

## Sélection d'attributs pour la classification d'objets 3D

François Meunier, Christophe Marsala, Cyril De Runz, Laurent Castanié

► **To cite this version:**

François Meunier, Christophe Marsala, Cyril De Runz, Laurent Castanié. Sélection d'attributs pour la classification d'objets 3D. LFA 2017 - 26èmes Rencontres Francophones sur la Logique Floue et ses Applications, Oct 2017, Amiens, France. <hal-01627024>

**HAL Id: hal-01627024**

**<http://hal.upmc.fr/hal-01627024>**

Submitted on 31 Oct 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Sélection d'attributs pour la classification d'objets 3D

## Feature selection for 3D-object classification

F. Meunier<sup>1,2</sup>

C. de Runz<sup>1</sup>

C. Marsala<sup>1</sup>

L. Castanié<sup>2</sup>

<sup>1</sup> Sorbonne Universités, UPMC Univ Paris 06, CNRS, LIP6 UMR 7606

<sup>2</sup> Total Exploration-Production, Tour Coupole, La Défense

### Résumé :

Dans cet article, une nouvelle approche de sélection d'attributs pour la classification est introduite. Elle est basée sur l'utilisation d'un degré de typicalité afin de mettre en avant des sous-parties prototypiques des classes. Combinée à une méthode existante de classification supervisée d'objets 3D, son efficacité est validée expérimentalement et permet une amélioration non négligeable des scores de prédiction.

### Mots-clés :

Classification supervisée, sélection d'attributs, typicalité, prototype, objets 3D.

### Abstract:

In this paper, we propose a new feature selection approach for classification. This proposal is based on the use of typicality degrees, in order to highlight some prototypical sub-parts for classes. Associated with an existing method of classification for 3D objects, its effectiveness is experimentally proved and allows a significant improvement of the prediction scores.

### Keywords:

Supervised classification, feature selection, typicality, prototype, 3D objects.

## 1 Introduction

La sélection d'attributs, réalisée en prétraitement d'un processus d'apprentissage artificiel, permet généralement une amélioration notable de ses performances. En effet, un nombre d'attributs trop important peut handicaper la tâche d'apprentissage, l'augmentation de la dimensionnalité du problème entraînant une instabilité du modèle appris [4].

Dans cet article, nous nous plaçons dans le cadre de la classification d'objets représentés dans un espace à 3 dimensions (3D) à l'aide d'un maillage et répartis en différentes classes

comme, par exemple, des objets fermés (visages, avions, bâtiments,...) ou des objets ouverts (géo-modèles,...). Le but est la mise en évidence d'éléments permettant de classer efficacement un objet quelconque. Les méthodes actuelles de traitement d'objets 3D utilisent souvent des descripteurs mettant en évidence des particularités de l'objet [1, 3, 17]. Ainsi, chaque objet est représenté en utilisant un modèle d'objet global. Ce type d'approches s'applique très bien dans le cas où l'objet est décrit par une structure fermée mais n'est pas adapté au traitement des structures ouvertes [17].

D'autre part, de telles approches ne mettent pas en évidence des propriétés inhérentes aux objets, comme des sous-parties caractéristiques, permettant de discriminer les classes. Dans certaines applications, comme par exemple dans le domaine de la recherche pétrolière dans les couches géologiques représentées par des géo-modèles 3D, il est important de pouvoir discriminer les différents types de failles, formes géométriques particulières, afin de permettre aux géologues de les détecter plus facilement, rendant plus aisée l'étude de la zone [9]. La mise en avant de sous-parties caractéristiques des classes offre un modèle interprétable de classification.

Dans notre approche, un ensemble de sous-parties de taille variable déterminées aléatoirement est extrait des objets 3D. Une

phase d'apprentissage est alors mise en œuvre sur les sous-parties des objets de la base d'apprentissage donnée afin de ne conserver que les sous-parties les plus pertinentes pour la reconnaissance de la classe d'un objet et obtenir ainsi des sous-parties caractéristiques de chaque classe. Toutes les sous-parties n'ont pas le même pouvoir de discrimination pour la reconnaissance de la classe d'objet et, à l'instar des approches de sélection d'attributs classiques, il est donc important de ne conserver que les sous-parties les plus discriminantes vis à vis des classes. Dans cet article, nous introduisons une nouvelle approche pour déterminer les sous-parties pertinentes en introduisant l'utilisation d'une mesure de typicalité pour la mise en évidence de sous-parties prototypiques de chaque classe [6, 14].

Le plan est le suivant. Tout d'abord, en Section 2, un rappel de l'approche 3DRESC de classification d'objets 3D est fait. Ensuite, les principales approches de sélection d'attributs sont présentées dans la Section 3. L'approche introduisant l'utilisation d'un degré de typicalité est présentée en Section 4. En Section 5, un ensemble d'expérimentations pour étudier les apports de la nouvelle approche est présenté. Puis, nous terminons en Section 6 par une conclusion et des perspectives.

## 2 La classification d'objets 3D

En classification, certaines structures de données nécessitent un traitement particulier, impossible à réaliser par les méthodes classiques, ainsi dans le cas des objets 3D, les approches actuelles ne permettent pas de comprendre quelles sous-parties d'un objet peuvent être discriminantes pour reconnaître sa classe. En effet, ces approches font le plus souvent appel à des descripteurs qui extraient des particularités globales de l'objet [1]. Pour remédier à cela, un certain nombre de questions se posent alors : Comment représenter les phénomènes locaux à détecter ? Comment classer les objets 3D ? Comment tenir compte des

sous-parties comme descripteurs ? Comment permettre à l'utilisateur de comprendre le choix du classifieur ?

**Approches par shapelets.** Proposées par [11], les approches par *shapelets* ont été introduites en traitement d'images. Dans ce type d'approches, une décomposition linéaire de chaque image en une série de fonctions élémentaires (des shapelets) décrivant des sous-parties locales est réalisée. À partir de données complexes, une telle approche offre une version simplifiée des données en les décomposant en entités plus simples à étudier. Cette méthode a été adaptée pour la classification supervisée de séries temporelles [18]. Dans l'approche shapelets, un ensemble des sous-parties possibles est extrait de chaque série temporelle de la base d'apprentissage, puis les sous-parties les plus discriminantes pour la classe à reconnaître sont déterminées. Un des avantages de cette méthode est qu'il est possible de comparer des séries temporelles de différentes tailles sans aucune connaissance *a priori*. Par contre, un inconvénient est que la sélection des sous-parties candidates s'effectue avec un temps de calcul prohibitif. Une amélioration notable a alors été d'introduire une sélection aléatoire des sous-parties afin de trouver celles qui sont potentiellement discriminantes [12].

**Classification d'objets 3D : l'algorithme 3DRESC.** L'efficacité de l'approche par shapelets pour la classification de séries temporelles a amené son utilisation pour d'autres types de données, en particulier les objets 3D. L'utilisation de descripteurs locaux [1, 3] et de mesures de distances de répartition [5, 7] offrent d'excellents résultats et permettent d'obtenir en complément, une visualisation des extraits les plus pertinents utilisés lors de cette classification, ce qui offre donc un moyen interprétable pour comprendre la classification. Les différentes étapes de la classification sont : 1) extraction aléatoire des sous-parties d'un objet donné ; 2) sélection des sous-parties les plus discriminantes selon un indicateur choisi (gain

d'information par exemple); 3) utilisation des sous-parties sélectionnées pour classer de nouveaux objets.

L'algorithme 3DRESC (*3D Random Extraction of Sub-parts for Classification*) mettant en œuvre tout ce processus a été introduit par [8]. On considère ici des objets 3D, représentés sous la forme d'un maillage triangulaire, associés à une classe. Pour chaque objet, un certain nombre de points sont choisis aléatoirement sur son maillage et une zone de taille variable (aléatoire) est déterminée autour de chacun de ces points. On obtient ainsi une sous-partie de l'objet considéré. Dans la seconde phase, on utilise l'espace de représentation créé à partir des sous-parties retenues : les valeurs du degré d'appariement de chaque objet à chaque sous-partie discriminante forment les attributs de description des objets 3D pour la phase de classification.

Durant la phase de sélection, afin de déterminer le degré d'appariement de chaque sous-partie à chaque objet, la méthode propose d'utiliser un descripteur de chaque sous-partie et de l'associer à une mesure de proximité. Le degré d'appariement entre une sous-partie  $S$  et un objet 3D  $O_j$  (parmi  $m$ ) est alors défini comme la distance minimale séparant leurs descripteurs :

$$Degre(S, O_j) = \min_{k \in [1, m]} (Dist(S, Extrait_{O_j, k}))$$

avec  $Extrait_{O_j, k}$  la  $k^e$  sous-partie extraite de l'objet  $O_j$  et  $Dist$  une mesure de distance entre une surface et un extrait. D'après une évaluation expérimentale des mesures de distances et des descripteurs proposée dans [10], il apparaît que la combinaison du descripteur appelé *Heat Kernel Signature* [16] et de la distance de *Matusita* [7] fournit de meilleurs résultats <sup>1</sup>.

1. Étant donné  $p = \{p_1, \dots, p_N\}$  et  $p' = \{p'_1, \dots, p'_N\}$ , deux distributions de probabilité, la distance de Matusita entre  $p$  et  $p'$  est définie par :

$$Matusita(p, p') = \sum_{i=1}^N (\sqrt{p(i)} - \sqrt{p'(i)})^2$$

Une fois le degré d'appariement défini, il nous faut maintenant déterminer comment faire la sélection des sous-parties les plus pertinentes. La section 3 présente différentes approches de la littérature pour la sélection d'attributs.

### 3 Approches pour la sélection d'attributs

Dans l'approche 3DRESC, les sous-parties pertinentes extraites des objets 3D jouent le même rôle que des attributs de description. La sélection de ces sous-parties pertinentes peut se comparer à une approche de sélection d'attributs. En effet, en apprentissage, la présence d'un trop grand nombre d'attributs peut réduire la performance du modèle appris lors de son utilisation en classification. Certains attributs peuvent avoir un effet néfaste sur l'apprentissage et il est nécessaire de sélectionner seulement les plus pertinents.

En sélection d'attributs deux grandes catégories d'approches peuvent être distinguées [2] : les approches de *sélection par filtrage* qui utilisent exclusivement la corrélation entre l'attribut évalué et la variable à prédire; et les *méthodes d'ensemble* qui, elles, considèrent des ensembles d'attributs candidats, et donc sélectionnent un sous-ensemble des attributs possibles en tenant compte des interactions entre ces derniers. Les méthodes de sélection par filtrage sont généralement plus intéressantes par rapport aux méthodes d'ensemble lorsqu'un grand nombre d'attributs est à traiter.

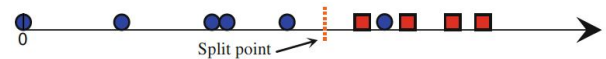


Figure 1 – Point de coupure sur l'axe des distances.

**Sélection par filtrage.** L'approche la plus courante utilise le gain d'information. Elle consiste à sélectionner les attributs selon leur pouvoir prédictif vis à vis de la classe. Dans le cadre des shapelets, pour chaque attribut, la distance qui

le sépare des autres éléments (dont on connaît la classe d'appartenance) est calculée et sert à les ordonner du plus proche au plus éloigné (cf. Figure 1). La capacité de discrimination de l'attribut est évaluée par le gain d'information maximal obtenu en prenant un point de coupure entre 2 éléments successifs sur cet axe [18].

**Méthodes d'ensemble.** Pour ce type d'approches, toutes les possibilités de sous-ensembles d'attributs possibles sont examinées afin de sélectionner le sous-ensemble optimisant l'information mutuelle avec la classe. L'inconvénient majeur de cette démarche est qu'elle implique de calculer cette information mutuelle pour un grand nombre d'attributs candidats, ce qui devient très coûteux en présence d'un très grand nombre d'attributs.

Pour pallier ce problème, des méthodes visant à réduire le temps de calcul des méthodes d'ensemble ont été imaginées. Ces méthodes sont dites *intégrées*. L'objectif est de chercher des attributs ayant un fort caractère discriminant pour la classification visée tout en étant peu corrélés avec les attributs déjà choisis.

## 4 Sélection par typicalité

L'algorithme 3DRESC présenté dans la Section 2 utilise une sélection d'attributs basée sur le gain d'information. Dans l'approche par *shapelets* [18], d'autres techniques de sélection d'attributs du même type ont été proposées, comme par exemple, une sélection aléatoire [13].

La nouvelle approche proposée dans cet article introduit l'utilisation d'un degré de typicalité [6, 14, 15] pour réaliser la sélection de sous-parties discriminantes. Dans ce cadre, le prototype d'une classe constitue l'élément le plus représentatif de cette classe. À partir d'un ensemble d'éléments, la construction d'un prototype s'effectue de la manière suivante :

1. calculer la ressemblance d'un élément avec les autres éléments de sa classe, et la dissimilarité avec les éléments des

autres classes et agréger ces 2 valeurs pour constituer le *degré de typicalité* de l'élément vis à vis de sa classe ;

2. ordonner les éléments d'une même classe en fonction de leur degré de typicalité ;
3. agréger les objets qui sont les plus typiques de leur classes afin de créer un prototype flou.

Dans cet article, le troisième point n'est pas considéré car nous nous concentrons seulement sur la modification de l'étape de sélection de sous-parties de l'algorithme 3DRESC.

Le calcul du degré de typicalité d'un élément se base sur le choix d'une mesure de ressemblance, d'une mesure de dissimilarité et d'un opérateur d'agrégation. Nous utilisons ici les mesures proposées dans [6].

Soit  $X = \{x_1, x_2, \dots, x_n\}$  un ensemble d'éléments et  $Y$  un ensemble de classes associées. Pour tout  $x_i \in X$ ,  $i = 1, \dots, n$ , on note  $y_i$  la classe associée à  $x_i$ . Étant donné  $x_i \in X$ , on note  $C_i$  l'ensemble des exemples de  $X$  qui possèdent la même classe que  $x_i$  :  $C_i = \{x_j \in X \mid x_j \neq x_i \text{ et } y_j = y_i\}$ . Dans la suite, on note  $\bar{C}_i = \{x_j \in X \mid x_j \notin C_i\}$ .

Soit  $\delta$ , une mesure de distance définie sur  $X$ . La dissimilarité  $d(x_i, x_j)$  entre deux éléments  $x_i$  et  $x_j$  de  $X$  est donnée par :

$$d(x_i, x_j) = \frac{\delta(x_i, x_j) - d_m}{d_M - d_m} \quad (1)$$

avec  $d_m$  et  $d_M$  des paramètres de normalisation fixés, par exemple, à 0 pour  $d_m$  et au maximum des distances entre élément de  $X$  pour  $d_M$ .

La ressemblance entre deux éléments  $x_i$  et  $x_j$  de  $X$  peut se déterminer à partir de la mesure de dissimilarité de deux façons différentes :

$$r(x_i, x_j) = 1 - d(x_i, x_j) \quad (2)$$

$$r(x_i, x_j) = \frac{1}{1 + d(x_i, x_j)^\gamma} \quad (3)$$

avec  $\gamma$  un paramètre fixé par l'utilisateur. Par exemple, dans cet article nous choisissons  $\gamma = 2$  pour nos premières expérimentations.

Le degré de typicalité d'un élément  $x_i \in X$ , comme par exemple celui proposé par [6], se calcule à partir de la ressemblance globale  $R$  de  $x_i$  à tous les éléments de sa classe et de sa dissimilarité globale  $D$  à tous les éléments qui ne sont pas de sa classe.

$$R(x_i, C_i) = \frac{1}{|C_i|} \sum_{x_j \in C_i} r(x_i, x_j)$$

et

$$D(x_i, C_i) = \frac{1}{|\bar{C}_i|} \sum_{x_j \in \bar{C}_i} d(x_i, x_j)$$

Le degré de typicalité  $T$  vis à vis de sa classe pour un élément  $x_i \in X$  se détermine en agrégeant ces deux mesures globales. Par exemple, on considère la mesure d'agrégation proposée dans [6] :

$$T(x_i, C_i) = \frac{R(x_i, C_i) + D(x_i, C_i)}{R(x_i, C_i) * D(x_i, C_i)} \quad (4)$$

D'autres choix pour les mesures de ressemblance, de dissimilarité et d'agrégation sont possibles et feront l'objet d'études ultérieures.

Au final, dans notre cadre, pour l'application de l'algorithme 3DRESC, les sous-parties de chaque objet sont ordonnées à l'aide de leur degré de typicalité vis à vis de leur classe.

## 5 Expériences

Les données utilisées pour évaluer cette nouvelle combinaison d'approches sont issues du *Princeton Shape Benchmark*<sup>2</sup> (PSB), dont les données proviennent d'ensembles d'objets 3D. Ces derniers sont ceux qui, en termes de formats et de types, se rapprochent le plus de

géo-modèles 3D qui seront, à terme, utilisés. Les classes les plus fines sont retenues (3 niveaux étant possibles), afin de rajouter encore davantage de contraintes et de difficultés à la prédiction lors des tests. Lors de ces tests, un même nombre d'objets issus de deux classes différentes sont sélectionnées aléatoirement. Une fois les sous-parties sélectionnées, nous utilisons l'algorithme des forêts aléatoires de la librairie Python *scikit-learn* pour construire le classifieur. Une validation croisée (*4-folds cross-validation*) est effectuée afin de vérifier la stabilité des résultats. La mesure utilisée pour évaluer la performance est l'*AUC* (*Area Under Curve ROC*). Les expérimentations ont été faites avec la mesure de dissimilarité de l'équation (1), les mesures de ressemblance des équations (2) et (3), et comme fonction d'agrégation, celle de l'équation (4).

Dans la table 1, les résultats de la comparaison de différentes approches de sélections sont présentés : une méthode de sélection aléatoire (comme *baseline*), la méthode de sélection utilisant le gain d'information et la méthode de sélection utilisant le degré de typicalité (en variant la mesure de ressemblance).

table 1 – Comparaison des approches

Méthode de sélection	Score (AUC)
Sélection aléatoire	0.594 ± 0.156
Gain d'information	0.762 ± 0.097
Deg. de typicalité (eq. (2))	0.622 ± 0.081
Deg. de typicalité (eq. (3))	0.757 ± 0.095

L'approche utilisant un degré de typicalité s'avère un peu moins performante que l'approche par gain d'information. Par contre, le choix de la mesure de ressemblance s'avère crucial. La mesure présentée dans l'équation (3) offrant de meilleurs résultats que celle de l'équation (2). Pour la suite, les expérimentations sont réalisées en utilisant cette seule mesure de ressemblance de l'équation (3).

2. <http://shape.cs.princeton.edu/benchmark/> (page consultée le 06/05/2017).

Un deuxième ensemble d'expérimentation permet de mettre en évidence l'apport de ce nou-

table 2 – AUC des classifications avec combinaisons de sélection d’attributs

Répartition des sous-parties sélect. (en % du total)		Nombre d’objets total			
G. I.	Typicalité	40 objets	60 objets	80 objets	100 objets
100%	0%	0.737 ± 0.097	0.762 ± 0.097	0.783 ± 0.097	0.801 ± 0.097
0%	100%	0.741 ± 0.099	0.761 ± 0.093	0.786 ± 0.091	0.803 ± 0.094
<b>50%</b>	<b>50%</b>	<b>0.787 ± 0.105</b>	<b>0.813 ± 0.102</b>	<b>0.821 ± 0.101</b>	<b>0.836 ± 0.094</b>
75%	25%	0.771 ± 0.111	0.805 ± 0.102	0.812 ± 0.088	0.827 ± 0.089
25%	75%	0.773 ± 0.108	0.799 ± 0.089	0.802 ± 0.101	0.824 ± 0.085

veau type de sélection par degré de typicalité en le combinant avec l’approche par gain d’information. Dans ce cas, on étudie l’apport que peut avoir l’utilisation conjointe de sous-parties trouvées par chacune des 2 approches. Ainsi, les sous-parties sélectionnées sont issues soit de l’approche par gain d’information (*G.I.*), soit de l’approche par typicalité. L’idée consiste donc à utiliser l’union des meilleurs extraits de chaque méthode de sélection d’attributs afin de couvrir au mieux les possibilités d’éléments pertinents selon chaque critère. Ainsi, les  $n$  extraits les plus pertinents sont choisis entre les mieux notés par différentes évaluations, selon une répartition (en pourcentage de l’ensemble) déterminée. Afin d’obtenir une vision d’ensemble des effets de cette combinaison, plusieurs proportions de sous-parties choisies pour chaque méthode de sélection ont été testées.

De cette manière, on remarque que les résultats s’améliorent (table 2), et permettent de confirmer l’intérêt d’une sélection par typicalité. L’amélioration des performances peut s’expliquer par l’introduction d’une importante diversité grâce à des critères de sélection différents. On constate que les répartitions de 75% – 25% améliorent la prédiction, mais n’égale pas celle d’un choix équilibré (50% – 50%). Cet équilibre s’avère être le meilleur compromis pour la prédiction.

Pour mieux comprendre les raisons d’une telle amélioration des résultats, il est présenté en table 3 le rapport du nombre de doublons au sein des extraits conservés sur le nombre

table 3 – Importance des doublons

Nombre d’objets	Ratio de doublons
40	0%
60	0%
80	0.5%
100	0%

total d’extraits dans le cas d’une répartition équilibrée (50% – 50%). On appelle doublon une sous-partie sélectionnée qui a été mise en avant à la fois par la méthode par gain d’information et par la méthode par typicalité. Les résultats permettent de se rendre compte que la présence de doublons est très faible et que le gain de performance est bien dû à la variété introduite par l’utilisation des approches de sélection.

Pour confirmer le fait que les 2 approches de sélection sont peu concordantes et relèvent bien des sous-parties très peu similaires la distribution des valeurs obtenues pour chaque sous-partie par chacune de ces deux méthodes de sélection est présentée dans la Figure 2).

Pour vérifier la non-corrélation des valeurs de gain d’information et de typicalité, une projection est présentée dans la Figure 3. Chaque point de ce tracé correspond à une sous-partie pour laquelle la valeur en abscisses est le gain d’information obtenu, et la valeur en ordonnée est son degré de typicalité.

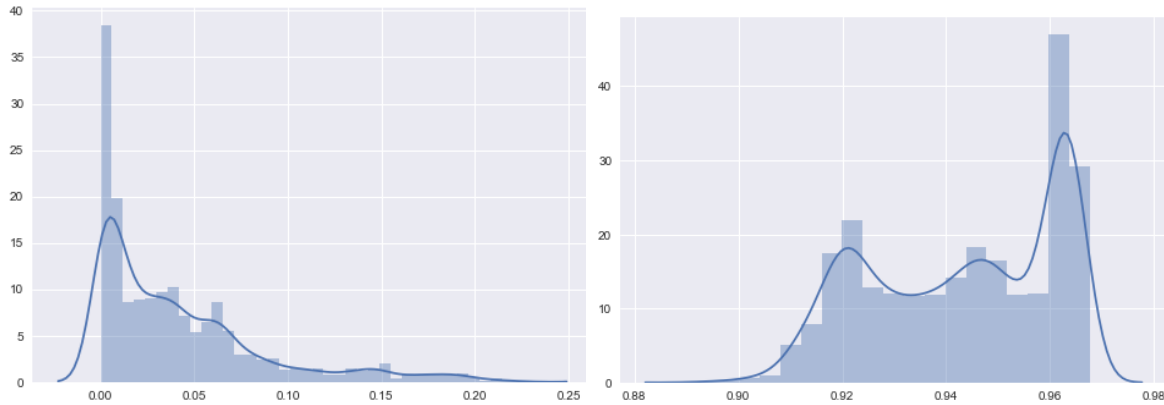


Figure 2 – Distribution des valeurs de gain d'information (gauche) ou de typicalité (droite)

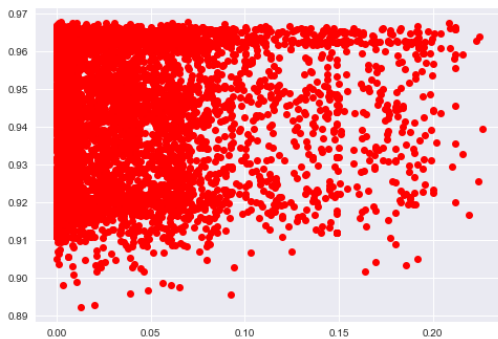


Figure 3 – Gain d'information vs typicalité

## 6 Conclusion

Dans cet article, nous avons proposé une nouvelle méthode de sélection d'attributs pertinents pour la classification supervisée d'objets 3D. Cette approche combine l'utilisation du gain d'information et d'un degré de typicalité pour retenir les attributs les plus pertinents pour la reconnaissance de la classe. Ainsi, l'introduction d'une plus grande diversité dans le choix des attributs sélectionnés est un critère pour l'amélioration des performances.

Une étude plus approfondie, en particulier des différences d'ordonnement provoquées par chacune de ces approches reste à mener. D'autres mesures de ressemblance, de dissimilarité ou d'agrégation doivent être encore étudiées afin de mettre en évidence d'éventuels apports pour l'amélioration des performances.

Finalement, une utilisation de cette approche est prévue dans le cadre d'une application réelle pour Total EP dans le domaine de la recherche pétrolière pour la reconnaissance de couches géologiques représentées par des géo-modèles 3D.

## 7 Remerciements

Ce travail a été réalisé dans le cadre d'une collaboration entre Total EP et l'UPMC. Cyril de Runz a participé à ce travail durant sa délégation CNRS au LIP6.

## Références

- [1] L. A. Alexandre. 3D descriptors for object and category recognition : a comparative evaluation. In *Workshop on Color-Depth Camera Fusion in Robotics at the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 1–6, 2012.
- [2] G. Chandrashekar and F. Sahin. A survey on feature selection methods. *Computers and Electrical Engineering*, pages 16–28, 2014.
- [3] Q. V. Dang. *Similarités dans des modèles BRep paramétriques : détection et applications*. PhD thesis, Université de Toulouse, 2014.
- [4] J. H. Friedman. On bias, variance, 0/1-loss, and the curse-of-dimensionality.



- Data mining and knowledge discovery*, 1(1) :55–77, 1997.
- [5] S. Kullback and R. A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22(1) :79–86, 1951.
- [6] M.-J. Lesot, M. Rifqi, and B. Bouchon-Meunier. Fuzzy prototypes : From a cognitive view to a machine learning principle. In H. Bustince, F. Herrera, and J. Montero, editors, *Fuzzy Sets and Their Extensions : Representation, Aggregation and Models*, pages 431–452. Springer, 2008.
- [7] K. Matusita. Decision rules, based on the distance, for problems of fit, two samples, and estimation. *Annals of Mathematical Statistics*, 26(4) :631–640, 1955.
- [8] F. Meunier and C. Marsala. Classification d’objets 3D par extraction de sous-parties discriminantes. In *La 5ème édition des journées ” Big Data Mining and Visualization ”*, Metz, France, June 2016.
- [9] F. Meunier, C. Marsala, and L. Castanie. Development of a recommendation system for decision support in geosciences. In *MATHIAS – TOTAL Symposium on Mathematics*, Paris, France, Oct. 2015.
- [10] F. Meunier, C. Marsala, L. Castanié, and B. Conche. Classification d’objets 3d par extraction aléatoire de sous-parties discriminantes pour l’étude du sous-sol en prospection pétrolière. In *Conf. Int. Francophone sur l’Extraction et la Gestion des Connaissances (EGC)*, pages 225–236, 2017.
- [11] A. Refregier. Shapelets : I. a method for image analysis. *Monthly Notices of the Royal Astronomical Society*, 338 :35, 2001.
- [12] X. Renard, M. Rifqi, and M. Detyniecki. Random-shapelet : an algorithm for fast shapelet discovery. In *IEEE Int. Conf. on Data Science and Advanced Analytics (DSAA’2015)*, pages 1–10, 2015.
- [13] X. Renard, M. Rifqi, G. Fricout, and M. Detyniecki. EAST representation : fast discovery of discriminant temporal patterns from time series. In *ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data*, 2016.
- [14] M. Rifqi. *Mesures de comparaison, typicalite et classification d’objets flous : théorie et pratique*. PhD thesis, Université Pierre et Marie Curie, Paris 6, 1996.
- [15] M. Rifqi. Mesures de similarité, raisonnement et modélisation de l’utilisateur. Technical report, Université Pierre et Marie Curie, Paris 6, 2010. Habilitation à diriger des recherches.
- [16] J. Sun, M. Ovsjanikov, and L. Guibas. A concise and provably informative multi-scale signature based on heat diffusion. In *Symposium on Geometry Processing, SGP’09*, pages 1383–1392. Eurographics Association, 2009.
- [17] H. Tabia. *Contributions to 3D-shape matching retrieval and classification*. PhD thesis, Université Lille 1, 2011.
- [18] L. Ye and E. Keogh. Time Series Shapelets : A New Primitive for Data Mining. In *15th ACM SIGKDD Int. Conf. on Knowl. Discovery and Data Mining, KDD ’09*, pages 947–956, 2009.