = a_1 a_2 \dots a_d$  et  $b = b_1 b_2 \dots b_d$  respectivement le point de départ et d'arrivée satisfaisant  $\|a, b\| = 2m+1$ . Car  $2m-1 \geq 1 \Rightarrow \|a, b\| = 2m+1 \geq 3$ , nous pouvons toujours trouver deux positions différentes  $i, j$  telles que  $a_i \neq b_i, a_j \neq b_j, i < j$ .

En examinant deux points :

$$\begin{aligned} a' &= a_1 a_2 \dots a_{i-1} \mathbf{a}_i a_{i+1} \dots \mathbf{a}_{j-1} \mathbf{a}_{j+1} \dots a_d \\ b' &= b_1 b_2 \dots b_{i-1} \mathbf{a}_i b_{i+1} \dots \mathbf{b}_{j-1} \mathbf{b}_{j+1} \dots b_d \end{aligned}$$

dans l'espace de dimension  $(d-1)$ , nous voyons que  $\|a', b'\| = 2m-1$ , donc, nous pouvons construire un motif primitif  $M = (a', \dots, b')$  entre  $a'$  et  $b'$ . En ajoutant  $a_j$  à la position  $j$  de chaque point appartenant à  $M$ ,  $M$  devient un autre parcours  $M'$  (en dimension  $d$ ) reliant  $a$  et  $b'' = b_1 b_2 \dots b_{i-1} \mathbf{a}_i b_{i+1} \dots b_{j-1} \mathbf{a}_j b_{j+1} \dots b_d$ ; notons que la distance entre deux points consécutifs de  $M'$  est toujours égale à 1 (cf. corollaire 10 - Annexe).

Examinons ensuite deux autres points en dimension  $d-1$  :

$$\begin{aligned} b^{(3)} &= b_1 b_2 \dots b_{i-1} \mathbf{a}_i b_{i+1} \dots \mathbf{b}_{j-1} \mathbf{b}_{j+1} \dots b_d \\ b^{(4)} &= b_1 b_2 \dots b_{i-1} \mathbf{b}_i b_{i+1} \dots \mathbf{b}_{j-1} \mathbf{b}_{j+1} \dots b_d \end{aligned}$$

nous avons  $d(b^{(3)}, b^{(4)}) = 1$ . Par conséquent, nous pouvons construire un motif primitif  $N = (b^{(3)}, \dots, b^{(4)})$  entre eux. En ajoutant  $b_j$  à la position  $j$  de chaque point appartenant à  $N$ ,  $N$  devient un parcours  $N'$  en dimension  $d$  reliant  $b^{(5)} =$

$b_1 b_2 \dots b_{i-1} a_i b_{i+1} \dots b_{j-1} b_j b_{j+1} \dots b_d$  à  $b$ , la distance entre deux points consécutifs de  $N'$  est toujours égale à 1.

Tous les points appartenant à un parcours donné ( $M'$  ou bien  $N'$ ) sont deux à deux différents. De plus, parce que  $a_j \neq b_j$ , tous les points dans  $M'$  sont deux à deux différents avec ceux de  $N'$ . Par conséquent, tous les points dans les deux parcours sont deux à deux différents. On est donc sûr de ne pas passer deux fois au même endroit. Car  $d(b'', b^{(5)}) = 1$ , la distance entre deux certains points consécutifs du parcours joint  $M'N' = (a, \dots, b'', b^{(5)}, \dots, b)$  est toujours égale à 1.

$M$  et  $N$  sont les motifs primitifs dans l'espace de dimension  $(d - 1)$ , ainsi, le nombre de points dans chaque motif est  $2^{d-1}$  et la totalité des points dans  $M'N'$  est  $2 * 2^{d-1} = 2^d$ . Ainsi on est sûr que l'espace est totalement rempli. En conclusion,  $M'N'$  est un motif primitif dans l'espace de dimension  $d$ .

Nous venons donc de démontrer qu'il est possible de construire un motif primitif entre deux points séparés par une distance impaire (et  $\leq d$ ). ■

De plus, il est possible d'en construire d'autres (à distance impaire fixée), en faisant simplement varier le choix des paires de position  $(i, j)$  parmi  $2m + 1$  positions possibles (cf. démonstration du théorème 2). Cette idée constitue le fondement de l'algorithme 1 (cf. annexe Annexe B). Notons que les motifs primitifs satisfont par construction la contrainte locale, ils sont donc des candidats aux sous-courbes quel que soit l'ordre  $n$  choisi.

#### 4.2. Le passage aux ordres supérieurs

Les développements suivants (théorème 4, corollaire 6) ont pour but de cerner (localement dans chaque sous-hypercube) les couples d'extrémités susceptibles de garantir la contrainte globale pour lesquels on est sûr qu'il existe une sous-courbe (motif primitif) réalisant leur connexion. Autrement dit, est-il possible de construire une sous-courbe avec deux extrémités fixées ?

THÉORÈME 4 (Distance paire). — *Entre deux extrémités séparées par une distance paire, il n'existe pas de motif primitif permettant de les connecter.*

PREUVE 5. — Si la distance entre deux extrémités est paire, alors cela signifie que le nombre de modifications à appliquer sur les coordonnées du point de départ pour atteindre celles du point d'arrivée est pair. Autrement dit, relier deux points séparés par une distance paire passe nécessairement par la connexion d'un nombre impair de points intermédiaires. Comme le nombre total de points d'un motif primitif de dimension  $d$  est de  $2^d$  donc toujours pair cela implique un nombre de points intermédiaires pair ( $2^d - 2$ ), par conséquent, il ne peut pas exister de chemin connectant deux extrémités séparées par une distance paire et donc il n'existe pas de motif primitif entre eux. ■

En conclusion, d'après les théorèmes 2 et 4, seules les extrémités séparées par une distance impaire (et  $\leq d$ ) peuvent être connectées par un motif primitif.

**COROLLAIRE 6.** — *Avec le point de départ fixé, le nombre de possibilités des points d'arrivée d'une sous-courbe est  $2^{d-2}$*

**PREUVE 7.** — Le point d'arrivée d'une sous-courbe doit être adjacent à la sous-courbe suivante. L'hyperplan orthogonal à l'axe liant ces deux sous-courbes coupe la sous-courbe actuelle en deux, une moitié approche la sous-courbe suivante. Si les coordonnées des points sont codées en binaire, seule une coordonnée des points approchant la sous-courbe suivante est identique,  $d - 1$  autres coordonnées sont différentes. Par conséquent, il existe  $2^{d-1}$  points approchant la sous-courbe suivante et donc candidats comme point d'arrivée.

Parmi les  $d - 1$  autres coordonnées, une fois les valeurs des  $d - 2$  coordonnées choisies, la valeur de la coordonnée restante fixe la distance au point de départ et donc sa parité. Par conséquent, il existe  $2^{d-2}$  points approchant la sous-courbe suivante. Ces points sont séparés du point de départ par une distance impaire.

Parce que la distance d'un point à lui-même (égale à 0) n'est pas impaire, le nombre de possibilité pour définir un point d'arrivée (toujours séparé par une distance impaire du point de départ) est  $2^{d-2}$ . ■

Une fois le point d'arrivée d'une sous-courbe choisi, le point de départ de la sous-courbe suivante est aussi fixé par adjacence. Une courbe d'ordre  $n$  est composée de  $2^{dn}$  sous-hypercubes, le passage de l'ordre  $n$  à l'ordre  $n+1$ , totalise au final  $(2^{d-2})^{2^{dn}}$  possibilités pour définir les points d'arrivée.

## 5. Diversité des courbes et localité

### 5.1. Diversité

Le processus de construction - éclairé par les analyses menées dans les sections 4.1 et 4.2 - permet de cerner les diverses possibilités pour la construction d'une courbe multidimensionnelle :

- choix du motif primitif,
- choix des extrémités (au niveau de chaque sous-courbe),
- choix du parcours entre ces extrémités (définition des sous-courbes).

Un choix de ces trois paramètres est appelé une configuration.

Son exécution délivre une famille de courbes, dont 5 sont données à titre d'exemple dans la figure 11. Chaque courbe correspond à une configuration de construction (notée T, BB, BS, CB, CS) *i.e.* à un choix d'un triplet parmi les degrés de liberté décrits ci-dessus. Ces configurations couvrent plusieurs cas :

- deux motifs primitifs sont utilisés, par exemple, la configuration T obéit au motif (b) illustré dans la figure 10, alors que les autres (BB,BS,CB,CS) obéissent au même motif RBG (a),

- BB et CB partagent les mêmes points de départ et d'arrivée, les mêmes motifs primitifs mais diffèrent par au moins un choix d'extrémités,
- BB et BS partagent les mêmes points de départ et d'arrivée, les mêmes motifs primitifs, les mêmes extrémités mais diffèrent toujours par le choix du parcours entre ces extrémités.

Finalement, on observe (figure 11) des courbes ayant (ou pas) les mêmes points de départ et d'arrivée et décrivant un parcours de l'espace différent.

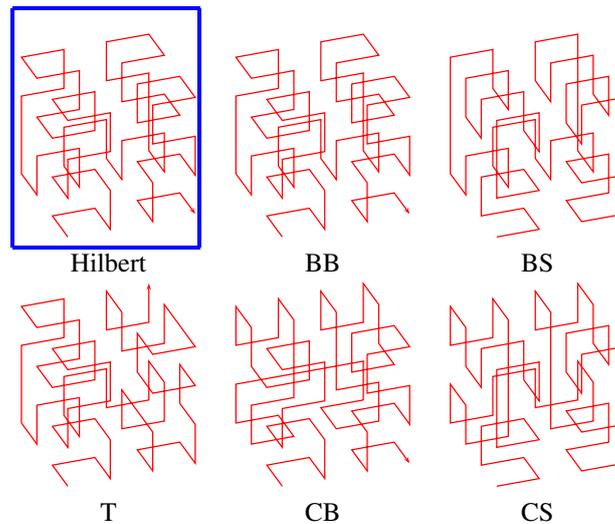


Figure 11. Exemple de 5 courbes ( $d = 3, n = 2$ ) résultant du processus de construction et positionnement par rapport à la courbe de Hilbert (encadré). Chaque courbe correspond à une configuration de construction (triplet d'entrée). On remarque que la courbe construite selon la configuration BB, correspond à la courbe de Hilbert. La courbe de Hilbert est donc une sortie du processus de construction

Dans la section suivante, le niveau de localité de chacune de ces courbes est estimé puis comparé à celui de la courbe de Hilbert. Cette dernière est prise comme référence car il a été montré, que parmi un ensemble de courbes remplissant l'espace (Peano, Lebesgue), elle possède le plus haut niveau de localité (Faloutsos, Roseman, 1989 ; Gotsman, Lindenbaum, 1996 ; Moon *et al.*, 2001).

### 5.2. Conservation de la localité des courbes

Plusieurs mesures ont été proposées dans la littérature (Mitchison, Durbin, 1986 ; Perez *et al.*, 1992 ; Voorhies, 1991) pour estimer la localité d'une courbe remplissant l'espace. L'idée centrale est d'estimer si le plongement d'un point  $d$ -dimensionnel sur une courbe donnée, est une opération conforme au sens de la conservation des distances inter-points avant et après plongement (espace original-espace des rangs).

Dans cet article, la localité de chacune des courbes - décrites dans la section 5.1 - résultant du processus de construction, est estimée par deux mesures, notées  $M_1$  et  $M_2$ .  $M_1$  obéit à la définition de (Perez *et al.*, 1992) et  $M_2$  correspond à une version de  $M_1$  établie dans le pire des cas (il s'agit de mesurer un rapport maximal et non plus moyen).

La mesure  $M_1$  quantifie la conservation moyenne de la localité des courbes :

$$M_1(C) = \frac{1}{N} \sum_{i,j \in R_C, i < j} \frac{|i - j|}{\mathcal{D}(C(i), C(j))}$$

avec  $C$  une courbe remplissant l'espace multidimensionnel,  $R_C$  l'espace unidimensionnel (l'espace des rangs),  $i, j$  les rangs affectés aux points  $C(i), C(j)$  sur la courbe  $C$ .  $N$  est le nombre de couples  $i, j$  et  $\mathcal{D}(C(i), C(j))$  est la distance euclidienne entre  $C(i)$  et  $C(j)$ .

La mesure  $M_2$  quantifie la conservation de la localité des courbes dans le pire des cas :

$$M_2(C) = \max_{i,j \in R_C, i < j} \frac{|i - j|}{\mathcal{D}(C(i), C(j))}$$

Une "bonne" courbe est une courbe qui minimise ces deux mesures.

Plusieurs courbes (30 au total) sont construites en dimensions 3 et 4 balayant les ordres  $n = 2, 3, 4$  selon les 5 configurations (T, BB, BS, CB, CS) décrites dans la section 5.1. La courbe dite de Hilbert résulte de l'application de l'algorithme de Butz (Butz, 1971).

À la lecture de ces résultats, dans le cadre de cette expérience, on observe que :

- quels que soient  $d, n$ , la configuration BB correspond à la courbe de Hilbert (vue par Butz), il est donc logique de retrouver les mêmes estimations,
- quels que soient  $d, n$ , le niveau de localité  $M_1$  correspondant aux 5 configurations est assez proche de celui de la courbe de Hilbert. L'écart maximum relatif reste toujours inférieur à 0,14963 %, ce qui tend à confirmer nos hypothèses sur la conservation de la localité (cf. section 3.1),
- il est possible de construire des courbes ayant un meilleur niveau de localité que la courbe de Hilbert et cela quel que soit le critère ( $M_1$  ou  $M_2$ ) considéré (configuration CB en dimension 3 selon  $M_1$ , T en dimensions 3 et 4 selon  $M_2$  - cf. tableau 1). Bien que numériquement le gain moyen selon  $M_1$  (engendré par CB) semble peu significatif (0,035 %), ce n'est pas le cas selon  $M_2$ , où la configuration T, en dimensions 3 et 4, procure un gain de localité moyen (respectivement maximum) par rapport à la courbe de Hilbert de 24,358 % (respectivement 27,118 %).

Tableau 1. Estimation de la localité (selon 2 mesures) d'une famille de courbes  $d$ -dimensionnelle et d'ordre  $n$  variable (application du processus de construction - selon 5 configurations). Comparaison avec la courbe de Hilbert (algorithme de Butz). Les chiffres en gras correspondent aux meilleurs niveaux de localités

Courbe	$M_1$	$M_2$
Hilbert	8.434138	59.0
Notre proposition	BB	8.434138
	BS	8.436178
	CB	<b>8.431279</b>
	CS	8.436178
	T	8.435158

$d = 3, n = 2$

Courbe	$M_1$	$M_2$
Hilbert	<b>28.667378</b>	247.0
Notre proposition	BB	<b>28.667378</b>
	BS	28.667798
	CB	28.679817
	CS	28.681465
	T	28.710273

$d = 4, n = 2$

Courbe	$M_1$	$M_2$
Hilbert	33.261915	471.0
Notre proposition	BB	33.261915
	BS	33.269824
	CB	<b>33.249778</b>
	CS	33.270267
	T	33.275373

$d = 3, n = 3$

Courbe	$M_1$	$M_2$
Hilbert	<b>226.085288</b>	3951.0
Notre proposition	BB	<b>226.085288</b>
	BS	226.088161
	CB	226.179165
	CS	226.192087
	T	226.421427

$d = 4, n = 3$

Courbe	$M_1$	$M_2$
Hilbert	132.809301	3767.0
Notre proposition	BB	132.809301
	BS	132.840889
	CB	<b>132.760701</b>
	CS	132.842745
	T	133.149893

$d = 3, n = 4$

Courbe	$M_1$	$M_2$
Hilbert	<b>1803.318995</b>	63215.0
Notre proposition	BB	<b>1803.318995</b>
	BS	1803.341506
	CB	1804.056001
	CS	1804.158155
	T	1805.982096

$d = 4, n = 4$

### 6. Conclusion

Dans cet article, nous proposons une nouvelle définition d'une courbe remplissant un espace multidimensionnel. Son contour est suffisamment large pour que plusieurs courbes la satisfassent. Elle constitue par divers aspects une généralisation de la définition de la courbe de Hilbert. Par exemple, il est désormais possible de choisir un autre motif primitif que le motif RBG standard, utilisé dans la courbe de Hilbert. De plus, à travers les ordres, il existe plusieurs manières de connecter le motif primitif (choix des extrémités). Cette définition est étayée par des éléments de preuve et son application est illustrée à travers de nombreux exemples. Des réflexions sur sa mise en œuvre (implémentation) sont menées ce qui se traduit par la proposition d'algorithmes. Leur exécution délivre une famille de courbes :

1. dont on a vérifié que figurent en dimension 2 la courbe originelle de Hilbert, et en dimensions 3, 4, la courbe dite de Hilbert résultant de l'application de l'algorithme de Butz (Butz, 1971),

2. de niveau de localité, mesuré section 5.2, comparable en moyenne à la courbe de Hilbert. Certaines courbes possèdent même un niveau de localité supérieur (cf. tableau 1).

Notons que les résultats présentés dans la section 5, sont confirmés pour des dimensions supérieures ( $d = 5, 6, 7$ ) y compris lorsque l'algorithme de Butz (Butz, 1971) est substitué par celui de Faloutsos (Faloutsos, Roseman, 1989). En ce sens, notre proposition réalise une généralisation aux dimensions arbitraires de la définition de Hilbert établie en dimension 2. Notons que nous nous sommes penchés sur l'article original de Hilbert (Hilbert, 1891) pour la formuler. Ces travaux s'inscrivent dans le cadre :

1. de la recherche de manières alternatives - aux célèbres courbes de Hilbert, Peano, Lebesgue - de remplir un espace, tout en maintenant un bon niveau de localité. Citons les travaux de Liu (Liu, 2004) dans cette direction,

2. de développement de schémas d'implémentation restant opérationnels au-delà de quelques dimensions, avec pour toile de fond le calcul du plongement.

Ces directions offrent des perspectives en termes d'applications. Dans le domaine du traitement des images, des travaux récents s'appuient sur la construction de courbe remplissant l'espace à des fins de caractérisation (Ebrahim *et al.*, 2009) ou de compression (Liang *et al.*, 2008). Néanmoins ces modèles opèrent dans de faibles dimensions (2, 3). Construire des courbes - remplissant des espaces - de dimensions supérieures pourrait permettre d'aborder d'autres problèmes, comme par exemple la recherche d'images dans de grandes bases. C'est la piste que nous explorons actuellement. Plus précisément, une image est représentée classiquement par diverses caractéristiques (de contours, régions, etc.) encapsulées dans un vecteur, dit vecteur de caractéristique (*feature vector*). Une collection d'images est synthétisée par un ensemble de points  $d$ -dimensionnels évoluant dans l'espace des caractéristiques. Le calcul du plongement de ces points sur une courbe remplissant l'espace (des caractéristiques) est un moyen d'ordonner le long d'une ligne (espace des rangs) les points (les images) en fonction de leur localisation dans l'espace. L'ordre ainsi obtenu contribue à réduire la complexité algorithmique de la recherche d'une image requête. La localité, borne combien de points proches sur la ligne correspondent à des images aux caractéristiques effectivement proches (métrique donnée) et donc visuellement compatibles. Elle peut donc inférer sur les résultats de la recherche et donc sur les performances notamment en termes de précision. Des tests à grandes échelles ( $\sim 15000$  images), menés sur les bases de référence MPEG-7 Core Experiment<sup>8</sup> et GREC 2011 Data Base<sup>9</sup>, semblent confirmer ces hypothèses. Notons que ces tests mettent en jeu plusieurs types d'images (naturelles, graphiques), aux modalités variées (binaire, couleur). Cette approche semble favorablement accueillie par la communauté (Nguyen *et al.*, 2012; Franco *et al.*, 2012).

8. <http://www.imageprocessingplace.com>

9. <http://grec2011.cau.ac.kr>

**Bibliographie**

- Angiulli F., Pizzuti C. (2005). An approximate algorithm for top-k closest pairs join query in large high dimensional data. *Data & Knowledge Engineering*, vol. 53, n° 3, p. 263–281.
- Asano T., Ranjan D., Roos T., Welzl E., Widmayer P. (1997). Space-filling curves and their use in the design of geometric data structures. *Theoretical Computer Science*, vol. 181, n° 1, p. 3–15.
- Bially T. (1969, novembre). Space-filling curves: Their generation and their application to bandwidth reduction. *Information Theory, IEEE Transactions on*, vol. 15, n° 6, p. 658 - 664.
- Bitner J., Ehrlich G., Reingold E. (1976). Efficient generation of the binary reflected gray code and its applications. *Communications of the ACM*, vol. 19, n° 9, p. 517–521.
- Butz A. (1969). Convergence with hilbert's space filling curve. *Journal of Computer and System Sciences*, vol. 3, n° 2, p. 128–146.
- Butz A. (1971). Alternative algorithm for hilbert's space-filling curve. *IEEE Transactions on Computers*, vol. 20, n° 4, p. 424–426.
- Chen H., Chang Y. (2011). All-nearest-neighbors finding based on the hilbert curve. *Expert Systems with Applications*.
- Ebrahim Y., Ahmed M., Abdelsalam W., Chau S. (2009). Shape representation and description using the hilbert curve. *Pattern Recognition Letters*, vol. 30, n° 4, p. 348–358.
- Faloutsos C., Roseman S. (1989). Fractals for secondary key retrieval. In *Proceedings of the eighth acm sigact-sigmod-sigart symposium on principles of database systems*, p. 247–252. New York, NY, USA, ACM. Consulté sur <http://doi.acm.org/10.1145/73721.73746>
- Franco P., Nguyen G., Mullot R., Ogier J. (2012). Space-filling curve for image dynamical indexing. In *27th international symposium on computer and information sciences*, p. 311–319. Paris, France, Springer.
- Gotsman C., Lindenbaum M. (1996). On the metric properties of discrete space-filling curves. *IEEE Transactions on Image Processing*, vol. 5, n° 5, p. 794–797.
- Hilbert D. (1891). Über die stetige abbildung einer linie auf ein flächenstück. *Mathematische Annalen*, vol. 38, n° 3, p. 459–460.
- Lawder J. (2000). *The application of space-filling curves to the storage and retrieval of multi-dimensional data*. Thèse de doctorat non publiée, University of London.
- Lawder J., King P. (2000). Using space-filling curves for multi-dimensional indexing. *Advances in Databases*, p. 20–35.
- Lawder J., King P. (2001). Using state diagrams for hilbert curve mappings. *International Journal of Computer Mathematics*, vol. 78, n° 3, p. 327–342.
- Lera D., Sergeev Y. (2010). Lipschitz and hölder global optimization using space-filling curves. *Applied numerical mathematics*, vol. 60, n° 1-2, p. 115–129.
- Liang J., Chen C., Huang C., Liu L. (2008). Lossless compression of medical images using hilbert space-filling curves abstract. *Computerized Medical Imaging and Graphics*, vol. 32, n° 3, p. 174–182.

- Liu X. (2004). Four alternative patterns of the hilbert curve. *Applied mathematics and computation*, vol. 147, n° 3, p. 741–752.
- Mitchison G., Durbin R. (1986). Optimal numberings of an  $n \times n$  array. *SIAM Journal on Algebraic and Discrete Methods*, vol. 7, n° 4, p. 571–582.
- Moon B., Jagadish H., Faloutsos C., Saltz J. (2001). Analysis of the clustering properties of the hilbert space-filling curve. *IEEE Transactions on Knowledge and Data Engineering*, vol. 13, n° 1.
- Muelder C., Ma K. (2008). Rapid graph layout using space filling curves. *IEEE Transactions on Visualization and Computer Graphics*, p. 1301–1308.
- Nguyen G., Franco P., Mullot R., Ogier J. (2012). Mapping high dimensional features onto hilbert curve: Applying to fast image retrieval. In *21th international conference on pattern recognition*, p. 425–428. Tsukuba, Japan, IEEE.
- Oliveira H. e, Petraglia A. (2011). Global optimization using space-filling curves and measure-preserving transformations. *Soft Computing in Industrial Applications*, p. 121–130.
- Peano G. (1890). Sur une courbe, qui remplit toute une aire plane. *Mathematische Annalen*, vol. 36, n° 1, p. 157–160.
- Perez A., Kamata S., Kawaguchi E. (1992). Peano scanning of arbitrary size images. In *Pattern recognition, 1992. vol. iii. conference c: Image, speech and signal analysis, proceedings., 11th iapr international conference on*, p. 565–568. The Hague, Netherlands, IEEE.
- Sagan H. (1994). *Space-filling curves* (vol. 2). New York, NY, USA, Springer-Verlag.
- Voorhies D. (1991). Space-filling curves and a measure of coherence. *Graphic Gems II*, p. 26-30.
- Yiu M., Tao Y., Mamoulis N. (2008). The b dual-tree: indexing moving objects by space filling curves in the dual space. *The VLDB Journal*, vol. 17, n° 3, p. 379–400.

## Annexe A. Distance et adjacence

Nous posons un hypercube découpé, pour déterminer les sous-hypercubes et choisissons un système de coordonnées binaires dans un repère cartésien avec pour origine un sous-hypercube choisi arbitrairement, les axes sont les arêtes de l'hypercube. Alors, les sous-hypercubes en dimension  $d$  sont notés  $a_1 a_2 \dots a_d$  avec  $a_i \in \{0, 1\}, \forall i = 1, 2, \dots, d$ , le sous-hypercube à l'origine est noté  $\underbrace{00\dots0}_d$ .

DÉFINITION 8 (Distance entre deux points). —

$$\|a_1 a_2 \dots a_d, b_1 b_2 \dots b_d\| = \sum_{i=1}^d |a_i - b_i|$$

Comme  $a_i, b_i \in \{0, 1\} \Rightarrow |a_i - b_i| \leq 1$ , par conséquent,

COROLLAIRE 9. — Si  $a$  et  $b$  sont deux sous-hypercubes d'un même hypercube en dimension  $d$  alors  $\|a, b\| \leq d$ .

Si de plus,  $a$  et  $b$  sont les hypercubes satisfaisant la condition  $C3$ , ils ont  $(d - 1)$  coordonnées communes, ainsi :

**COROLLAIRE 10 (Adjacence).** — *Dans un hypercube découpé, si les sous-hypercubes  $a$  et  $b$  partagent une hyperface commune alors  $\|a, b\| = 1$ .*

## Annexe B. Générer tous les motifs primitifs en dimension $d$ arbitraire

---

**Algorithme 1** Chercher un motif primitif  $d$ -dimensionnel entre deux extrémités  $a = a_1a_2\dots a_d$  et  $b = b_1b_2\dots b_d$ .

---

**Entrée:** Dimension  $d$ , deux extrémités  $a = a_1a_2\dots a_d$  et  $b = b_1b_2\dots b_d$ .

**Sortie:** `Motif` : un motif primitif  $d$ -dimensionnel.

```

1:  $P = \{p_1, p_2, \dots, p_n, p_i < p_{i+1}\} \leftarrow$  tous les positions  $p_i : a_{p_i} \neq b_{p_i}$ 
2: si  $n$  est paire alors
3:   quitter
4: fin si
5:  $p^* \leftarrow$  une position quelconque parmi  $P$ 
6:  $P^- \leftarrow P \setminus \{p^*\}$ 
7:  $C = \{c_1, c_2, \dots, c_m : c_i < c_{i+1}\} \leftarrow m$  impaire positions quelconques de  $P^-$ 
8:  $R = \{r_1, r_2, \dots, r_{d-m-1} : r_i < r_{i+1}\} \leftarrow P^- \setminus C$ 
9:  $a' = a'_1a'_2\dots a'_{d-1} \leftarrow a_1\dots a_{p^*-1}a_{p^*+1}\dots a_d$ 
10:  $a'' = a''_1a''_2\dots a''_{d-1} \leftarrow a'$ 
11: pour  $c_i \in C$  faire
12:    $a''_{c_i} \leftarrow 1 - a''_{c_i}$ 
13: fin pour
14:  $b' = b'_1b'_2\dots b'_{d-1} \leftarrow a''$ 
15: pour  $r_i \in R$  faire
16:    $b'_{r_i} \leftarrow 1 - b'_{r_i}$ 
17: fin pour
    /* application de cet algorithme pour la dimension  $d - 1$  */
18: sousMotif1  $\leftarrow$  un motif entre  $a'$  et  $a''$ 
19: sousMotif2  $\leftarrow$  un motif entre  $b'$  et  $b''$ 
20: Ajouter  $a_{p^*}$  à la position  $p^*$  de chacun des points composant le sousMotif1
21: Ajouter  $b_{p^*}$  à la position  $b^*$  de chacun des points composant le sousMotif2
22: Motif  $\leftarrow$  joindre sousMotif1 et sousMotif2

```

---

*Giap Nguyen est doctorant à l'Université de La Rochelle depuis Octobre 2009. Il est membre du Laboratoire Informatique, Image et Interaction (L3i). Ses recherches actuelles portent sur les courbes remplissant l'espace et leurs applications dans le traitement d'images.*

*Patrick Franco est Maître de Conférence à l'Université de La Rochelle depuis Septembre 2004. Il est responsable de parcours du Master Informatique ICONÉ (Ingénierie des Contenus Numériques en Entreprise) et également responsable des stages. Menées au sein du L3I, ses recherches portent sur les courbes fractales remplissant l'espace, et plus largement, sur le développement de signatures pour l'image et leurs applications dans des problèmes gravitant autour de la reconnaissance de formes.*

*Rémy Mullot est Professeur à l'Université de La Rochelle depuis Septembre 2000. Il a été directeur du laboratoire L3i de 2007 à 2011. Ses recherches traitent de l'analyse de documents numériques et sont menées dans l'équipe Systèmes d'Informations Documentaires (SID) dont il est co-fondateur depuis 2001. Il est membre du Groupe de Recherche en Communication Écrite depuis 1998, et également animateur du groupe de recherche SCATI du GDR ISIS depuis 2003. Les principales contributions portent avant tout sur l'utilisation d'approches textures pour l'analyse de documents anciens, sans connaissance a priori.*

*Jean-Marc Ogier a obtenu son diplôme de doctorat en informatique de l'Université de Rouen en 1994, dans le cadre d'un partenariat avec la société Matra Systèmes et Informations. Depuis 2001, il est professeur titulaire à l'Université de La Rochelle. Les intérêts de ses recherches actuelles portent sur la reconnaissance, l'interprétation, et l'indexation de documents numériques. Jean-Marc Ogier est directeur du laboratoire L3I de l'Université de La Rochelle, qui comporte 90 chercheurs. Il a été le leader national ou international de plusieurs projets nationaux financés par l'Agence Nationale pour la recherche nationale (MADONNE, NAVIDOMASS...) et par la communauté européenne (Eureka !). Il est directeur adjoint du GDR I3 du CNRS depuis 2005 (structure regroupant environ 800 personnes). Il est le président général du comité technique numéro 10 (reconnaissance graphique - 300 chercheurs dans le monde) de l'Association internationale pour la Reconnaissance des Formes (IAPR), et est membre du Governing Board de l'IAPR (association internationale scientifique regroupant plus de 10 000 chercheurs dans le monde), représentant la France. Il est également vice-président de l'Université de La Rochelle.*