

Explorer3D

Manuel d'Utilisation

Version 1.2

Matthieu EXBRAYAT, Lionel MARTIN
LIFO, Université d'Orléans

3 janvier 2013

Table des matières

1	Présentation générale	4
2	L'essentiel...	4
3	Installation	5
3.1	Prérequis	5
3.2	Installation et démarrage	5
4	Format des données en entrée	5
4.1	Notion d'attributs complémentaires	5
4.2	Mode multisources	6
4.3	Sous-ensemble d'objets	6
4.4	Fichiers attributs / valeurs	6
4.4.1	Cas général	6
4.4.2	Attribut séparateur spécifique	7
4.4.3	Sous-ensemble d'objets	7
4.4.4	Gestion des attributs descriptifs symboliques	8
4.5	Fichiers de distances	8
4.6	Fichier de coordonnées	9
4.6.1	Sous-ensembles d'objets	10
4.7	Fichiers ne comportant que des attributs complémentaires	10
4.8	Noms d'attributs réservés (attributs complémentaires)	10
4.9	Importation de données	10
5	Présentation de l'interface de visualisation	11
5.1	Prise en main	11
5.2	Menus de la fenêtre de contrôle	13

6	Fonctions générales	15
6.1	Méthodes de projection	15
6.1.1	A partir d'un fichier attributs / valeurs	15
6.1.2	À partir d'une matrice de distance	16
6.1.3	À partir d'une liste de coordonnées 3D	17
6.2	Utilisation/affichage d'informations complémentaires	17
6.2.1	Présentation générale	17
6.2.2	Objets multi-classes	17
6.2.3	Affichage des images	17
6.2.4	Sources d'images multiples	18
6.2.5	Gestion des couleurs	18
6.3	Utilisation de la souris	19
6.3.1	Manipulation de la scène 3D	19
6.3.2	Action sur les objets (raccourcis)	19
6.3.3	Menu contextuel	20
6.4	Sélection d'objets	20
6.4.1	Sélection multiple	20
6.5	Loupe	21
6.6	Zoom sur un sous-ensemble d'objets (Crop)	21
6.7	Options	22
6.8	Légende	23
6.8.1	Légende multiclassées	24
6.8.2	Légende sur attribut continu	24
6.8.3	Légende à base de textures	25
6.9	Axes de projection	25
6.10	Observations	27
6.11	Sauvegarde de projet	27
7	Outils principaux	27
7.1	Manipulation multi-vues	27
7.1.1	Système multi-vues	28
7.1.2	Effet des commandes sur les vues 3D	28
7.1.3	Explorateur de vues	29
7.1.4	Vues et objets sélectionnés	29
7.2	Affichage tabulaire de la description des objets	29
7.3	Clustering	30
7.3.1	Mélange de lois	30
7.3.2	Nuées dynamiques (kmeans)	31
7.3.3	Nuées dynamiques floues	32
7.3.4	Classification par densité (DBSCAN)	32
7.3.5	Arbre minimal de recouvrement et classification hiérarchique	33
7.4	Reconfiguration spatiale	34
7.4.1	Comparaison	34
7.4.2	Anomalie	35
7.4.3	Déplacement	35
7.4.4	Axes d'observation	36
7.5	Exploration à partir des images	36

7.6	SVM	37
7.7	Noyaux	40
7.8	Visualisation 2D	41
7.9	Commande via le réseau	41
7.10	Importation de données	43
8	Outils secondaires	45
8.1	Voyage à travers les dimensions	45

1 Présentation générale

Explorer3D est à l'origine un logiciel de visualisation. Il permet de visualiser des objets dans un espace 2D ou 3D grâce à différentes techniques de projection spatiale :

- Lorsque les objets sont décrits par un ensemble d'attributs, **Explorer3D** procède à une réduction de dimensions supervisée (e.g. Analyse Discriminante) ou non supervisée (e.g. ACP).
- Lorsque l'on dispose d'une information de distance (ou de dissimilarité) entre objets, **Explorer3D** utilise des techniques de positionnement multidimensionnel (MDS).

Différents outils de classification non supervisée sont disponibles, afin d'aider l'utilisateur dans son interprétation graphique.

Explorer3D est globalement orienté vers une utilisation interactive, permettant des extractions de sous-ensembles d'objets, des ajouts de contraintes de positionnement, etc.

2 L'essentiel...

- **Charger un fichier** : aller dans le menu "Fichiers / charger un fichier". Si vos données ne sont pas à un format reconnu par Explorer3D, un outil d'import sera automatiquement ouvert (voir section 4.9).
- **Visualisation d'images associées à vos objets** : pour utiliser cette fonctionnalité, votre fichier source doit comporter une colonne indiquant le nom de l'image associée à chaque objet (voir section 4). Une fois votre fichier chargé (et visualisé), rendez-vous dans la sous-fenêtre "Attributs complémentaires" et indiquez, pour "Données associées", le nom de la colonne correspondant aux noms de fichiers.
- **Affichage permanent des images associées aux points 3D** : faire un clic droit sur le point et sélectionner "Afficher / cacher image". Suivant le paramétrage des options ("Afficher les images hors 3D"), les images seront affichées soit dans une fenêtre à part, soit dans la vue 3D.
- **Affichage contextuel des images associées aux points 3D** : il est possible d'afficher dynamiquement une image lorsque l'on passe sur le point correspondant, et de la cacher lorsque l'on quitte ce point. Pour ce faire, cocher "Affichage dynamique des images" dans les options.
- **paramétrer les options** : menu "Outils / options".
- **Explorer à partir des images** : il est possible de cacher tous les points 3D, puis de ne visualiser que ceux correspondant à des images choisies. Ce mode d'exploration se lance à partir de "Outils / explorer à partir des images".
- **Eloigner / rapprocher des points dans la vue 3D** : L'exploration à partir des images permet de modifier le positionnement relatif des objets dans l'espace. Il existe une deuxième méthode offrant le même service mais par le biais d'une interface différente. Aller dans "Outils / interaction", et consulter la rubrique correspondante pour le mode d'emploi.

3 Installation

3.1 Prérequis

Vous devez disposer d'une version récente de java Standard Edition (> 1.6u12) dotée de la librairie Java3D. JavaSE et Java3D peuvent être respectivement téléchargées depuis :

- <http://www.oracle.com/technetwork/java/javase/downloads/index.html>
- <http://www.oracle.com/technetwork/java/javase/tech/index-jsp-138252.html>

Si java3D est mal installé ou absent, un message d'erreur doit s'afficher. Toutefois, un démarrage qui n'aboutit pas à l'affichage de la fenêtre de contrôle, ou des affichages qui ne se produisent pas correctement, peuvent aussi être un symptôme de l'absence, de la mauvaise utilisation ou d'une version inadaptée de java ou de java3D.

3.2 Installation et démarrage

Explorer3D se présente sous la forme d'un fichier java archive (explorer3d.jar). Il se lance donc grâce à la commande “ `java -jar explorer3d.jar` ”.

Dans divers environnements (e.g. MacOS) on pourra lancer explorer3D en cliquant sur explorer3d.jar. Toutefois, ce mode de lancement risque de masquer les éventuels problèmes de démarrage. Le lancement par double clic ne fonctionne pas sous Windows 7. Une astuce simple consiste à créer un fichier “explorer.bat” dans le répertoire où se trouve explorer3D.jar, et de placer dans ce fichier la commande “ `java -jar explorer3d.jar` ”. Le double clic sur explorer.bat permettra de lancer **Explorer3D** .

Le recours à cette astuce est conseillé également dans le cas d'une installation personnalisée de java3D nécessitant l'utilisation du classpath. L'option -jar interdisant l'accès du classpath, il faut placer explorer3D dans le classpath et lancer directement la classe explorer.Explorer3D. Une bonne solution consiste alors à écrire un script du type (exemple pour shell bash et java3d 64 bits) :

```
export LD_LIBRARY_PATH = $LD_LIBRARY_PATH:chemin_d_acces_j3d/lib/amd64
export CLASSPATH = $CLASSPATH:chemin_d_acces_j3d/lib/ext/j3dcore.jar:\
  chemin_d_acces_j3d/lib/ext/j3dutils.jar:\
  chemin_d_acces_j3d/lib/ext/vecmath.jar:\
  chemin_d_acces_explorer/explorer3D.jar
java explorer.Explorer3D
```

4 Format des données en entrée

4.1 Notion d'attributs complémentaires

Dans les fichiers de données, une partie des informations sert à la projection proprement dite (e.g. attributs de description des objets, matrice de distances, coordonnées 3D). Ces informations peuvent être accompagnées de données complémentaires, indiquant par exemple le nom d'un fichier image associé à un objet, la classe d'un objet dans le cas de données étiquetée, etc. De telles données sont appelées *attributs complémentaires*. Chaque fichier peut comporter 0, 1 ou plusieurs attributs complémentaires.

4.2 Mode multisources

Si l'utilisateur dispose, pour un jeu de données, de plusieurs fichiers sources contenant des descripteurs différents, il peut charger ses différents fichiers et observer simultanément les projections 3D obtenues pour ces différents fichiers.

L'accès à ce mode est contrôlé par les options d'*Explorer3D* (cf. section 6.7). Si cette option est désactivée, chaque chargement de fichier conduit à la réinitialisation de l'environnement et donc à la suppression des projections en cours, car on considère que chaque fichier porte sur des objets différents. Dans le cas contraire, l'utilisateur se voit proposer de charger les nouveaux fichiers dans le projet courant (multi source) ou dans un nouveau projet (réinitialisation).

En mode multisources, les différents fichiers peuvent contenir des données de natures différentes (attributs/valeur, distance, etc.). **les objets doivent se présenter dans le même ordre dans les différents fichiers, sauf si l'on a eu recours à un fichier contenant un sous-ensemble des objets.**

4.3 Sous-ensemble d'objets

Plusieurs des types de données supportés permettent le chargement de données pour un sous-ensemble des objets. Cela a surtout un sens en mode multisources, où l'on peut ne disposer de certaines informations que pour un sous-ensemble des objets, mais où l'on souhaite identifier ces objets avec ceux chargés précédemment à partir d'autres fichiers. Un fichier contenant un sous-ensemble des objets est repéré par la présence du mot-clé SUBSET dans son contenu (voir syntaxe des différents formats de fichiers). Chaque objet doit alors être accompagné de son rang dans l'ensemble complet des objets. Dans la version courante, un fichier contenant tous les objets doit être chargé avant de pouvoir utiliser un fichier de sous-ensemble.

Les fichiers sous-ensemble peuvent contenir des attributs complémentaires. La valeur de ceux-ci sera initialisée à "UNDEFINED" pour les objets manquants.

4.4 Fichiers attributs / valeurs

Il s'agit du cas le plus courant, où des objets sont décrits par une liste d'attributs. La visualisation 3D correspond alors à une projection dans un sous-espace calculé à partir des attributs initiaux.

4.4.1 Cas général

Explorer3D supporte des fichiers au format texte, avec la structuration suivante :

```
[SUBSET [START WITH x]]
nombre d'objets
nombre d'attributs descriptifs
nom des attributs descriptifs et noms des attributs complémentaires
Description des objets (un par ligne)
```

Les noms et valeurs d'attributs sont séparés par un caractère espace (il est toutefois possible de définir explicitement le caractère séparateur utilisé, cf. section 4.4.2).

Voici un exemple de fichier (début de la description du très classique jeu de données "iris") :

```

150
4
A B C D classe
5.1 3.5 1.4 0.2 Iris-setosa
4.9 3.0 1.4 0.2 Iris-setosa
4.7 3.2 1.3 0.2 Iris-setosa
...

```

Dans cet exemple il y a 4 attributs descriptifs, nommés “A”, “B”, “C” et “D”, et un attribut complémentaire nommé “classe”. Les attributs descriptifs sont ici de type numérique (réel). **Remarque : attention à ne pas laisser de caractères non significatifs en fin de ligne, notamment, pour les deux premières lignes, veiller à ne pas laisser de caractères “espace” en fin de ligne.**

4.4.2 Attribut séparateur spécifique

Dans le cas où l’espace ne convienne pas comme séparateur, on peut insérer une ligne dans le fichier de données indiquant le caractère utilisé. Cette information s’insère en troisième ligne, après le nombre d’attributs, et avant la liste de noms d’attributs. Par exemple, si l’on choisi “|” comme séparateur, le fichier devient :

```

150
4
|
A|B|C|D|classe
5.1|3.5|1.4|0.2|Iris-setosa
4.9|3.0|1.4|0.2|Iris-setosa
4.7|3.2|1.3|0.2|Iris-setosa
...

```

4.4.3 Sous-ensemble d’objets

Si “SUBSET” est écrit dans la première ligne du fichier, alors le fichier est considéré comme ne contenant qu’un sous-ensemble des objets. Une valeur numérique doit être fournie au début de chaque ligne (description d’objet) donnant le rang global de cet objet.

La numérotation globale des objets commence à 0 dans **Explorer3D** . Si le système de numérotation de l’utilisateur ne commence pas à 0, SUBSET doit être suivi de “START WITH”, puis de l’offset de l’utilisateur. Par exemple, si l’utilisateur numérote à partir de 1, il doit ajouter “START WITH 1”.

Par exemple :

```

SUBSET START WITH 1
3
4
A B C D classe
5 5.1 3.5 1.4 0.2 Iris-setosa
20 4.9 3.0 1.4 0.2 Iris-setosa
110 4.7 3.2 1.3 0.2 Iris-setosa

```

L'utilisateur ne fournit ici que trois objets, de rang 5, 20 et 110 dans son système de numérotation qui démarre à 1. Il s'agit donc des objets 4, 19 et 109 dans **Explorer3D** .

4.4.4 Gestion des attributs descriptifs symboliques

Par défaut, les attributs de description sont de type numérique (réel). **Explorer3D** supporte cependant les attributs symboliques. Par symbolique, on entend tout attribut ne prenant pas ses valeurs dans un domaine continu (il peut donc s'agir de chaînes, mais aussi d'entiers). Afin d'indiquer leur présence, le nom de l'attribut doit se terminer par ".S". Ceci est illustré par l'exemple suivant :

```
151
5
R1.S R2.S R3.S R4.S R5 Classe
A Vert OUI + 19 VRAI
B Rouge NON - 17 VRAI
C bleu NON - 49 FAUX
...
```

On peut constater que les 4 premiers attributs sont symboliques. Concernant les attributs numériques, une extension ".N" est possible, mais elle est optionnelle.

Concrètement, les attributs symboliques sont traités par binarisation : la liste des valeurs possibles est générée par un premier parcours du fichier de données. Chaque attribut symbolique est ensuite remplacé par une liste d'attributs, correspondant chacun à une valeur de l'attribut symbolique, prenant la valeur 0 ou 1. Ces attributs portent le nom de l'attribut d'origine, avec le suffixe "\$valeur".

Par exemple, dans le jeu de données ci-dessus, R1 est remplacé par deux attributs, "R1\$1" et "R1\$2", qui pour le 1er objet valent respectivement 1 et 0, et pour le second objet 0 et 1. Cette décomposition est transparente pour l'utilisateur, et les "sous attributs" générés ne sont manipulés directement par ce dernier que dans le cadre de tâches assez avancées.

4.5 Fichiers de distances

Il s'agit de fichiers contenant des matrices de distances entre objets. La visualisation correspond alors au placement des objets dans un espace à trois dimensions, tel que les distances des objets dans cet espace soient les plus proches possible de leurs distances dans la matrice fournie. Le format du fichier est le suivant :

```
nombre d'objets [COMPLETE]
Liste des distances entre objets (un objet par ligne)
Nombre ou Liste des attributs complémentaires
Valeurs des attributs complémentaires
```

Les distances entre objets sont données à raison d'un objet par ligne. Par défaut on ne fournit ici que la matrice triangulaire supérieure (si on a trois objets a b et c, et que $d_{i,j}$ est la distance entre les objets i et j, alors le fichier contient, dans cet ordre : $d_{a,b}$ et $d_{a,c}$ sur la première ligne, $d_{b,c}$ sur la seconde ligne).

Si, sur la première ligne, on a indiqué le mot-clé optionnel “COMPLETE”, alors la matrice complète est fournie (en reprenant notre exemple, la première ligne contiendrait $d_{a,a}$, $d_{a,b}$ et $d_{a,c}$, la seconde $d_{b,a}$, $d_{b,b}$ et $d_{b,c}$, etc.).

Les attributs fournis ici (deux dernières lignes du format ci-dessus) sont équivalents aux “attributs complémentaires” décrits dans la section précédente : ils ne servent pas au placement des objets, mais apportent des informations complémentaires. Si on indique un nombre d’attributs plutôt qu’une liste de noms, les attributs sont automatiquement nommés “Att1”, “Att2”, etc.

Voici un exemple de fichier (extraits d’un fichier de distances entre enluminures) :

```
166
...
1.0694574 1.1302139 1.0019832 1.0004523 ...
1.0656028 0.96607274 1.1858556 ...
...
image classe
ms0001_1.jpg ms0001-Mazarine-Fr-SW-Début-12eme
ms0001_2.jpg ms0001-Mazarine-Fr-SW-Début-12eme
...
```

Il y a ici 166 objets. pour chaque objet on dispose de deux informations : le nom d’un fichier image associé, et le nom du document d’origine (ce fichier fournit des distances entre enluminures ; les images ne sont pas stockées dans ce fichier, mais les noms d’images peuvent être utilisés par la suite pour accéder aux fichiers images).

Le premier objet appartient au document “ms0001 - Mazarine - Fr - SW - Début 12eme”, et est accompagné de l’image “ms0001_1.jpg”. Le second provient du même manuscrit et est accompagné de l’image “ms0001_2.jpg”. La distance entre le premier et le second objet est 1.0694574, la distance entre le premier et le troisième objet est “1.1302139”, etc.

4.6 Fichier de coordonnées

Contrairement aux autres formats, ce type de fichier ne commence pas par le nombre d’objets contenus (celui-ci est calculé automatiquement). *Cette disposition est susceptible de changer dans les versions ultérieures.*

Optionnellement, le fichier peut comporter une première ligne indiquant le nom de chaque attribut. En son absence, les noms sont automatiquement attribués.

Chaque ligne suivante décrit un objet. Elle comporte trois réels correspondant aux coordonnées de l’objet. Ces trois réels peuvent être suivis d’attributs complémentaires. Pour l’affichage, les coordonnées seront automatiquement réduites (toute valeur de coordonnées se trouve entre -1 et 1).

Voici un exemple de fichier contenant 3 points :

```
0.5 0.5 0.5
-0.5 -0.5 -0.5
0 0 -2
```

4.6.1 Sous-ensembles d'objets

Si le fichier ne contient qu'un sous-ensemble des objets, le mot clé "SUBSET" doit être ajouté, seul, en première ligne du fichier. Les lignes suivantes seront donc constituées de quatre valeurs, la première étant le rang global de l'objet. Comme pour les fichiers attributs / valeurs, SUBSET peut être suivi de "START WITH" pour spécifier un offset. Par exemple :

```
SUBSET START WITH 1
4 0.5 0.5 0.5
2 -0.5 -0.5 -0.5
1 0 0 -2
```

On fournit ici des coordonnées pour les objets 4, 2 et 1, avec un offset de 1, c'est à dire qu'il s'agit des objets de rang global respectif 3, 1 et 0.

4.7 Fichiers ne comportant que des attributs complémentaires

Il est possible de charger des fichiers ne contenant que des attributs complémentaires, sans aucune information de projection. On utilise alors le format des fichiers attributs / valeurs, en positionnant à 0 le nombre d'attributs descriptifs.

4.8 Noms d'attributs réservés (attributs complémentaires)

Il est possible d'utiliser des noms d'attributs spécifiques dans le fichier de données. Ces noms sont réservés pour désigner des types d'attributs particuliers, et ont un impact sur le fonctionnement du logiciel : ils lient directement ces attributs à des éléments visuels (tâche qui est habituellement faite à la main, comme décrit en section 6.2.1). Les deux noms réservés sont, à ce jour :

- **ImgFileName** : la colonne contient les noms de fichiers images. Le chemin d'accès aux images est toujours relatif au répertoire du fichier de données (par exemple, dans le fichier de données de la section 4.5, les fichiers images sont supposés être dans le même répertoire que le fichier de données). Un nommage absolu à partir de la racine est possible. Ces images peuvent ensuite être visualisées de manière contextuelle (au passage de la souris sur l'objet) ou permanente, à la demande. (exemple : "/data/images/enluminures/img0001.png").
- **Class** : classe associée à chaque objet. Dans **Explorer3D**, la classe est matérialisée par une couleur. La classe peut être représentée par une valeur textuelle ou numérique. La correspondance classe-couleur est définie automatiquement par le logiciel de visualisation (**Explorer3D** offrant toutefois la possibilité de modifier à la volée les correspondances).

4.9 Importation de données

Dans le cas de données sources dont le format n'est pas adapté à **Explorer3D**, un outil d'import est disponible. Il sera lancé automatiquement en cas de tentative de chargement de données non reconnues, et peut également être lancé par le menu "Fichiers / Outil d'importation de données". Cet outil permet de réorganiser les lignes et colonnes de votre fichier, et vous guidera pas à pas pour adapter vos données. Toutefois, il ne

fonctionne actuellement que pour des fichiers source au format texte (pas de fichier de données binaires), et se limite à l'import de fichiers "attributs / valeurs", c'est à dire de description d'objets (pas de matrice de distances ou de coordonnées 3D directes). Ceci devrait néanmoins couvrir la plupart de vos besoins.

Pour importer des données depuis un fichier de tableur (e.g. Excell ou OpenOffice Calc), passez par votre tableur et sauvez votre fichier au format CVS (qui est un format texte); ensuite, ouvrez-le dans l'outil d'import.

5 Présentation de l'interface de visualisation

5.1 Prise en main

Cette section illustre une utilisation simple d'Explorer3D dans le cadre d'une ACP. Au lancement d'Explorer3D s'ouvre la fenêtre de contrôle (cf. fig. 1).

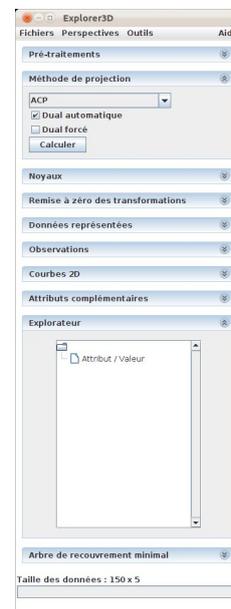
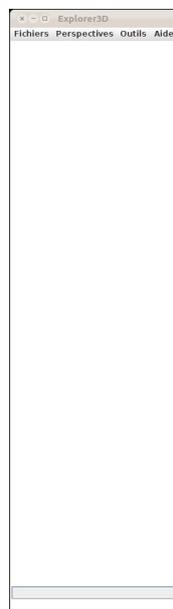


FIGURE 1 – Fenêtre de contrôle d'Explorer3D (au lancement) FIGURE 2 – Les commandes de l'ACP

Pour charger le fichier de données, allons dans le menu "Fichier / Charger un fichier local (nouveau projet)", et choisissons le fichier *iris.csv*. La fenêtre de contrôle se modifie et propose désormais les commandes liées à la projection de "n" dimensions vers 3 dimensions (cf. fig. 2).

Afin de réaliser une ACP classique, nous commençons par vérifier, dans la sous-fenêtre "pré-traitements", que "Variables centrées" est bien coché, et que "Variables réduites" est sélectionné dans la liste "Méthode de réduction", puis nous cliquons sur le bouton "Calculer" de la sous-fenêtre "Méthode de projection". La fenêtre de visualisation 3D s'ouvre et présente les objets sous leur forme par défaut, c'est à dire comme des sphères bleues sur fond noir (fig. 3).

Nous allons maintenant ajouter une information visuelle, en colorant les sphères en fonction de la classe (ou groupe, ou étiquette, ou modalité, ou catégorie, etc.) de l'objet

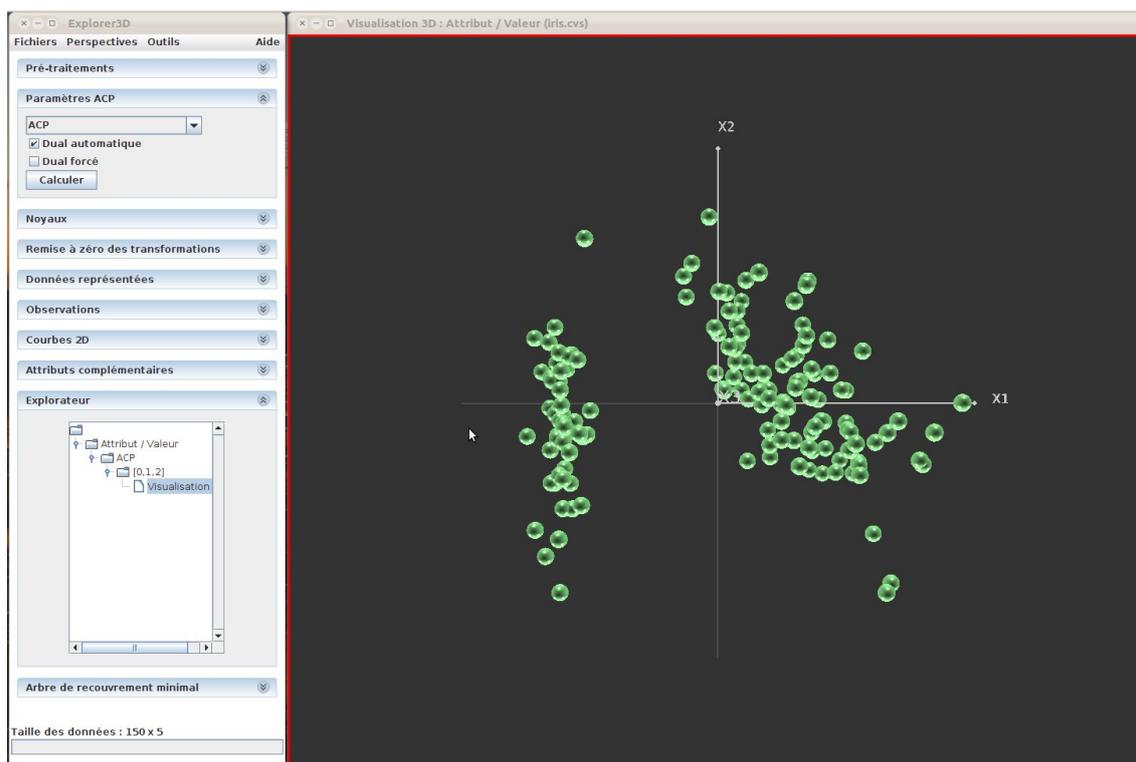


FIGURE 3 – Visualisation 3D par défaut

représenté (chaque sphère représente un iris, et le fichier contient des iris de trois variétés différentes). La classe de l’iris (c’est à dire sa variété) est indiquée par un attribut inclus dans le fichier de données, intitulé “classe”. Nous allons dérouler la sous fenêtre “Attributs complémentaires”, puis dans la liste “Classe”, nous allons choisir “classe” (il s’agit ici du seul attribut proposé). Les sphères sont alors colorées et nous voyons apparaître une légende (fig. 4).

Nous pouvons constater que la légende présente des cases à cocher. En décochant une case de la colonne “Classe”, nous dissimulons les objets de ce groupe. En cochant une case de la colonne “Ellipsoïdes”, nous faisons apparaître une ellipsoïde représentative du groupe correspondant (fig. 5). (Sommairement, cette ellipsoïde est centrée sur le centre de gravité du groupe, et ses diamètres correspondent à la variance du groupe suivant ses trois axes de plus grande variance). La figure 5 illustre une visualisation conjointe des objets et de leurs ellipsoïdes représentatives.

À noter la première ligne de cases à cocher (fond bleu) qui permettent de cocher/décocher l’ensemble des cases de la colonne.

Il est également possible de faire apparaître l’enveloppe convexe d’un groupe. Pour cela, il suffit de cocher les case de la colonne “Enveloppes convexes”. La figure 6 illustre une visualisation où les objets ont été dissimulés et où seules les enveloppes convexes apparaissent.

Pour terminer, nous pouvons manipuler la vue 3D en utilisant, à tour de rôle, les trois boutons de la souris (gauche : rotation ; centre : zoom ; droit : translation).

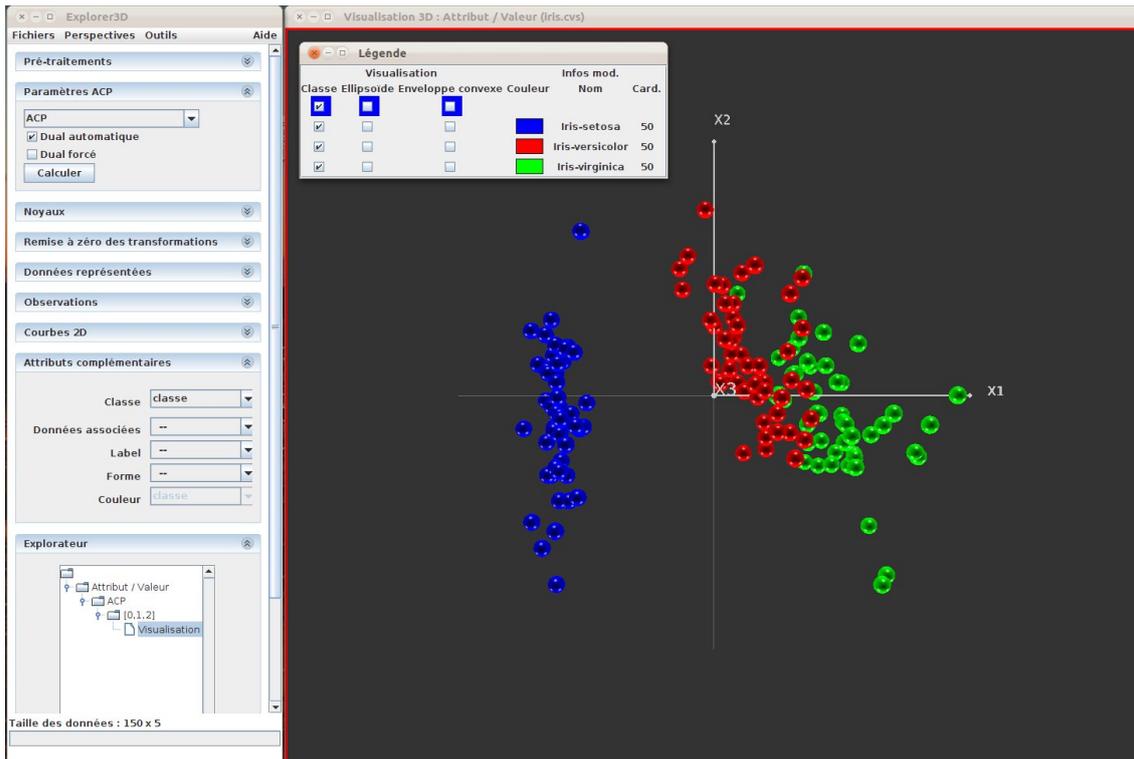


FIGURE 4 – Visualisation 3D avec couleur et légende

5.2 Menus de la fenêtre de contrôle

- **Fichiers** : donne accès aux fonctions de chargement de données
 - **Charger un fichier local** : permet de charger les différents types de fichiers de données supportés. Le type de fichier est automatiquement déterminé par le logiciel. Si aucun format standard n'est reconnu, un message d'erreur est affiché, et l'utilisateur est redirigé vers l'outil d'importation de fichiers, afin de le rendre compatible avec **Explorer3D**. Rappelons les trois types de fichiers reconnus :
 - Fichier attributs / valeurs : fichier contenant les descriptions d'objets. L'affichage se fera typiquement grâce à une technique de réduction de dimension (e.g. ACP).
 - Fichier matrice de distance : lorsque les données transmises par les prétraitements sont des distances entre objets. La projection se fera grâce à un MDS.
 - Fichier de points 3D : lorsque les données transmises sont directement des coordonnées dans l'espace 3D.
 - **Charger un fichier distant** : permet de charger un fichier stocké à distance, sur un serveur HTTP. L'outil permet d'indiquer l'url du site et de se déplacer dans une arborescence de répertoires. Les formats supportés sont les mêmes que pour un fichier local.
 - **Outil d'importation de données** : Cet outil permet de visualiser le contenu d'un fichier texte et de le rendre compatible avec Explorer3D.
 - *liste de fichiers* : après une première utilisation, les fichiers accédés sont mémorisés et peuvent être rechargés rapidement à partir du menu **Fichiers**.
 - **Quitter** : permet de fermer proprement l'application

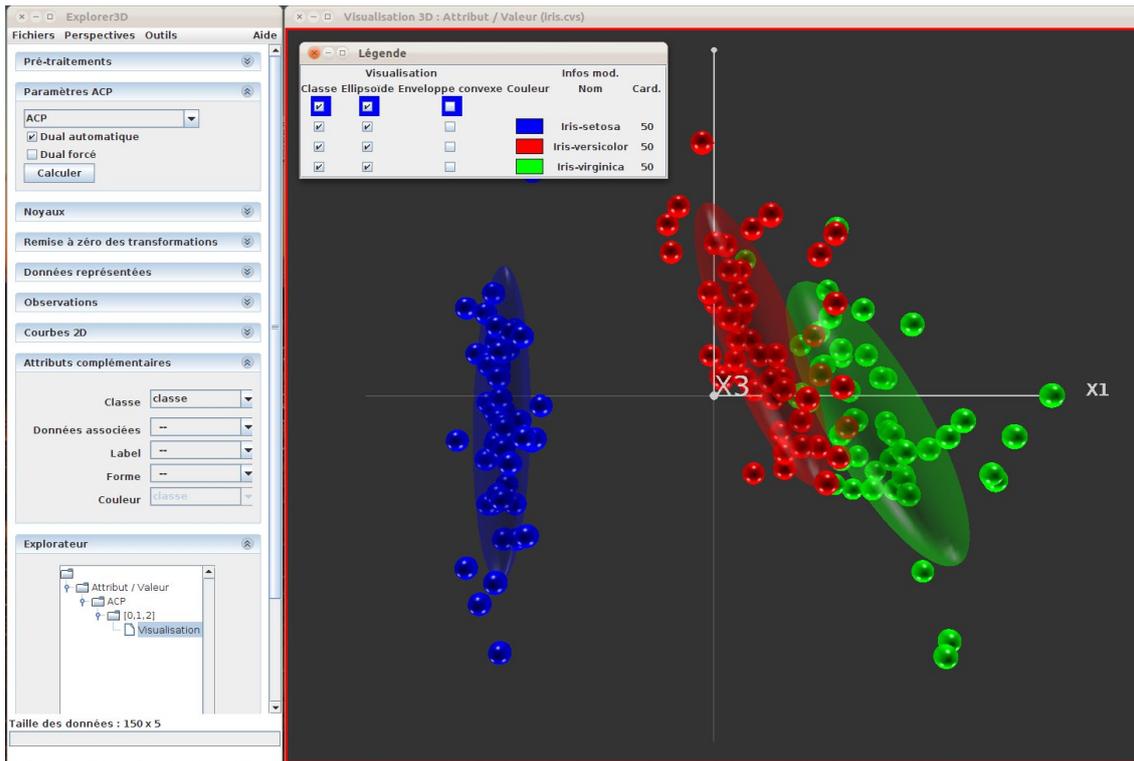


FIGURE 5 – Ellipsoïdes de groupes

- **Perspectives** : Ce menu regroupe les différents types de fenêtres de contrôle que l'on peut faire apparaître dans la fenêtre principale.
- **ND -> 3D** : permet d'accéder aux fenêtres de contrôle de la projection à partir d'un fichier attributs / valeurs (projection de "N" dimensions vers 3 dimensions).
- **Distance -> 3D** : permet d'accéder aux fenêtres de contrôle de la projection à partir d'un fichier de distance.
- **Aspect** : permet d'accéder à la fenêtre de réglage des paramètres de visualisation (e.g. couleur d'arrière plan, dilatation des objets,...).
- **Classification** : permet d'accéder à la fenêtre de lancement de techniques de classification automatique (e.g. kmeans, mélange de lois,...).
- **SVM** : permet d'accéder aux outils de calcul de SVM, c'est à dire d'hyperplans séparateurs entre objets de catégories différentes (voir section 7.6).
- **Outils** :
 - **Crop** : donne accès au Crop rectangulaire. Il s'agit d'un outil de sélection d'objets, permettant de générer une projection limitée aux objets sélectionnés. En choisissant cet outil, on ouvre la fenêtre de Crop décrite en section 6.6.
 - **Interaction** : donne accès à l'outil de modification de la projection. Cet outil permet d'indiquer des erreurs de placement spatial et de demander à l'ordinateur de calculer une nouvelle projection les corrigeant. Cet outil est décrit en section 7.4
 - **Voyage à travers les dimensions** : donne accès à un outil changement manuel de projection, dont le fonctionnement est décrit en section 8.1.
 - **Explorer à partir des images** : Cet outil permet de visualiser l'espace de projection, non pas dans son entier, mais en ne faisant apparaître que les objets associés à

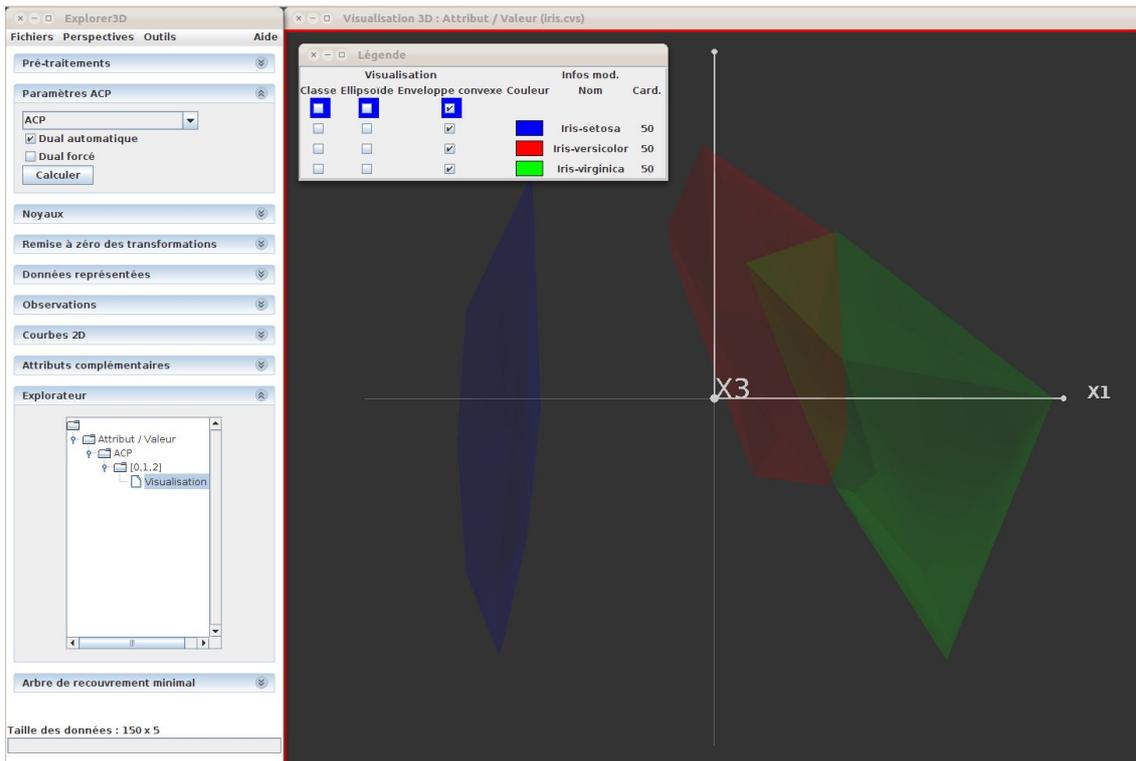


FIGURE 6 – Enveloppes convexes de groupes

des images sélectionnées par l'utilisateur (voir section 7.5

- **Options** : permet de régler divers options, liées notamment à l'affichage. Les options sont décrites en section 6.7.
- **Aide** Ce menu comporte deux sous-rubriques :
 - **Manuel en ligne** : donne accès à une version en ligne de ce manuel.
 - **A propos d'Explorer3D** : affiche quelques informations de version

6 Fonctions générales

6.1 Méthodes de projection

6.1.1 A partir d'un fichier attributs / valeurs

Méthode de projection Plusieurs techniques sont proposées dans la sous-fenêtre “Méthode de projection” accessible par le menu “Perspectives / ND -> 3D”. Le choix de la méthode se fait via une liste déroulante. Les techniques disponibles actuellement sont :

- **ACP** il s'agit de la méthode par défaut ; elle projette les objets de manière à ce que leurs distances dans l'espace de projection corresponde au mieux à leurs distances dans l'espace d'origine (on considère que l'ensemble des descripteurs fournis correspond aux coordonnées dans un espace de grande dimension). Lorsque l'ACP est sélectionnée, on dispose de deux options permettant de gérer l'exécution duale de l'ACP. Sans entrer dans les aspects techniques, la méthode duale fournit les mêmes résultats que l'ACP, mais est plus rapide lorsque le nombre de dimensions est plus grand que le nombre

d'objets. A noter que cette méthode duale interdit ensuite l'utilisation de certains outils, comme la correction de projection par interactions. On dispose de deux options : **dual automatique** détecte si le nombre de dimensions est supérieur au nombre d'objets, et utilise la méthode duale le cas échéant ; **dual forcé** force l'utilisation de la méthode duale.

- **Analyse discriminante** cette méthode et toutes les suivantes supposent que l'on dispose d'une étiquette pour chaque objet : on sait quels objets appartiennent à quel classe. En d'autres termes, cela suppose que l'on a spécifié un attribut complémentaire "classe" (voir section 6.2.1). La projection va donc se faire en cherchant à rapprocher les objets d'une même classe, tout en éloignant les différentes classes. Cela se fait en limitant toutefois les déformations par rapport à l'espace d'origine.
- **Analyse r-discriminante** similaire à la précédente, mais on cherche uniquement à éloigner les groupes.
- **Analyse rw-discriminante** similaire à la précédente, mais en plus la distance entre en compte : on cherchera surtout à éloigner des objets très proches mais n'étant pas de même groupe, en s'intéressant moins aux objets déjà éloignés. Le facteur sigma règle l'importance de l'éloignement. Plus sigma est élevé, moins on s'intéresse aux objets éloignés (avec $\sigma=1$, on fait une analyse r-discriminante).
- **ACP k plus proches voisins** variante expérimentale non documentée.
- **LLE** variante expérimentale non documentée

Pré-traitements A la sous-fenêtre "Méthode de projection" est associée une autre sous-fenêtre "Pré-traitements", permettant d'effectuer ce que les mathématiciens appellent un conditionnement des données. Cette fenêtre gère les traitements appliqués aux données sources avant la mise en oeuvre de la méthode de projection.

Une case à cocher permet de choisir le centrage des variables : les valeurs de chaque attributs sont décalées de manière à être centrées à zéro. Ce pré-traitement est notamment nécessaire à l'ACP.

Une liste déroulante permet de choisir le réglage de l'amplitude des valeurs : **Données brutes** ne modifie pas l'amplitude ; **Variabes réduites** règle l'amplitude de manière à ce que la variance de chaque attribut soit égale à 1 ; **Distances normalisées** règle l'amplitude des valeurs à 2, et donc avec l'option de centrage sélectionnée, conduit à des valeurs entre -1 et +1.

6.1.2 À partir d'une matrice de distance

La seule méthode par défaut est un MDS métrique, c'est à dire d'une projection linéaire qui place globalement les objets de manière à respecter au mieux les distances. En théorie on ne peut l'appliquer raisonnablement que si les données fournies sont bien des distances, au sens mathématique du terme. Il n'y a pas de conditionnement disponible. Récemment, nous avons ajouté des méthodes à noyaux. Celles-ci font l'objet d'une section spécifique (7.7). Sommairement, ces méthodes, accessibles à partir de la sous-fenêtre "noyaux", permettent de calculer les distances non euclidiennes. Par exemple, si l'on applique le noyau "isomap", on obtiendra une projection différente, ne tenant compte pour la projection que des distances entre objets proches (ce type de noyau est en général utilisé lorsque l'on sait que les objets, dans leur espace d'origine, sont répartis le long d'une surface non plane, ce que l'on appelle en mathématiques une "variété").

6.1.3 À partir d'une liste de coordonnées 3D

Dans le cas de coordonnées 3D, la projection reproduit à l'identique les coordonnées fournies. Les outils de centrage et dilatation dans la vue 3D restent disponibles (voir section 6.4).

6.2 Utilisation/affichage d'informations complémentaires

6.2.1 Présentation générale

Les attributs complémentaires sont les attributs présents dans les fichiers en entrée mais ne servant pas nécessairement au positionnement de l'objet (par exemple, la classe de l'objet, le nom du fichier image associé, etc.). Il existe 5 types d'attributs complémentaires dans **Explorer3D** :

- *classe* : indique la classe de l'objet. Par défaut, le même attribut est utilisé pour la coloration des objets. Il est possible d'utiliser de la classification floue, c'est à dire que chaque objet se rattache à plusieurs classes avec un certain degré d'appartenance. Ce cas particulier est décrit dans la section suivante.
- *données associées* : dans la version actuelle, il s'agit uniquement de fichiers images associés aux objets ;
- *label* : un texte descriptif de l'objet ;
- *forme* : une forme différente est proposée pour chaque valeur différente de cet attribut. Actuellement quatre formes sont disponibles ; il faut donc éviter de l'utiliser si l'attribut peut prendre plus de 4 valeurs ;
- *couleur* : une couleur différente est proposée pour chaque valeur différente de cet attribut. Il n'y a pas de limite au nombre de couleurs (ou plus exactement cette limite est plus élevée que le nombre de teintes perceptibles par l'utilisateur).

6.2.2 Objets multi-classes

Afin d'identifier des groupes d'objets, on peut soit faire une affectation en dur (chaque objet appartient à un seul groupe), soit une affectation souple, dite "floue", ou chaque objet appartient à un certain degré à chaque groupe. Il est possible de visualiser ce genre d'affectation. Pour cela, dans la liste déroulante "Classe", choisir "–Multiclasses–". Cela fait apparaître une sous-liste dans laquelle on peut choisir plusieurs attributs. L'affichage qui en résulte est similaire à celui des nuées dynamiques floues (voir section 7.3.3).

Les attributs utilisés pour le multi-classe doivent être de type numérique (réel), chacun indiquant le degré d'appartenance de l'objet à une classe donnée ; la somme des degrés d'appartenance de chaque objet doit être égale à 1.

6.2.3 Affichage des images

L'affichage des images est commandée par le menu contextuel ouvert lors du clic droit de la souris (voir section 6.3.3).

Si, dans la fenêtre des options (section 6.7) la case "Afficher les images hors 3D" est cochée, les images sont affichées dans des fenêtres séparées (figure 7) ; dans le cas contraire, elles s'affichent dans la scène 3D (figure 8).

Dans la figure 7, on peut observer que les objets correspondant aux images ont été sélectionnés (fond rouge des images, et coloration claire des objets). On peut aussi voir que

l'on fait apparaître une des images dans la vue 3D, de manière contextuelle, au passage de la souris. Cette image disparaîtra quand le pointeur de la souris ne sera plus sur l'objet. Par défaut, les images apparaissent hors de la fenêtre de visualisation 3D.

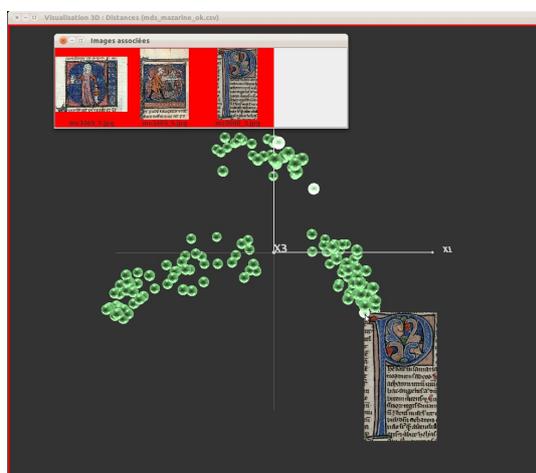


FIGURE 7 – Affichage d'images hors-3D + une image contextuelle (passage de la souris sur l'objet).

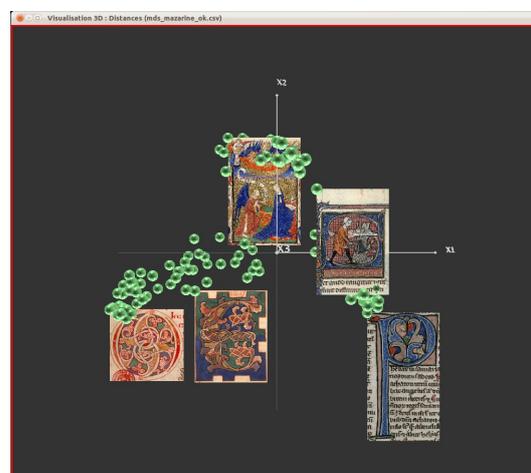


FIGURE 8 – Affichage d'images dans la scène 3D. Les images sont rattachée à l'objet se trouvant sur leur coin supérieur gauche.

6.2.4 Sources d'images multiples

Dans le cas où l'on dispose de plusieurs images par objet, il est possible de les visualiser simultanément dans des vues différentes. Plus précisément, il est possible de définir un attribut d'image spécifique à chaque vue 3D.

Par défaut, un seul attribut est utilisé globalement pour toutes les vues. Si plusieurs attributs complémentaires contiennent des noms de fichier image, et que l'on souhaite utiliser plusieurs de ces attributs, on peut en associer un différent à chaque vue 3D. Pour cela, il faut tout d'abord rendre active la vue 3D à laquelle on souhaite associer un attribut différent, puis cocher la case "D.A. spécifiques" dans la sous-fenêtre "Attributs complémentaires", et enfin choisir l'attribut souhaité. Toutes vues pour laquelle la case "D.A. spécifiques" n'a pas été cochée partagent le même attribut d'image. A l'inverse, plusieurs fenêtre peuvent avoir un attribut d'image spécifique.

Cette option peut par exemple être utile en cas de données multisources apportant chacune leur image, ou dans le cas plus général où l'image n'est qu'une information complémentaire de l'objet étudié, et non pas l'objet dont sont extraits les descripteurs.

6.2.5 Gestion des couleurs

Par défaut, les couleurs sont associées aux classes, c'est à dire que la désignation d'un attribut de classe conduit à l'utilisation du même attribut pour les couleurs.

Néanmoins, il est possible de choisir l'attribut de couleur directement, et donc de le dissocier de l'attribut de classes. Cela ouvre notamment l'accès à une coloration par dégradé de couleurs (voir fig. 9). Celle-ci s'active automatiquement si l'attribut choisi est un attribut numérique. **attention** : *Explorer3D* considère par défaut que les attributs

complémentaires sont de type symbolique. Pour qu'un attribut puisse déboucher sur une coloration par dégradé, il faut qu'il soit connu explicitement comme attribut numérique, et donc que dans le fichier source, son nom comporte le suffixe ".N".

Les couleurs peuvent également être remplacées par des textures en noir et blanc. L'utilisation des textures est déclenchée en cochant la case "Texture" dans la sous-fenêtre "Attributs complémentaires". Une dizaine de textures sont actuellement disponibles. L'utilisateur peut changer la texture associée à un groupe en cliquant sur cette texture dans la légende. Les textures sont reproduites sur les ellipsoïdes et les enveloppes convexes, si celles-ci sont tracées.

En revanche, les couleurs ont été conservées, dans la version actuelle, pour distinguer les objets sélectionnés (un objet sélectionné apparaîtra coloré, même si on utilise les textures ; il redeviendra texturé s'il n'est plus sélectionné).

La figure 10 illustre l'utilisation de textures. On peut voir que l'utilisateur est en train de redéfinir la texture associée à un groupe (en l'occurrence : iris-virginica).

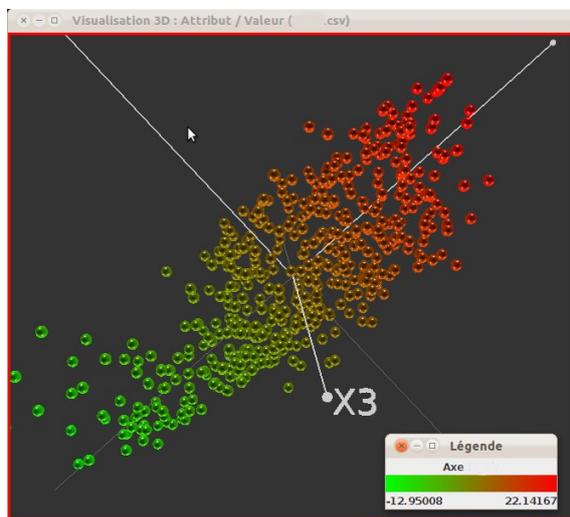


FIGURE 9 – Coloration par dégradé

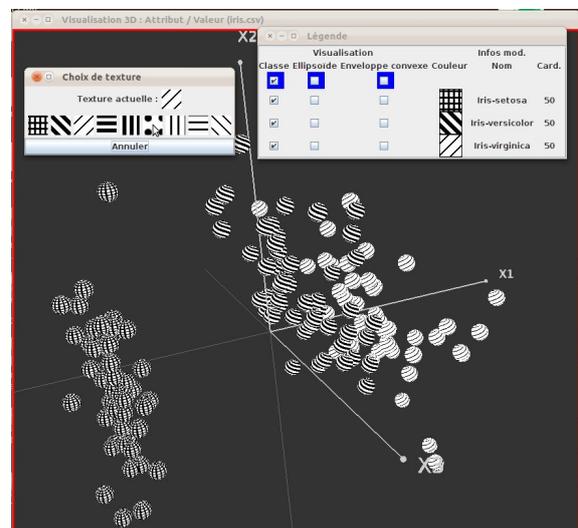


FIGURE 10 – Utilisation de textures

6.3 Utilisation de la souris

Il est conseillé d'utiliser une souris à 3 boutons. De nombreuses manipulations sont accessibles par le biais de la souris : sélection d'objets, affichage d'informations, etc.

6.3.1 Manipulation de la scène 3D

En cliquant sur le fond 3D (c'est à dire sans que le pointeur de souris se trouve sur un objet particulier), on obtient les comportements suivants :

- *Bouton gauche* : rotation
- *Bouton central* : zoom
- *Bouton de droite* : translation

6.3.2 Action sur les objets (raccourcis)

Quelques actions sont accessibles en un simple clic :

- *Bouton gauche + shift* : sélectionne / désélectionne l’objet situé sous le pointeur (un objet sélectionné apparaît en surbrillance). Les objets sélectionnés sont par exemple utiles pour les crops, mais également pour retrouver l’image correspondante, etc.
- *Bouton droit* : donne accès au menu contextuel présenté dans la section suivante.

6.3.3 Menu contextuel

Au contraire des versions précédentes, le déclenchement d’actions à partir de la souris a été revu et simplifié. Le clic droit ouvre un menu contextuel, permettant de choisir diverses actions. Lorsque l’on fait un clic droit alors que la souris se trouve au dessus d’un objet, les actions proposées sont :

- afficher / cacher l’image associée à l’objet (a un effet uniquement si un attribut correspondant a été spécifié) ;
- afficher / cacher toutes les images (même remarque) ;
- afficher / cacher le label associé à l’objet. Si un attribut label a été spécifié, on affiche la valeur de cet attribut. Sinon, on affiche le numéro de l’objet ;
- afficher / cacher tous les labels ;
- afficher / cacher les labels des objets sélectionnés ;
- effectuer un crop circulaire (ce crop est centré sur l’objet sous la souris). Voir section 6.6 ;
- effectuer une sélection multiple. Cette action est décrite dans la section 6.4.1 ;
- centrer la vue sur l’objet se trouvant sous la souris ;
- activer la loupe (voir section 6.5).

Si le menu contextuel est ouvert alors que la souris ne se trouve pas sur un objet, la liste d’actions proposées est réduite aux actions globales (afficher / cacher tout, sélection multiple).

6.4 Sélection d’objets

Un mécanisme de sélection uniformisé a été mis en place. Il existe plusieurs façons de sélectionner un objet : à partir de la vue 3D, à partir de la présentation tabulaire, à partir de la liste d’images affichées...

Dorénavant, la sélection dans une des fenêtres entraîne la sélection dans les autres fenêtres. Par exemple, lorsqu’un objet est sélectionné dans la zone 3D, la ligne correspondante apparaît sur fond gris dans le tableau, et l’image associée au point apparaît sur fond rouge. A l’inverse, en cliquant sur une ligne du tableau, on sélectionne (ou on désélectionne l’objet correspondant dans la scène 3D (et dans liste d’images).

6.4.1 Sélection multiple

Il est possible de sélectionner plusieurs objets en une seule opération. En effectuant un clic gauche, alors que la souris ne se trouve sur aucun objet, on génère un rectangle translucide (on génère son coin supérieure gauche). Par la suite, le déplacement de la souris permet de matérialiser le rectangle (la position de la souris correspond au coin inférieur droit, voir fig. 11). Un second clic sur la souris fixe le rectangle, et sélectionne tous les objets situés à l’intérieur du rectangle (voir fig. 12).

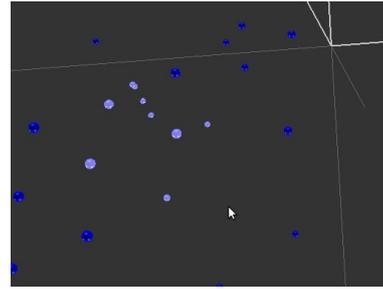
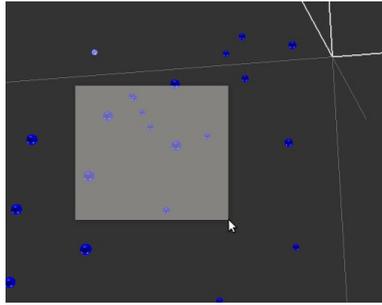


FIGURE 11 – Sélection multiple (rectangle) FIGURE 12 – Sélection multiple (effectuée)

6.5 Loupe

Cet outil, activable par le menu contextuel, permet de voir plus en détail les objets projetés. Cependant, son intérêt premier vient du fait que les objets sous la loupe se retrouvent sélectionnés. On peut alors suivre cette liste d'objets sélectionnés dans toutes les vues, s'il y en a plusieurs. Les paramètres de la loupe (diamètre, coefficient) sont réglable à partir de la sous-fenêtre "Outil loupe", accessible par le menu "Perspectives / Aspect". La figure 13 illustre l'utilisation de la loupe.

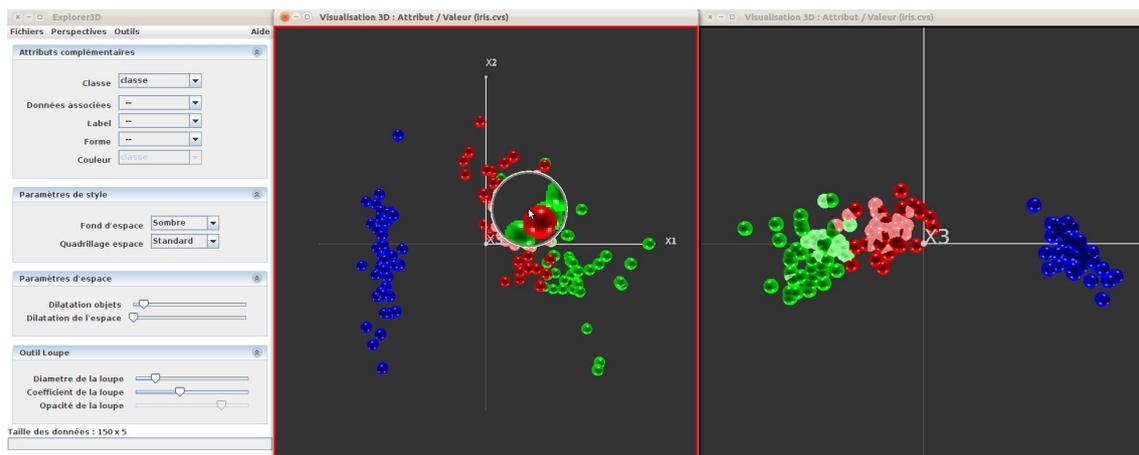


FIGURE 13 – Loupe. Dans la vue de droite, les objets sélectionnés sont ceux qui se trouvent sous la loupe dans la vue centrale. A noter la fenêtre de paramétrage de la loupe dans la fenêtre principale de commande, à gauche.

6.6 Zoom sur un sous-ensemble d'objets (Crop)

Le mécanisme de crop consiste à sélectionner un sous ensemble d'objets et à recalculer la projection spatiale pour ce seul groupe d'objets. La projection obtenue peut être assez différente de la projection pour l'ensemble des objets.

Il existe deux manières de déclencher un crop :

- **Crop rectangulaire** la zone d'espace sélectionnée est un pavé. Ce type de crop est lancé à partir du menu "Outils / Crop". Lorsque l'on clique sur "Nouveau Crop", un pavé est créé au centre de la projection 3D. On peut déplacer / redimensionner ce pavé

grâce aux curseurs glissants de la fenêtre (figure 14). Si des objets sont sélectionnés avant que l'on clique sur "Nouveau Crop", alors le pavé créé englobe ces objets.

Lorsque le pavé est correctement placé, on clique sur "Crop !" pour calculer la nouvelle vue 3D. On peut enchaîner plusieurs zoom successifs.

- **Crop circulaire** il s'agit à proprement parler d'un crop sphérique, la zone de sélection étant centrée sur un point (figure 15). On active ce type de crop par le biais du menu contextuel.

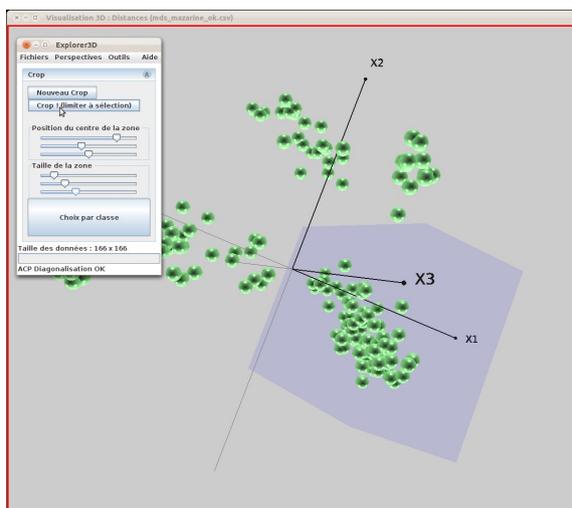


FIGURE 14 – Crop rectangulaire

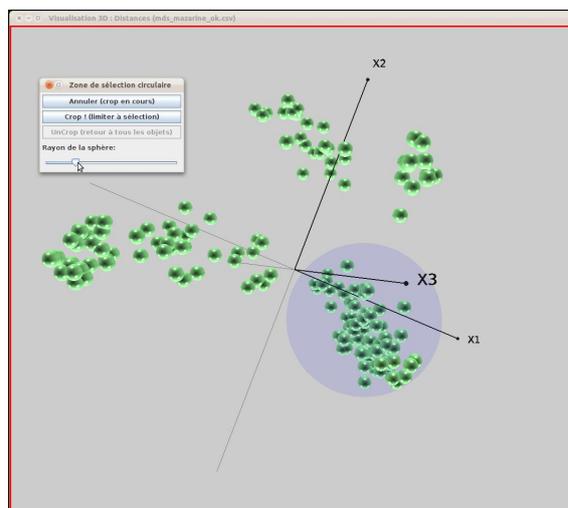


FIGURE 15 – Crop sphérique

6.7 Options

Accessible par "Outils/options". Il offre les options suivantes :

- **couleur par défaut** des objets : Le rectangle coloré est cliquable, et le clic ouvre un dialogue de sélection de la couleur par défaut ;
- **Valorisation de la 3ème dimension au démarrage** : permet d'ouvrir la scène 3D avec une légère rotation (mise en valeur des trois axes) ;
- **Afficher les axes originaux avant projection** : Les axes d'origine sont représentés dans le sous-espace de projection. Cette fonction n'est valable que pour la réduction de dimension. Elle montre comment les axes d'origine se projettent dans le sous-espace de projection. Elle ne constitue pas vraiment une option, et devrait prochainement être placée hors du menu des options.
- **Afficher les images hors 3D** : permet d'afficher les images liées aux objets dans la scène 3D ou en dehors de celle-ci (dans une fenêtre indépendante, ce qui est l'option par défaut) ;
- **Calcul de la vue 3D au chargement** : permet de calculer directement la scène 3D lors du chargement d'un fichier. Dans le cas d'une réduction de dimensions, c'est l'ACP qui est calculée ;
- **Activer du système d'envoi distant** : cette fonctionnalité concerne surtout la visualisation sur un mur d'images. Elle ne devrait donc pas être activée pour une utilisation courante. Il s'agit d'une fonctionnalité ancienne et non maintenue dont le bon fonctionnement n'est pas garanti dans la version courante.

- **Affichage des objets sources (tableau)** : permet d'afficher, dans un tableau, les informations complémentaires des objets visualisés (cf section 7.2).
- **Affichage dynamique des images** : permet de faire apparaître dynamiquement, dans la vue 3D, l'image associée à un objet lorsque le pointeur de la souris s'arrête sur celui-ci. Bien entendu, cela ne fonctionne que si l'attribut indiquant les fichiers images a été défini (cf. section 6.2.1).
- **Vision stéréoscopique** : Sur une machine dotée d'une carte (et d'un écran) adapté(s), active la vision stéréoscopique. Sans effet sur une machine sans carte adaptée.
- **Taille des labels** : réglage de la taille des informations textuelles de la vue 3D.
- **Taille des objets dynamique** : Si cette option est sélectionnée, le logiciel choisi lui-même la taille des objets en fonction du nombre d'objets à afficher, et l'option "Taille des objets par défaut" est désactivée.
- **Taille des objets par défaut** : si l'option "Taille des objets dynamique" n'est pas sélectionnée, l'utilisateur peut régler ici la taille des objets.
- **Mode multisource** : Autorise l'utilisateur à manipuler des données multisources (see 4.2).
- **Commande via le réseau** : si cette option est activée, **Explorer3D** se met en écoute sur un port (par défaut, le port 50000) pour recevoir des commandes. Le protocole utilisé est présenté en section 7.9.
- **N° de port** : réglage du port d'écoute. Non accessible dans cette version.
- **Légende activée** : Lorsque cette option est désactivée, les légendes ne s'affichent pas. Pour plus d'informations sur les différents types de légende et leurs fonctionnalités, voir section 6.8.

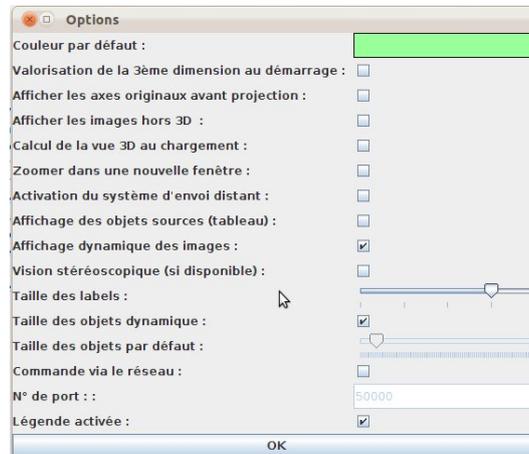


FIGURE 16 – Options

6.8 Légende

La légende se décline en plusieurs aspects suivant la nature de l'information de coloration. La légende standard a déjà été présentée en section 5.1. Elle correspond au cas simple d'une coloration reposant sur la classe des objets sur la base d'une couleur par classe. Il existe actuellement deux autres types de légendes : la légende multiclassés, et la légende sur attribut continu.

6.8.1 Légende multiclassés

La légende multiclassés est activée dans le cas d'objets multi-classes. Cela se produit lorsque l'on indique à la main une liste d'attributs de classe (voir section 6.2.2) ou lorsque l'on utilise un outil de classification floue (voir section 7.3.3). La coloration des objets se fait en fonction de la classe à laquelle ils appartiennent "le plus". Les actions complémentaires de cette légende ne sont pas fonctionnelles actuellement, du fait d'évolutions en cours du logiciel ; néanmoins, en mode normal, on peut faire apparaître le centroïde de chaque classe (sous forme d'une croix de la même couleur que la classe). On peut également observer la diffusion de la classe parmi tous les objets grâce aux enveloppes convexes : en sélectionnant "activer les enveloppes convexes" et en jouant sur le slider en bas de la fenêtre de légende, on voit les enveloppes se déplacer à mesure pour englober tous les objets dont le degré d'appartenance à une classe est supérieure à la valeur courante du slider. La colonne "ellipsoïde" n'est pas utilisée actuellement.

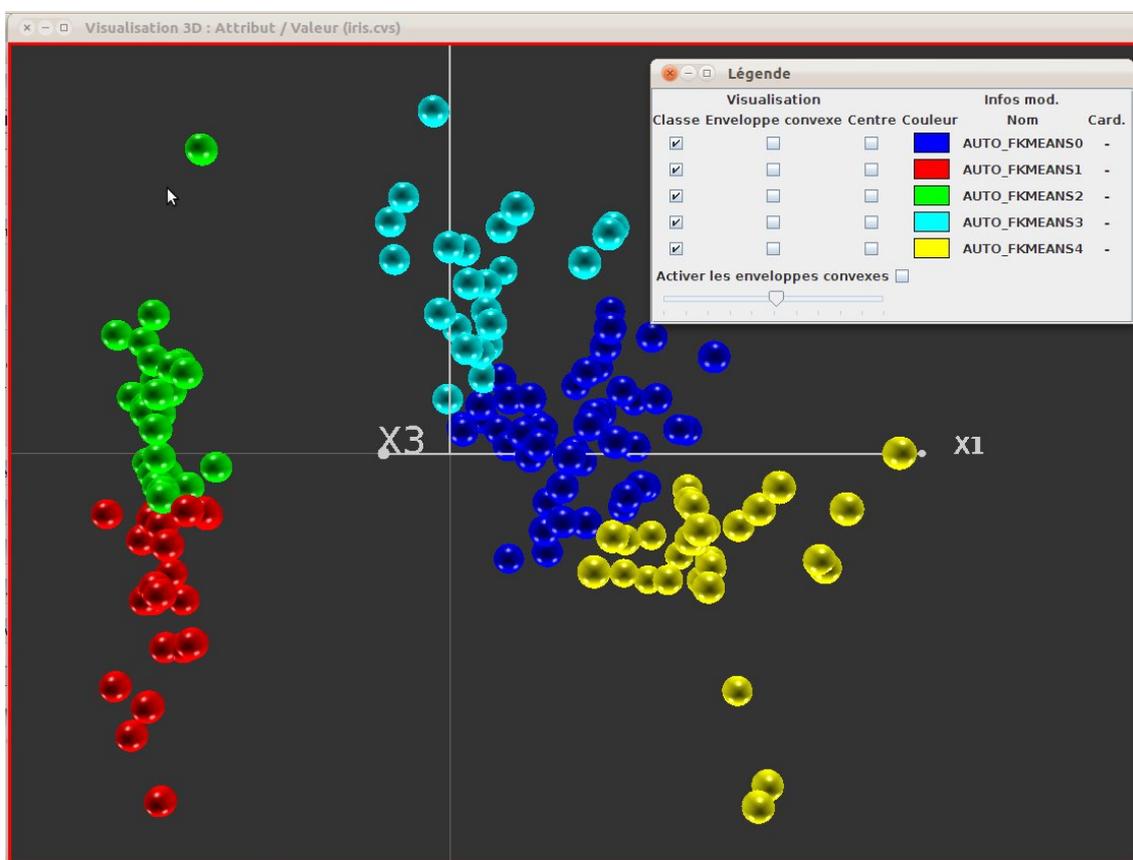


FIGURE 17 – Légende multiclassés. Ici on observe le résultat d'un kmeans flou. Chaque objet est coloré en fonction de sa classe principale.

6.8.2 Légende sur attribut continu

Il est possible de définir une coloration qui n'est pas liée à une classe, mais uniquement à un attribut numérique à valeurs réelles. Un dégradé de couleurs est alors calculé avec comme extrêmes les valeurs min et max de l'attribut choisi (voir figure 18). Chaque objet

est coloré en fonction de la valeur qui lui est associé. Cette coloration peut être déclenchée à partir de trois actions : la sélection d'un attribut complémentaire explicitement numérique ; la visualisation d'un SVM par coloration des objets (en fonction de leur distance à l'hyperplan séparateur, voir section 7.6) ; et la matérialisation par coloration des objets des valeurs d'un axe de projection (voir section 6.9). Cette légende comporte tout simplement le dégradé avec les valeurs min et max correspondant aux extrêmes.



FIGURE 18 – Légende sur attribut continu. On peut remarquer ici que la coloration correspond à la position sur l'axe x1.

6.8.3 Légende à base de textures

Comme présenté en section 6.2.5, les couleurs peuvent être remplacées par des textures. La légende permet de modifier la texture associée à un groupe.

6.9 Axes de projection

La sous-fenêtre “données représentées” permet différentes actions sur la projection des objets dans les vues 3D. Les deux premières option proposées sont :

- **Zoom dimensions peu discriminantes** : suivant la méthode de projection, il est fréquent que les objets soient bien dispersés suivant l'axe x1, moins suivant l'axe x2, et encore moins suivant l'axe x3 (écrasement autour de $x3=0$). Afin de mieux répartir les objets dans l'espace, on peut cocher cette option qui va redilater les axes les moins exploités (voir figure 19).

- **Centrage nuage / image** : cette option permet de déplacer les points afin que l’origine des axes ne corresponde plus au centre de gravité (vue standard) mais à la valeur médiane suivant chaque axe.

Ces deux options s’appliquent à la vue active et à toutes les vues créées par la suite (jusqu’à désélection de ces options).

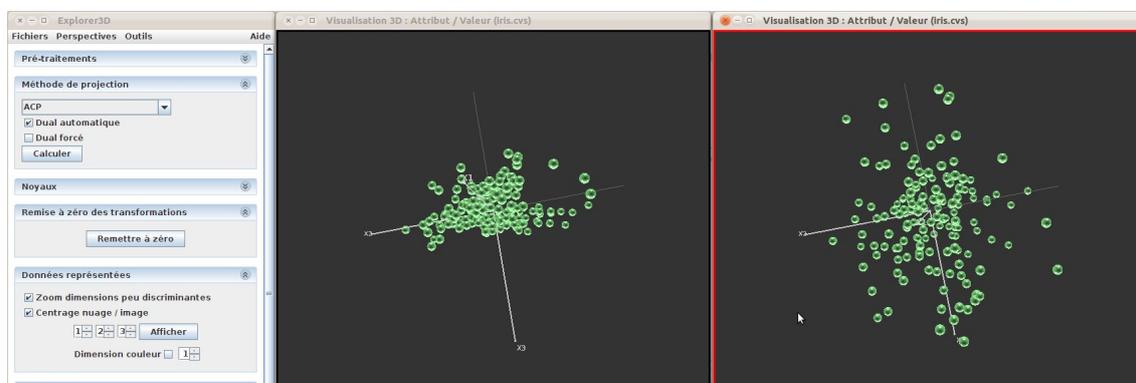


FIGURE 19 – Zoom dimensions peu discriminantes et centrage nuage / image. La vue de gauche correspond à la vue standard d’une ACP, celle de droite à l’application des deux options sur la même ACP. Les vues ont été tournées afin de voir les axes x_2 et x_3 . On peut voir que les objets sont plus écartés à droite sur l’axe x_3 (zoom dimension peu discriminante) et dans une moindre mesure suivant l’axe x_2 , et que le nuage est décalé vers la gauche (centrage)

Cette fenêtre permet également de choisir les axes de projection utilisés. En général ce choix n’a de sens que lorsque le nombre d’axes disponibles est supérieur à 3, c’est à dire pour une projection de type “ $nD \rightarrow 3D$ ”. Toutefois, il permet également de modifier l’ordre des axes, ou d’utiliser plusieurs fois les même axe. L’outil permet de choisir les trois axes dans l’ordre x_1 (largeur), x_2 (hauteur) et x_3 (profondeur). La projection n’est recalculée qu’après avoir cliquer sur le bouton “Afficher”. La projection s’applique à la vue courante. Elle ne génère pas de nouvelle vue 3D, mais un déplacement des objets dans la vue courante.

Les axes disponibles sont numérotés de 1 à n , n étant ici le nombre d’axes fournis par la méthode de projection (il s’agit bien des axes calculés par cette méthode, et non des dimensions d’origine fournies par le jeu de données). Les sélecteurs d’axe fournissent une valeur supplémentaire, “X”, qui indique qu’il n’y aura pas de projection suivant l’axe correspondant. Par exemple, si on indique “X” pour le troisième axe, alors les objets seront tous dans le plan x_1 - x_2 .

Enfin, cette fenêtre permet de sélectionner une dimensions comme support à la coloration des objets. Il s’agit d’une des dimensions calculées par la méthode de projection. En cochant la case “Dimension couleur”, on active la coloration par valeur continue, en utilisant comme valeur, pour chaque objet, la valeur de projection correspondant à l’axe indiqué dans le sélecteur. Il s’agit ici encore des axes calculés par la méthode de projection, et non des dimensions d’origine fournies par le jeu de données (cette disposition est susceptible d’évoluer, et pourrait être étendue aux attributs sources dans les versions ultérieures). La figure 18 illustre la coloration en fonction d’un axe de projection. Dans cet exemple on a utilisé un des axes réellement utilisé pour la projection (celui de x_1), mais on aurait

pu utilisé un autre axe, non utilisé pour la projection (par exemple le quatrième ou le cinquième axe fourni par la méthode de projection).

6.10 Observations

Cette sous-fenêtre visualise la part de variance portée par les axes issus de la méthode de projection (figure 20). Un histogramme donne la part de variance par axe ; les axes utilisés pour la projection sont en rouge, les autres en bleu. Un camembert donne la part de variance restituée dans la vue 3D active.

Si un attribut de classe est sélectionné, une indication de la pureté des groupes est fournie (indice de Dunn).

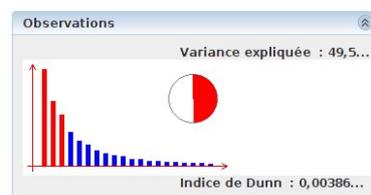


FIGURE 20 – Sous-fenêtre “Observations”. Nous pouvons voir ici que la projection a donné lieu à de nombreuses dimensions. Les trois dimensions les plus significatives sont utilisées pour la projection (en rouge dans l’histogramme), ce qui correspond à 49,5 % de la variance (camembert). Un attribut de classe a été sélectionné (l’indice de Dunn en indique la qualité).

6.11 Sauvegarde de projet

L'utilisateur peut sauvegarder le projet en cours, afin de le recharger plus tard. Cette sauvegarde, dans sa version actuelle, comporte les sources de données et les vues 3D issues directement de ces sources ou de sous-ensembles d'objets. Sont également sauvegardés les attributs complémentaires actifs (classe, couleur, etc.).

Pour sauvegarder le projet courant, aller dans le menu “Fichiers” et choisir “sauver le projet”. Pour charger un projet, choisir “charger un projet”.

Les projets sont par défaut sauvegardés avec l’extension “.e3d”. Techniquement parlant ce sont des fichiers texte organisés suivant le format JSON.

7 Outils principaux

7.1 Manipulation multi-vues

Explorer3D a été profondément modifié afin de permettre une manipulation multi-vues interactive. Ce mécanisme se décline en plusieurs aspects : d’une part le fait de pouvoir ouvrir simultanément plusieurs fenêtres de visualisation, et d’autre part le fait de pouvoir suivre les objets sélectionnés entre ces différentes vues.

7.1.1 Système multi-vues

Par vue nous désignons ici la fenêtre 3D. On peut ouvrir plusieurs fenêtre 3D pour un même jeu de données, notamment pour les raisons suivantes :

- voir simultanément la projection sous différents angles ;
- effectuer plusieurs projections différentes (ACP, ALD, etc.) ;
- voir simultanément une vue d'ensemble et un ou plusieurs zoom locaux (crops).

7.1.2 Effet des commandes sur les vues 3D

Certaines commandes sont globales et vont affecter toutes les vues : il s'agit par exemple de la coloration de sujets (notamment pour identifier les classes). En revanche, beaucoup de commandes n'ont pas vocation à s'appliquer globalement, telles les opérations de clustering, les crops, etc., qui s'appliquent sur une vue donnée.

Ceci conduit à définir la notion de vue active : la vue active est la vue sur laquelle s'appliquent les commandes non globales. Comment désigner et identifier cette vue ?

On peut choisir la vue à activer en donnant le focus à sa fenêtre. Une fois la vue active, sa fenêtre est bordée de rouge (voir figure 21). Une vue reste active tant que l'on ne désigne pas explicitement ou implicitement une nouvelle vue active. Lorsque l'on crée une nouvelle vue, celle-ci devient automatiquement active. A tout instant, il y a au plus une vue active.

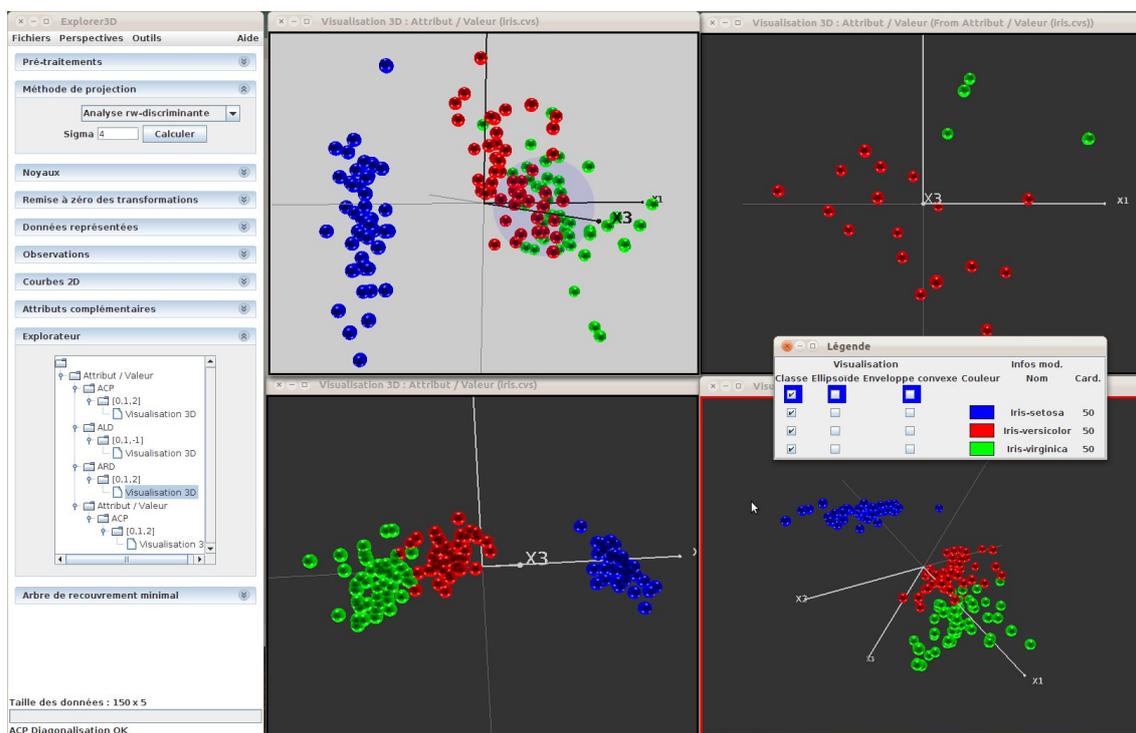


FIGURE 21 – Plusieurs vues simultanées. La vue active (en bas, à droite) est repérée par sous liseré rouge. On peut remarquer ici l'utilisation de commandes globales (la coloration en fonction des classes) et locales (la mise sur fond clair de la vue en haut à gauche)

7.1.3 Explorateur de vues

La liste des vues disponibles peut être consultée dans la sous-fenêtre “Explorateur” de la fenêtre principale. Cette sous-fenêtre contient une arborescence, avec à la racine la source de données utilisée, puis la méthode appliquée sur cette source, enfin la projection effectuée (indiquant les axes de projection choisis), et enfin la vue calculée sur cette base. La vue active est mise en avant dans cette arborescence. Le fait de cliquer sur l’une des vues dans l’arborescence rend celle-ci active. La figure 22 présente l’explorateur de vues. On peut constater ici que l’on manipule des données décrites sous-forme “attribut / valeur”. On peut constater que pour l’ACP on a projeté suivant les trois axes principaux. On pourra observer ici les axes de projection de la seconde méthode. Le troisième axe, ici, vaut “-1”, ce qui indique que l’on n’a pas de projection suivant le troisième axe : tous les objets sont projetés dans le plan x-y.

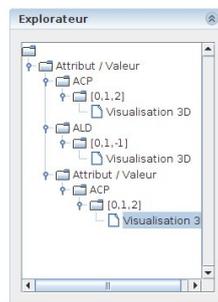


FIGURE 22 – L’explorateur des vues. Ici, nous avons une source de données de type “Attribut / Valeur”, sur laquelle on a appliqué une projection par ACP et une par ALD ; un zoom local a été appliqué sur une des vues (on voit qu’il y a une source “Attribut / Valeur” dans les descendants de la source principale, signe d’une sélection d’un sous-ensemble d’objets. C’est la vue correspondant au zoom local qui est active (fond gris)

7.1.4 Vues et objets sélectionnés

Les objets sélectionnés apparaissent en surbrillance. La liste des objets sélectionnés est transverse à toutes les vues. En d’autres termes, on sélectionne dans une vue et on voit la surbrillance dans toutes les vues. Cela peut être assez utile pour rechercher un même objet dans les différentes vues, ou pour observer la dispersion d’un groupe d’objets à travers des projections diverses.

7.2 Affichage tabulaire de la description des objets

Il est possible, par le biais des options, d’activer l’affichage d’un tableau contenant les informations complémentaires des objets manipulés. Comme défini précédemment, les attributs “complémentaires” sont les informations fournies mais non utilisées dans le calcul de la projection spatiale. Il peut s’agir par exemple de la classe si une classe est fournie, d’un nom de fichier image associé, etc. Plusieurs attributs peuvent être fournis.

Ce tableau contient par défaut tous les objets de la scène 3D, triés suivant leur ordre dans le fichier de données. Par défaut, tous les attributs complémentaires apparaissent.

On peut choisir les attributs affichés grâce au menu **Visibilité colonnes / attributs complémentaires**. Il est également possible de faire apparaître les coordonnées de projection des objets, grâce au menu **Visibilité colonnes / coordonnées de projection**. Par défaut, ces colonnes ne sont pas affichées.

Il est prévu dans une version ultérieure de pouvoir afficher les attributs ayant servi à calculer la projection. Toutefois, le nombre d'attributs étant potentiellement large, la mise en oeuvre technique nécessite réflexion.

Il est possible de sélectionner les objets dans cette liste, et de réduire les lignes visibles aux seuls objets sélectionnés (voir fig. 23). Pour cela, cocher "Filtrer (sélection)" dans le menu "Visibilité lignes".

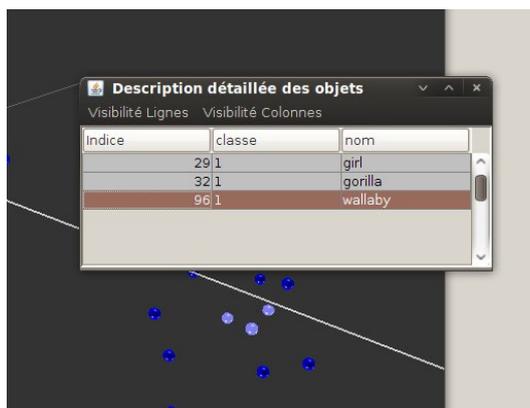


FIGURE 23 – Visualisation d’attributs complémentaires (avec filtre)

Le mécanisme de sélection est décrit en section 6.4 .

7.3 Clustering

Plusieurs méthodes de classification sont désormais disponibles. Elles sont accessibles à partir de la perspective “classification” (voir fig.24) qui peut être ouverte à partir des menus. Les méthodes de classification proposées ont un défauts commun : elles sont plus ou moins instables, et d’une exécution à l’autre, les groupes générés peuvent changer. Il est donc conseillé de les relancer plusieurs fois (bouton ”recalculer”) afin de se faire une meilleure idée de la stabilité des groupes (notons que le bouton “recalculer” permet aussi, pour certaines méthodes, de relancer la classification automatique avec de nouveaux paramètres, comme un nombre de groupes différents).

7.3.1 Mélange de lois

La fonction de clustering par mélange de lois considère que la répartition spatiale des objets se fait selon un ensemble de lois (multi) normales. On cherche donc à calculer les lois sous-jacentes, et on les visualise sous formes d’ellipsoïdes centrées sur les centres des lois et étirées suivant les directions et écart-type de celles-ci (cf figure 25). La méthode implantée calcule automatiquement le nombre de lois.



à l'ouverture

après choix d'une méthode

FIGURE 24 – Fenêtre “classification”

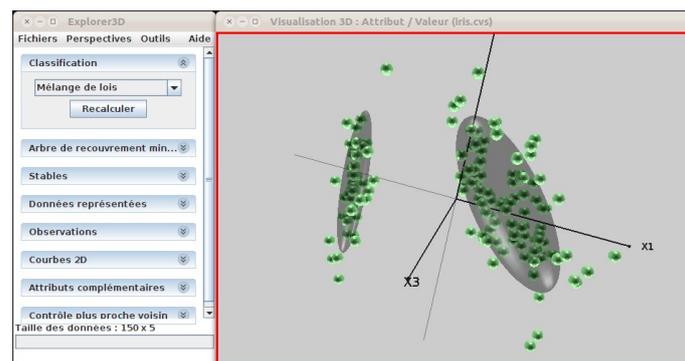


FIGURE 25 – Mélange de lois

7.3.2 Nuées dynamiques (kmeans)

Dans cette méthode classique, on fixe un nombre de groupes, dont les centres sont initialement positionnés de manière aléatoire. Chaque objet est rattaché au centre le plus proche. Les centres sont déplacés jusqu'à obtenir un placement optimal, c'est à dire, dans ce cas, générant des groupes les moins dispersés possible. Dans l'implantation proposée, l'utilisateur peut choisir le nombre de centres, et donc le nombre de groupes générés.

Les nuées dynamiques étant instables, un second outil a été implanté, accessible à partir de la sous-fenêtre “Stables”. Un stable est un ensemble d'objets qui sont toujours classés ensemble (ici, par les nuées dynamiques). Dans cet outil, on choisit le nombre de classes (comme pour kmeans), mais également le nombre de fois où kmeans est exécutée. Au terme de ces calculs, on obtient un ensemble de classes, chacune contenant des objets qui ont été classés ensemble lors de toutes les exécutions de kmeans.

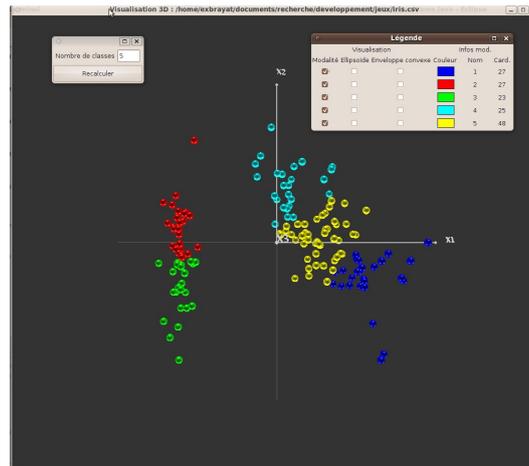


FIGURE 26 – Nuées dynamiques

7.3.3 Nuées dynamiques floues

Il s’agit d’une méthode dérivée de la précédente, à ceci près que chaque objet appartient à tous les groupes, avec un degré d’appartenance lié à sa distance au centre du groupe. Comme précédemment, l’utilisateur peut choisir le nombre de centres. L’intérêt d’une telle méthode consiste à laisser plus de marge d’interprétation à l’utilisateur. Elle est en revanche plus difficile à visualiser. Nous avons donc développé une technique de visualisation spécifique. Il convient de faire apparaître différentes informations : le groupe principal d’un objet, mais également les objets appartenant “à un certain point” à chaque groupe. Nous avons donc développé une visualisation reposant sur la technique des enveloppes convexes (voir figure 27) : pour chaque groupe (ou un sous ensemble de groupes sélectionnés), on fait apparaître l’enveloppe convexe contenant tous les objets appartenant à un groupe “à au moins x pourcents”. Le seuil x peut être réglé dynamiquement par le biais d’un composant graphique, et l’on peut donc observer dynamiquement l’évolution des enveloppes en fonction de x . Il est à noter que cet outil peut également être utilisé lorsque la classification floue n’est pas fournie par un outil interne, mais directement dans les données transmises au logiciel (donc effectuée par un outil tiers).

7.3.4 Classification par densité (DBSCAN)

Contrairement aux techniques précédentes, qui fonctionnent suivant la distance au centre, et génèrent donc des groupes à structure sphérique, les techniques de classification par densité consistent à considérer que les centres de groupes sont les zones spatiales les plus denses. L’agrégation des objets aux groupes se fait ensuite par voisinage. De telles techniques sont utiles lorsque les objets sont structurés par zones de fortes densité, et suivant des formes complexes.

Lors de l’utilisation de cette méthode, on choisit deux paramètres : le rayon dans lequel on considère les voisins, et le nombre minimum de voisins pour considérer que l’on se trouve dans un groupe. Afin d’aider l’utilisateur à régler le rayon, une sphère est visualiser au centre de la vue 3D ; son rayon correspond au rayon utilisé pour le clustering.

La figure 28 représente le résultat d’une telle méthode. Notons que le premier groupe

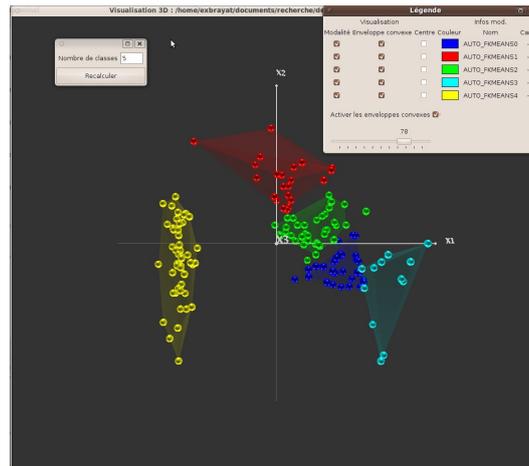


FIGURE 27 – Nuées dynamiques floues (avec enveloppes et seuil)

correspond aux objets isolés. Il peut éventuellement être vide.

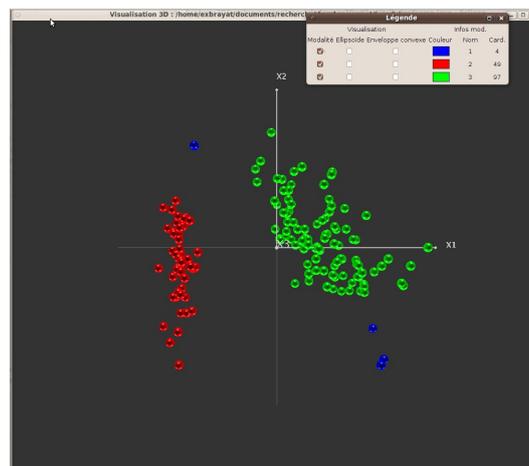


FIGURE 28 – Classification par densité (dbscan)

7.3.5 Arbre minimal de recouvrement et classification hiérarchique

L'arbre minimal de recouvrement est un arbre reliant les objets visualisés. Sans rentrer dans les détails, disons que cet arbre matérialise les proximités entre objets, et que deux objets sont reliés s'ils sont "les plus proches". Visualiser cet arbre permet donc de voir comment s'organise les objets de proche en proche. Un lien long montrera que l'on a des groupes éloignés, etc.

Une classification hiérarchique consiste à regrouper les objets de proche en proche. On commence par regrouper les deux objets les plus proches. Puis on recommence en cherchant les éléments les plus proches, qu'ils s'agisse de deux objets, de deux groupes déjà constitués, ou d'un objet et un groupe. Les distances impliquant un ou deux groupes peuvent être calculées de plusieurs manières. Nous avons retenu celle reposant sur la notion de "saut minimal" : la distance entre un objet et un groupe est la distance entre

l'objet et l'objet du groupe qui en est le plus proche. Entre deux groupes, il s'agit de la distance entre les objets les plus proches de ces deux groupes. On produit ainsi un arbre de classification, également appelé dendrogramme. Il s'agit du type de structure rencontrée pour décrire, par exemple, la classification des espèces. En général, cette structure est représentée à plat. Ici, elle est superposée à la vue 3D (voir fig. 29).

On voit qu'il existe un lien entre les concepts d'arbre de recouvrement minimal et de classification hiérarchique. Le nombre de classes est fixé par l'utilisateur. Il revient à couper l'arbre, en partant du sommet, afin d'avoir un nombre de branches égal au nombre de classes souhaitées.

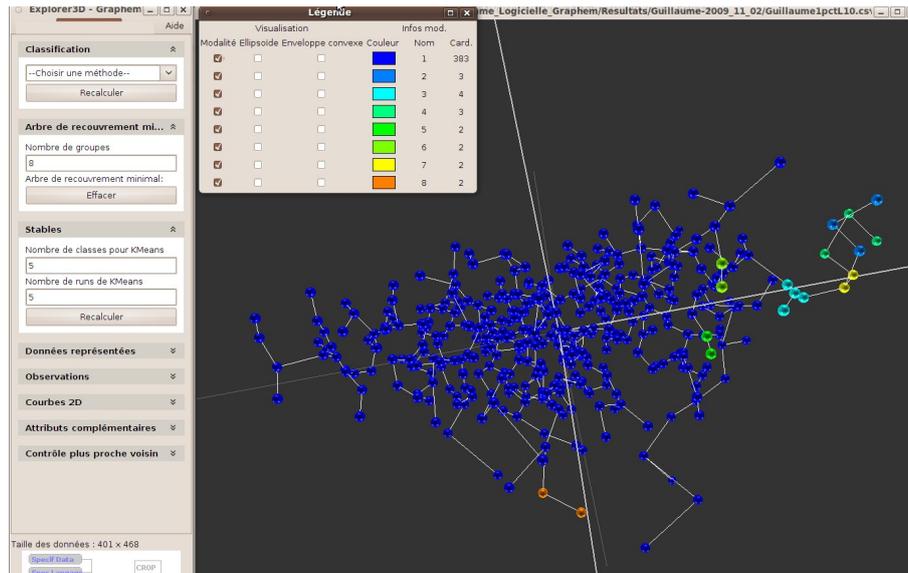


FIGURE 29 – Arbre de recouvrement minimal et classification hiérarchique

7.4 Reconfiguration spatiale

Il s'agit d'une fonctionnalité expérimentale que nous vous invitons à tester. Cet outil, accessible à partir des menus (Outils/Interaction), permet de poser des contraintes sur la proximité ou l'éloignement spatial de divers objets, et donc de reconfigurer l'espace de projection afin de respecter au mieux les nouvelles contraintes (tout en offrant par ailleurs une projection limitant les distorsions). Nous vous invitons dans un premier temps à manipuler l'outil "Comparaison".

7.4.1 Comparaison

Procédez comme suit :

- Cochez la case "ObjetA", puis cliquez sur un objet pour le sélectionner. Une "cible" rouge apparaît autour de l'objet (voir fig. 30). Une cible verte apparaît également. Elle désigne le second objet (celui que l'on souhaite rapprocher ou éloigner de l'objet A). Par défaut, il s'agit de l'objet portant le numéro 0.

- Cochez la case “ObjetB”. Cliquez ensuite sur l’objet dont vous souhaitez modifier la distance par rapport à l’objet A. La cible verte se déplace alors sur cet objet (voir fig. 30).
 - A l’aide de l’ascenseur horizontal, posez la contrainte de distance : vers la gauche, les objets doivent se rapprocher. Vers la droite, ils doivent s’éloigner (voir fig. 30).
 - Cliquez sur “Valider” pour valider votre contrainte.
 - Cliquez sur “Recalculer”, pour générer une vue 3D tenant compte de votre contrainte.
 - Lorsque les calculs sont terminés, les cibles disparaissent et sont remplacées par les numéros des objets. On observe ensuite la déformation de l’espace (voir fig. 31).
- Plusieurs contraintes peuvent être saisies pour un même calcul. Il est toutefois déconseillé, pour l’instant, d’ajouter beaucoup de contraintes, ou de mettre des contraintes opposées.

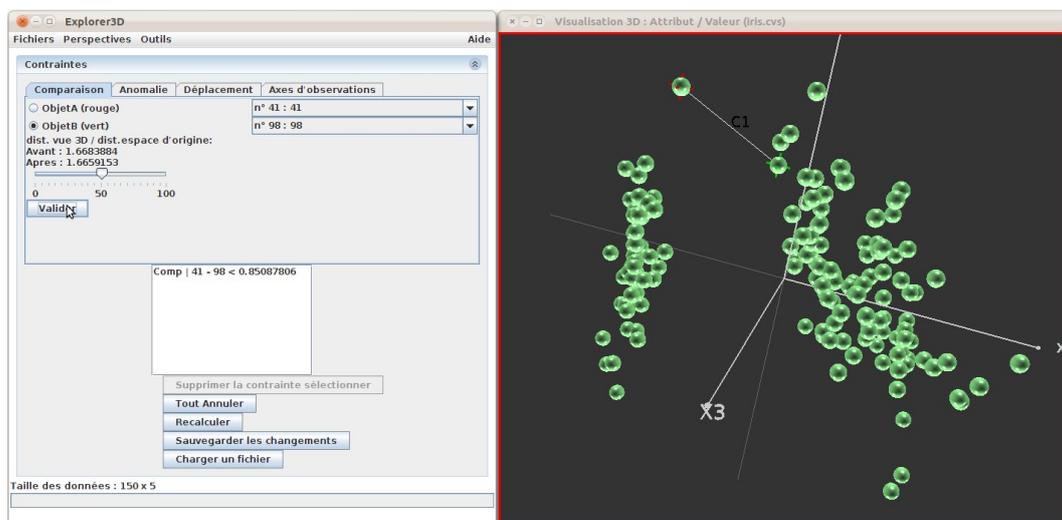


FIGURE 30 – Correction de distance : On peut voir les deux points concernés dans la vue 3D (point A-41 : croix rouge ; point B-98 : croix verte). La contrainte saisie demande que les deux points soient rapprochés (slider de la fenêtre de gauche, et contrainte explicite exprimée en dessous (Comp | 41-98 < 0.85...))

7.4.2 Anomalie

L’outil Anomalie permet de modifier le positionnement relatif de trois objets. Pour ce type de reconfiguration, on manipule trois objets A, B et C. Le principe est de modifier la distance entre A et C relativement à la distance entre A et B. En d’autres termes, il s’agit de rapprocher ou d’éloigner C par rapport à A, de manière à ce qu’il soit plus proche ou plus lointain de A que B.

On procède comme précédemment pour sélectionner les objets.

7.4.3 Déplacement

L’idée de ce type de reconfiguration consiste à modifier le voisinage d’un objet. Cette méthode est assez complexe et son utilisation est déconseillée pour l’instant. Le mode opératoire est le suivant : Il faut commencer par définir le voisinage : on sélectionne les

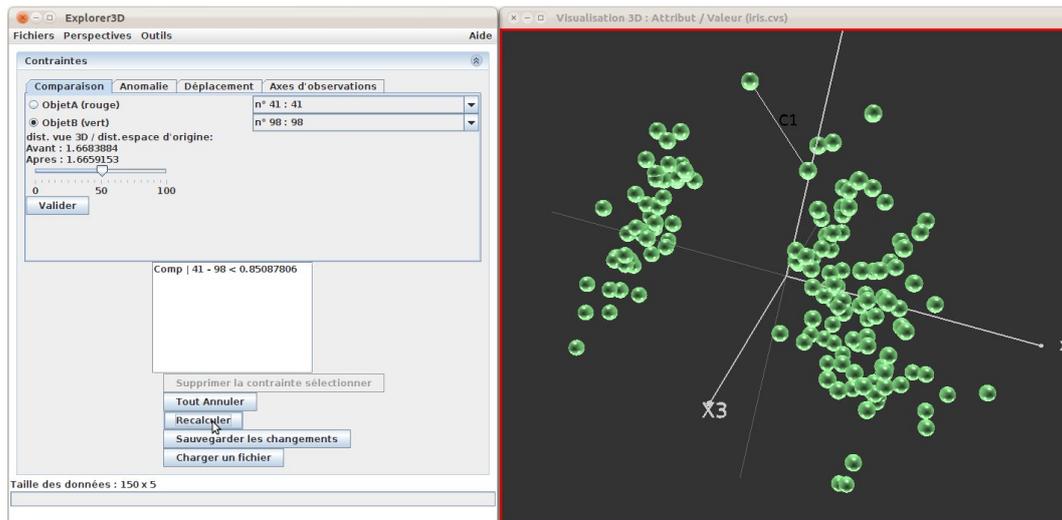


FIGURE 31 – Correction de distance : la contrainte a été appliquée. On peut voir que la projection a changé et que les deux points concernés ont été rapprochés.

objets du voisinage un par un puis on clique sur “ajouter”. Ensuite on coche “objet à déplacer”, et on indique l’objet que l’on souhaite éloigner de ce voisinage. Il suffit ensuite, comme dans les cas précédents, de cliquer sur “Valider”, puis sur “Recalculer”.

7.4.4 Axes d’observation

Afin d’observer un sous-ensemble d’objets, il est possible de mettre en place des axes d’observations. Pour cela, on se rend dans l’onglet “Axes d’observation” (voir figure 32), puis on sélectionne les deux objets formant les extrémités de l’axe (ou, dans un sens plus strict, du segment) observé. On règle le nombre de points à afficher (par défaut, 10), et on clique sur “Valider”. Les dix objets les plus proches du segment (et se projetant sur celui-ci) sont alors sélectionnés, et un tableau indiquant leur numéro et leurs attributs complémentaires apparaît (fenêtre “caractéristiques...”). On peut gérer simultanément plusieurs axes d’observation, et les dissimuler / supprimer grâce à la fenêtre “Axes :”

7.5 Exploration à partir des images

Afin de rendre l’espace de projection plus lisible, il est possible de ne faire apparaître qu’un ensemble d’objets choisis à partir des images associées. Pour cela, il faut bien entendu avoir défini l’attribut complémentaire correspondant aux images associées. Il suffit ensuite de cocher l’élément “Exploration à partir des images” dans le menu “Outils”. Les objets sont alors tous retirés de la vue 3D (ils sont rendus invisibles). Une fenêtre “Images associées” s’ouvre dans laquelle les images apparaissent (fig. 33). L’utilisateur choisit les objets qui l’intéressent soit grâce à la liste déroulante située en haut de la fenêtre (onglet “Liste” + validation), soit en saisissant directement son nom (onglet “Nom” + validation). A chaque ajout, l’image correspondante est affichée et l’objet est réaffiché dans la scène 3D. Lorsqu’une image est cliquée, elle est considéré comme sélectionnée, et il en va donc de même de l’objet correspondant qui est placé en surbrillance dans la

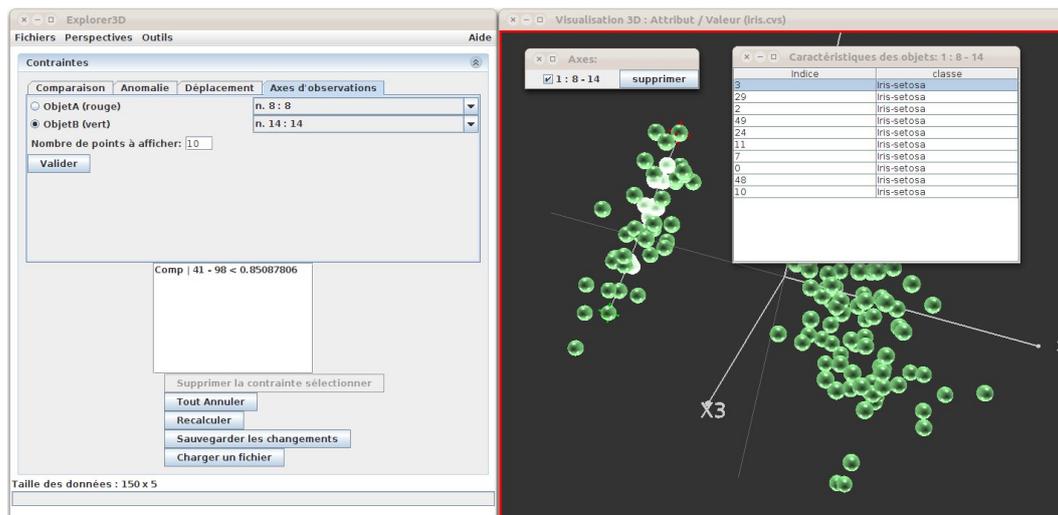


FIGURE 32 – Mise en place d’un axe d’observation. On observe les dix objets les plus proche de l’”axe” 8-14.

fenêtre de visualisation. On peut sélectionner plusieurs images. Pour désélectionner une image, il suffit de cliquer dessus.

L'utilisateur peut choisir de visualiser les voisins, dans l'espace, de chaque objet exploré à partir de son image : lorsqu'une image est sélectionnée, elle apparaît, ainsi que ses "n" plus proches voisins, "n" étant fixé à 0 par défaut. Le nombre de voisins affichés et réglés dans l'onglet "Voisins". Si l'image est désélectionnée, ses plus proches voisins sont aussi dissimulés.

Cette exploration pourra également être superposée à l'option générale "Affichage dynamique des images", qui permet de visualiser une image (dans la vue 3D) lorsque le pointeur de la souris s'arrête sur l'objet correspondant.

7.6 SVM

Les SVM ou séparateurs à vaste marge sont un outil de classification cherchant à séparer au mieux deux catégories d'objets. Ils sont donc utilisés lorsque l'on connaît la classe d'au moins une partie des objets.

Au plus simple, la séparation entre les deux groupes est un plan, tel que de chaque côté on trouve uniquement les objets d'un groupe. De plus, ce plan est le plus éloigné possible des objets des deux groupes (il est donc, en quelque sorte, "médian"). Une séparation aussi simple n'étant pas toujours disponible, on peut tolérer une marge d'erreur (quelques objets du mauvais côté) et/ou utiliser un hyper plan situé dans un espace plus complexe que l'espace de description ou l'espace de projection. Pour cela, on a généralement recours à un noyau (voir section 7.7). Une fois l'(hyper)plan séparateur défini, on peut estimer le groupe d'appartenance des objets non étiquetés, suivant le côté ou ils se trouvent (et leur éloignement par rapport à l'(hyper)plan.

Par ailleurs, si l'on a plus de deux groupes, on peut soit définir des SVM entre deux groupes, ou entre un groupe et tous les autres.

Dans Explorer3D les fenêtres de manipulation des SVM sont accessibles à partir du menu

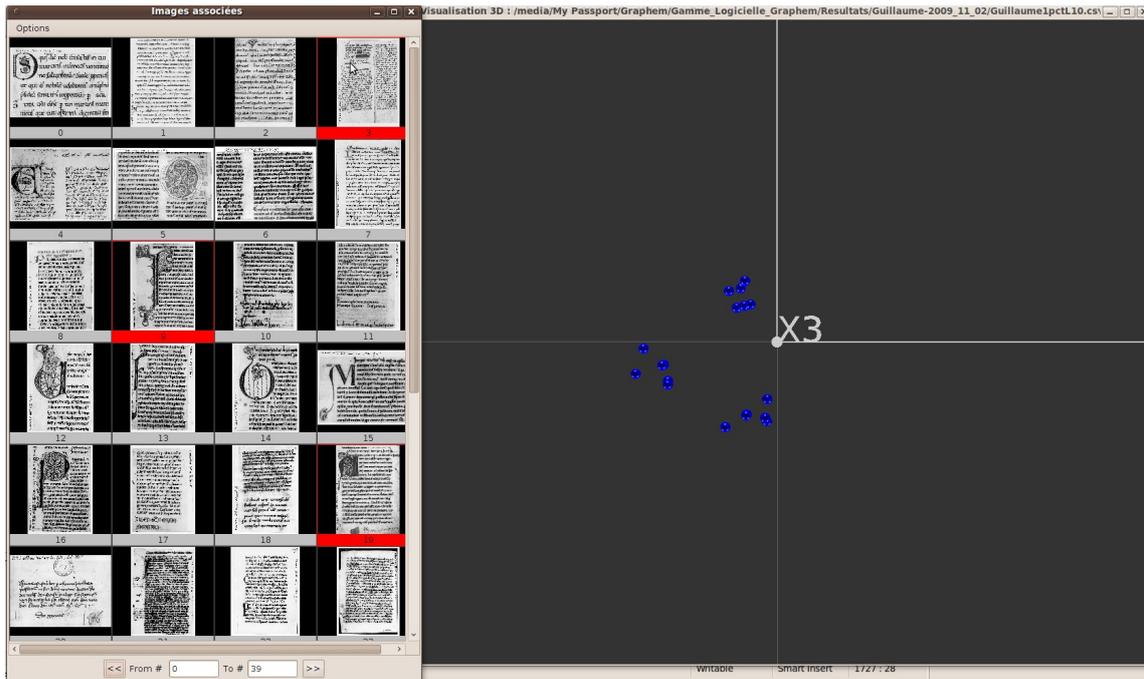


FIGURE 33 – Exploration à partir des images (affichage des 5 plus proches voisins activé)

“Perspectives/SVM”. Pour les utiliser, il faut auparavant avoir défini l’attribut de classe. La figure 34 montre un exemple d’utilisation des SVM dans un cas simple : le très classique jeu de données “iris”. Nous pouvons voir dans la fenêtre de contrôle que nous avons tracé un SVM entre le groupe “iris-setosa” et le reste des objets, qu’il se fait sur les données projetées, et qu’il est visualisé dans la fenêtre 3D. Notons que la visualisation sous forme de plan ne peut se faire que pour les données projetées.

Les fonctionnalités disponibles sont les suivantes :

- fenêtre “SVM” : permet essentiellement de choisir les objets allant de chaque côté du plan séparateur.
 - “constant” : laisser la valeur par défaut (1) ;
 - objets du premier côté : une classe parmi les classes connues ;
 - objets du second côté : une classe parmi les classes restantes, ou l’ensemble des classes restantes ;
- fenêtre “Mode d’affichage du SVM”
 - mode de calcul : on peut calculer le SVM pour l’ensemble des attributs de description des objets (tels que fournis dans le fichier) ou pour leur projection. Le SVM sur l’espace d’origine sera probablement plus précis, mais ne pourra pas être visualisé dans l’espace 3D (puisqu’il est défini dans l’espace d’origine). Dumoins, il ne pourra pas être visualisé sous la forme d’un plan (voir point suivant).
 - mode d’affichage : le mode de visualisation le plus simple, dans la fenêtre 3D, est le plan. Toutefois, il n’est pas toujours utilisable. Les autres modes disponibles sont : “coloration” et “remplacer 3ème axe”.

Avec le mode coloration, les objets sont colorés suivant un dégradé de couleur (rouge d’un côté, vert de l’autre). La figure 35 illustre ce mode de représentation.

Avec le mode “remplacer 3ème axe”, la projection des objets ne se fait plus que suivant

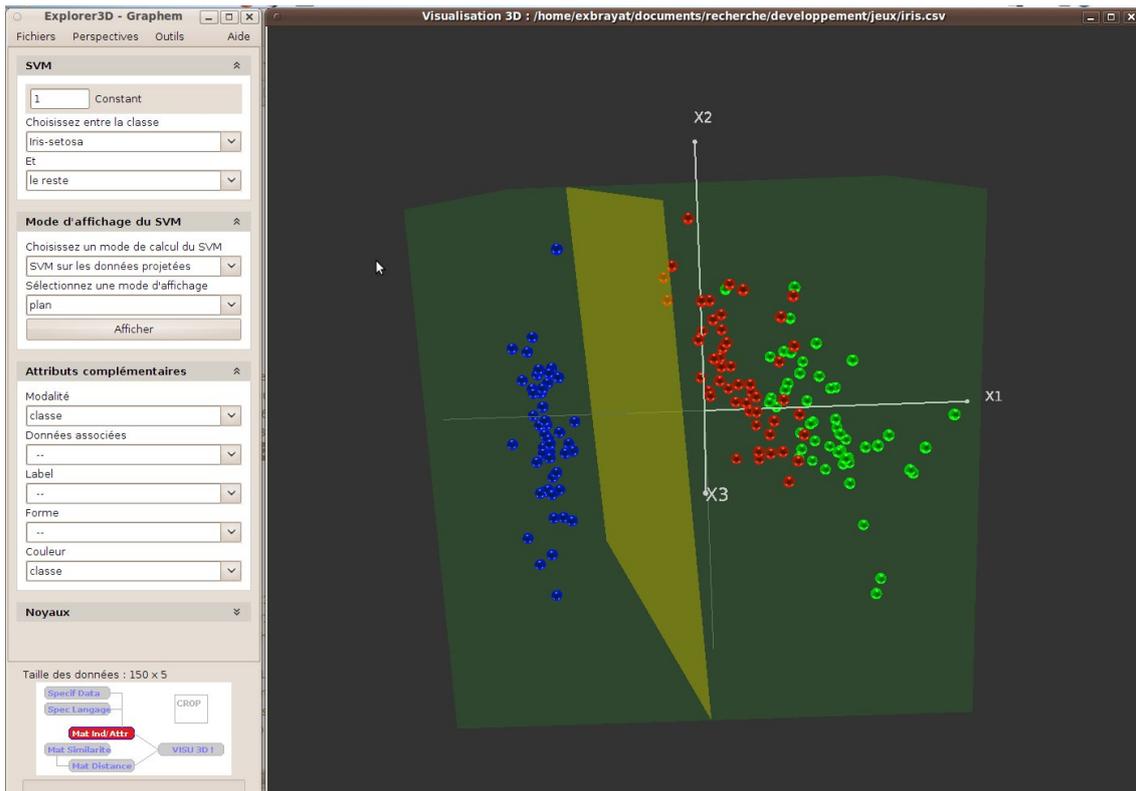


FIGURE 34 – Exemple de SVM avec plan séparateur (iris)

les axes “x1” et “x2” ; leur coordonnée suivant l’axe “x3” correspond à leur éloignement de l’hyperplan séparateur. La figure 36 illustre ce mode de représentation.

Ces deux derniers modes peuvent être utilisés avec des SVM calculés dans l’espace d’origine ou à l’aide d’un noyau.

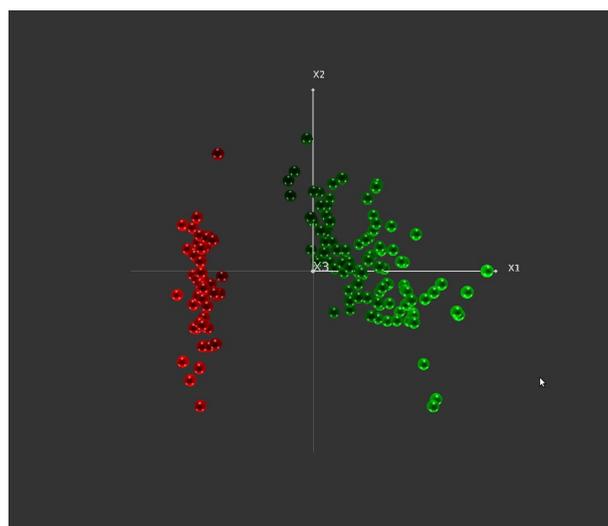


FIGURE 35 – Exemple de SVM avec coloration

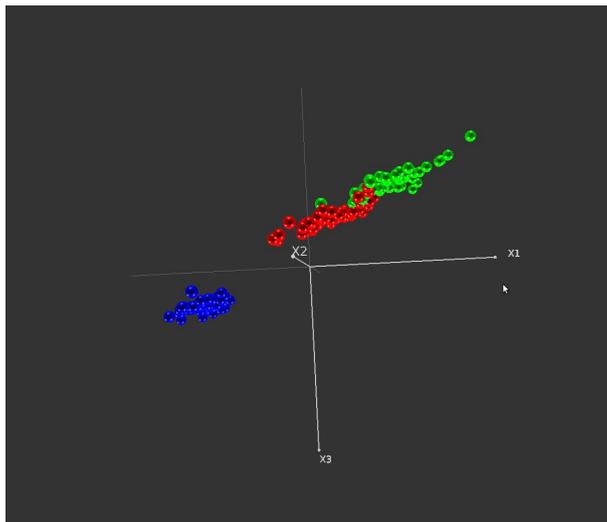


FIGURE 36 – Exemple de SVM 3ème axe

7.7 Noyaux

Les noyaux sont un outil puissant, dont l’usage en apprentissage automatique s’est principalement répandu grâce aux SVM. En effet, l’espace d’origine et l’espace de projection ne permettent pas toujours de trouver un hyperplan séparateur.

Les noyaux reposent sur l’idée de calculer de nouveaux attributs par combinaison des attributs existants. Ces combinaisons ne sont pour la plupart pas linéaires, mais doivent respecter certaines propriétés. Pour ces aspects techniques, on invitera l’utilisateur à consulter la bibliographie du domaine.

Les noyaux sont disponibles dans les perspectives “SVM” et “ND->3D”. La sous-fenêtre “Noyaux” permet de sélectionner un noyau, et le cas échéant de fixer les paramètres associés. Le noyau est choisi dans une liste déroulante. Lorsqu’un noyau est choisi, il s’applique à tout nouveau calcul de projection (clic sur le bouton “Calculer” de la sous-fenêtre “Méthode de projection”).

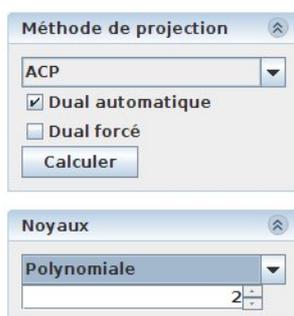


FIGURE 37 – Fenêtre de contrôle des noyaux. On voit qu’un noyau polynomial a été choisi, avec un polynôme de niveau 2.

Nous donnerons uniquement et sommairement ici la liste des noyaux disponibles dans *Explorer3D* :

- noyau linéaire : ce noyau ne fait rien. Il permet d'utiliser l'espace d'origine lorsque les noyaux sont activés.
- noyau polynomial
- noyau RBF
- Sigmoide
- Isomap
- LLE

Notons que même l'utilisateur néophyte à la recherche, par exemple, d'une projection satisfaisante peut tester ces différents noyaux.

7.8 Visualisation 2D

La sous-fenêtre "Courbes 2D" permet de visualiser diverses courbes complémentaires :

- **Diagramme de Shepard** ce diagramme donne le rapport entre les distances dans l'espace d'origine et dans l'espace de projection.
- **2D (symboles)** et **2D (labels)** correspondent à la même projection qu'en 3D, mais sur les deux premiers axes. En mode *symboles*, chaque objet est associé à un symbole correspondant à sa classe. En mode *labels*, on affiche le label des objets (par défaut leur numéro).
- **Contributions**
- **Coscarrés**
- **Cellules de Voronoi** les outils de ce type servent à analyser la vue 3D. Elles prennent un instantané de la vue 3D, dans lequel les cellules de Voronoi sont tracées autour de chaque objet, et colorées différemment suivant les cas : **Proximité** mesure la proximité des objets dans l'espace d'origine. Il s'agit de visualiser la distance des objets par rapport à un objet de référence, qu'il faut sélectionner avant de tracer la courbe ; **Distorsion** mesure les distorsions issues de la projection 3D. Ces vues ne reposent pas sur les cellules de Voronoi, mais sur la triangulation de Delaunay, qui constitue son complémentaire. Les objets voisins (vue 2D) sont reliés par un quadrilatère, dont la couleur varie du blanc (pas de distorsion) au noir (forte distorsion). Ces deux types de coloration peuvent être utilisés pour mesurer les proximités ou distorsions entre les vues 2D et 3D, 2D et ND, 3D et ND (espace d'origine). Ces outils ne sont pas complètement fonctionnels dans cette version du logiciel.

7.9 Commande via le réseau

Cet outil permet notamment d'interfacer **Explorer3D** avec un autre logiciel, afin de visualiser les résultats de ce dernier. Cette fonction est activée via la fenêtre des options (voir section 6.7). **Explorer3D** se met en écoute sur le port 50 000, et reçoit les commandes via le protocole json.

La liste des commandes disponibles est assez simple pour l'instant et se concentre sur la réception d'un tableau d'objets et sur la mise à jour de la classe d'appartenance de ces objets (et de la visualisation de celle-ci par coloration).

- **chargement d'objets** : Les objets sont sous forme de tableau attribut/valeur. Le message comporte deux rubriques : le nom de la commande ("msg" : "set-data"), puis la description des objets ("val" : [[objet1], [objet2], ...]). Voici un exemple avec 4 objets à 4 dimensions chacun :

```

{
  "val": [
    [
      0.09995818,
      0.8450656,
      0.31611204,
      0.9885413
    ],
    [
      0.20870173,
      0.6817836,
      0.30197513,
      0.20966005
    ],
    [
      0.95933825,
      0.7643002,
      0.23336774,
      0.7409621
    ],
    [
      0.075930595,
      0.12991834,
      0.6995411,
      0.092056274
    ]
  ],
  "msg": "set-data"
}

```

- **fourniture de la classe des objets (liste complète)** : On fournit la classe de tous les objets. Le message comporte trois rubriques : le nom de la commande ('`msg`': '`set-classes`'), puis le nombre de classes ('`nbclasses`': valeur) et enfin la description des classes ("val" : [[classe_objet1], [classe_objet2], ...]). Voici un exemple avec nos 4 objets; il y a 2 classes; les 3 premiers objets appartiennent à la classe 1, le dernier à la classe 0 :

```

{
  "val": [
    1,
    1,
    1,
    0
  ],
  "msg": "set-classes",
  "nbclasses": 2
}

```

- **fourniture de la classe des objets (liste partielle)** : cette commande permet une mise à jour des classes et *ne peut être invoquée qu'après une première invocation de la*

commande précédente (“set-classes”). Le message comporte deux rubriques : le nom de la commande (‘ ‘msg’ ’ : ‘ ‘set-classes-fast’ ’) et la liste des modifications. Celle-ci se présente sous forme d’un tableau (“val” : [[classe_objet1], [classe_objet2], [saut_objets] dont l’encodage est le suivant : une valeur positive indique une nouvelle valeur de classe pour l’objet correspondant ; une valeur négative indique un saut dans la liste des objets (objets dont la classe ne change pas). Voici un exemple avec nos 4 objets ; la classe des deux premiers objets n’est pas changée (saut de 2). Le troisième passe en classe 0 ; la classe du dernier objet ne change pas (saut de 1).

```
{
  "val": [
    -2,
    0,
    -1
  ],
  "msg": "set-classes-fast"
}
```

7.10 Importation de données

L’outil d’import est ouvert dans le cas d’un jeu de données non reconnu, ou directement par le menu “Fichiers / outil d’importation de données”. Cet outil assez intuitif commence par chargement du fichier source (fig. 38) ; cette étape n’est explicite que dans le cas de l’accès via le menu. L’utilisateur indique ensuite les délimiteurs utilisés dans le fichier source : séparateur de champs, symbole de décimale et séparateur de texte (fig. 39). Une vue du fichier source est présentée durant cette étape. Le jeu de données utilisé pour cet exemple est le jeu “Parkinson” disponible sur le site de l’UC Irvine.

A l’étape suivante, le contenu du fichier est présenté sous forme tabulaire (fig. 40) ; l’uti-

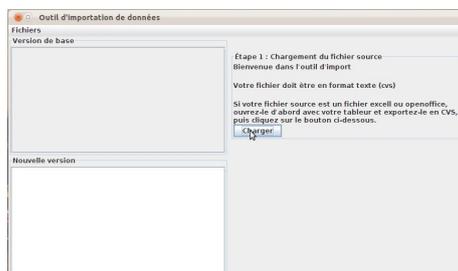


FIGURE 38 – Import : fenêtre initiale

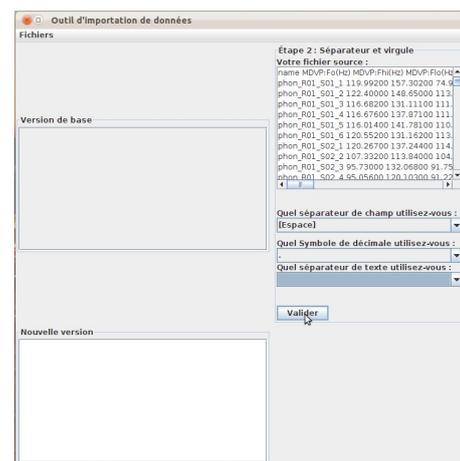


FIGURE 39 – Import : réglage des délimiteurs

lisateur indique le type de données importées. A l’heure actuelle, la seule option prise en charge concerne les fichiers attributs/valeurs.

L’utilisateur indique ensuite le rôle des lignes et colonnes dans le fichier source (fig. 41). Dans l’exemple présenté, les objets sont décrit à raison de 1 par ligne. Les attributs descriptifs occupent les lignes 2 à 196 et les colonnes 2 à 24. La première colonne contient les

noms des objets ; si l’agit donc d’un attribut complémentaire. La première ligne contient les noms des attributs. Dans la version actuelle, les attributs complémentaires et les attributs descriptifs ne peuvent pas être entrelacés.

L’outil d’import propose ensuite un récapitulatif des données importées (fig. 42), puis

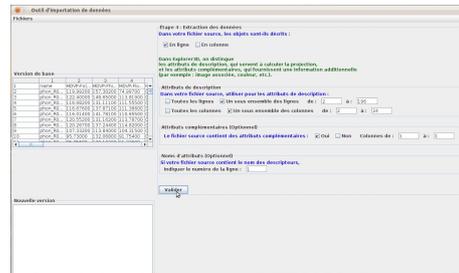
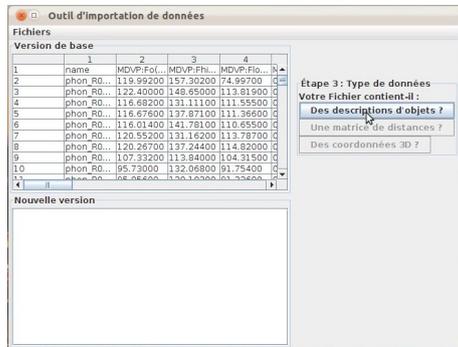


FIGURE 40 – Import : type de données

FIGURE 41 – Import : description du contenu

la vue finale du fichier au format **Explorer3D** (fig. 43). L’utilisateur a le choix entre une sauvegarde pure et une sauvegarde suivi du chargement du fichier. La figure 44 présente l’ACP issue de ce jeu de données après import.

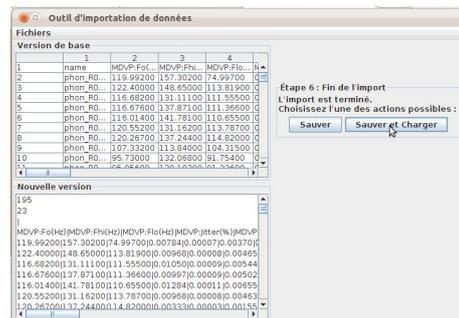
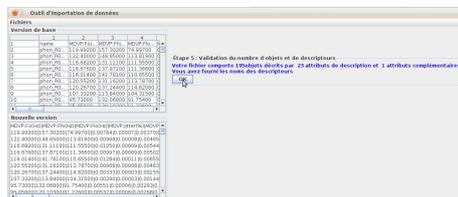


FIGURE 42 – Import : récapitulatif

FIGURE 43 – Import : fin d’import

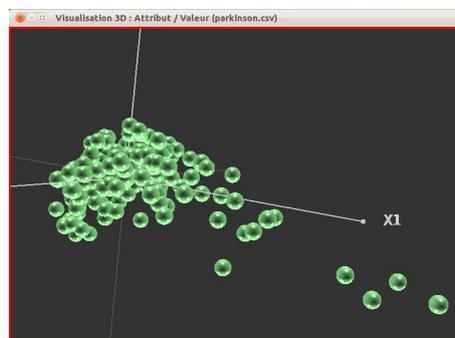


FIGURE 44 – Projection après import (jeu de données “Parkinson”)

8 Outils secondaires

8.1 Voyage à travers les dimensions

Cette fonction, située dans le menu "Outils", permet de choisir à la volée les dimensions visualisées (réduction de dimensions). Il est possible de choisir ce qui est projeté sur chaque axe, et ce avec des effets de transition (voir figures 45 à 47). Dans les cas les plus simples, cela s'apparente aux interactions offertes par la fenêtre "Axes de projection" (section 6.9). toutefois, les possibilités offertes ici sont plus étendues : on peut mélanger les axes, par exemple faire en sorte que l'axe x3 porte les informations de plusieurs axes calculés (somme).

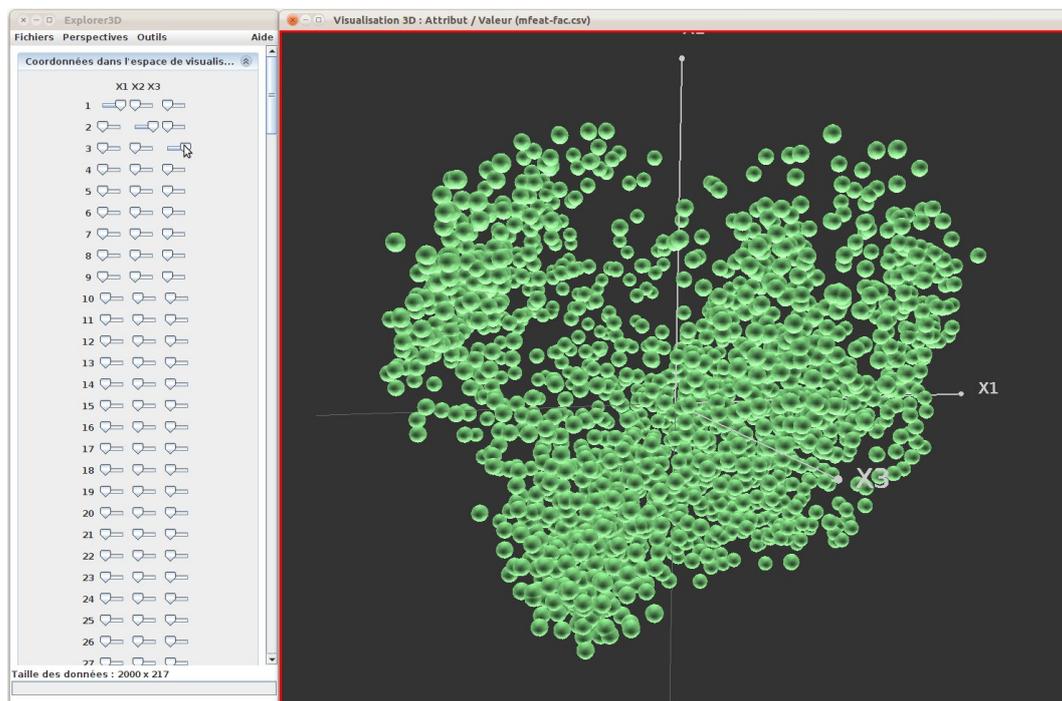


FIGURE 45 – Voyage à travers les dimensions : les objets sont projetés suivant les trois axes les plus représentatifs.

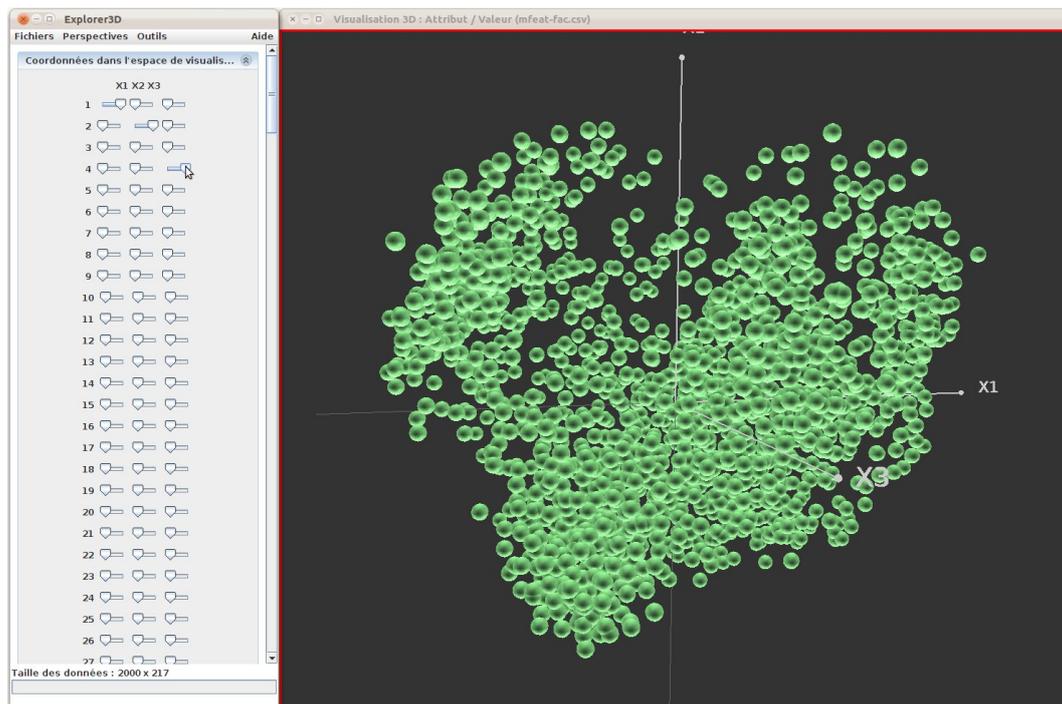


FIGURE 46 – Voyage à travers les dimensions : on choisit de remplacer le troisième axe par le quatrième.

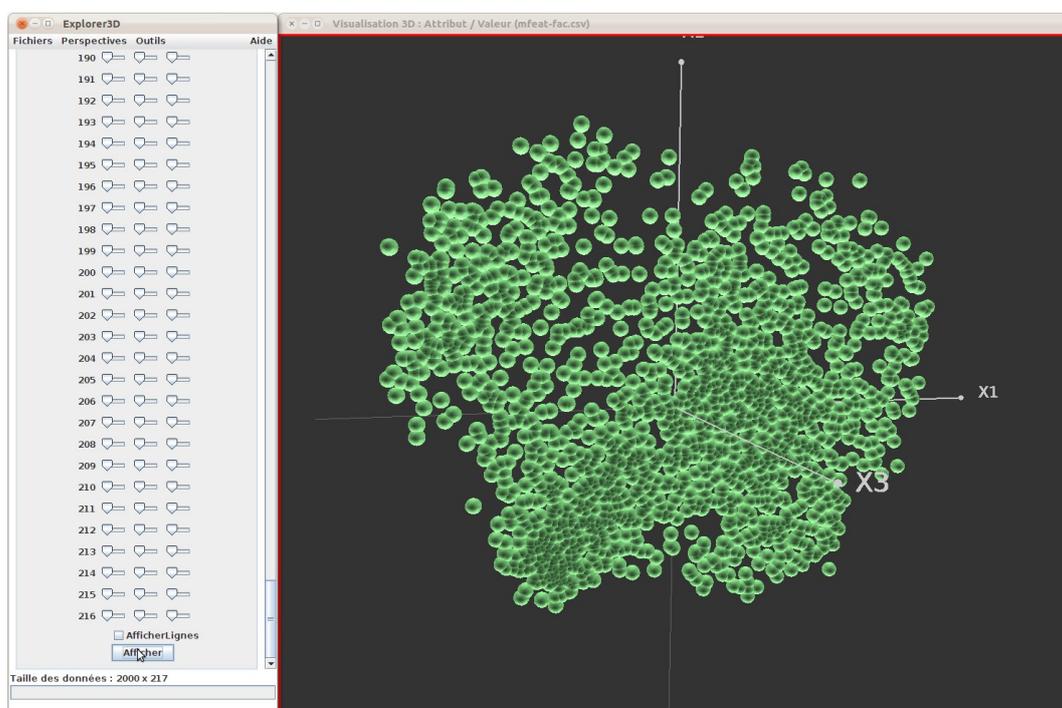


FIGURE 47 – Voyage à travers les dimensions : après le changement d'axe.

Index

- analyse discriminante linéaire, 14
- analyse en composantes principales, 13
- arbre minimal de recouvrement, 31
- attributs complémentaires, 15
- axes de projection, 23
 - choix, 24
 - coloration, 24
 - voyage à travers, 43

- centrage nuage / image, 24
- classification hiérarchique, 31
- coloration
 - par dégradé, 16
 - texture, 17
- conditionnement de données, 14
- crop
 - circulaire, 20
 - rectangulaire, 19

- dimension couleur, 24

- explorateur de vues, 27

- import de données, 41

- légende multiclassés, 22
- légende sur attribut continu, 22
- loupe, 19

- noyaux (méthodes à), 38

- observations, 25
- options (gestion des), 20

- positionnement multidimensionnel (MDS),
 - 14
- pré-traitement, 14
- pré-traitements, 9
- projection
 - coordonnées 3D, 15
 - matrice de distances -> 3D, 14
 - ND -> 3D, 13
- projection interactive
 - anomalie, 33
 - axes d'observation, 34
 - comparaison, 32
 - déplacement, 33

- Sheppard (diagramme de), 39
- svm, 35

- visualisation 2D, 39
- Voronoi (cellules de), 39
- vue active, 26

- zoom dimensions peu discriminantes, 23