

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieur et de la recherche scientifique

Université de 8 Mai 1945 – Guelma -

Faculté des Mathématiques, d'Informatique et des Sciences de la matière

Département d'Informatique



## Mémoire de Fin d'études Master

**Filière :** Informatique

**Option :** Systèmes Informatiques

**Thème :**

---

### Réduction de dimensionnalité et variété en mosaïquage

---

**Encadré Par :**

Dr. GUEROUI Nadia

**Présenté par :**

RAHDOUNE Iskander

**Septembre 2021**

## Remerciements

*Tout d'abord je tiens à remercier Dieu, le tout puissant et miséricordieux, qui m'a donné la force, l'intelligence et la patience d'accomplir ce modeste travail.*

*Un grand remerciement à Mme GUEROUI Nadia pour son encadrement. Elle était toujours montrée à l'écoute et très disponible tout au long de la réalisation de ce mémoire, ainsi pour l'inspiration, l'aide et le temps qu'elle a bien voulu me consacrer, je te dis merci.*

*Des remerciements à ma famille et mes amis pour leurs encouragements, grâce à eux j'ai pu surmonter tous les obstacles.*

*Enfin, mes vifs remerciements vont également aux membres de jury pour l'intérêt qu'ils ont porté à mon projet en acceptant d'examiner et juger mon travail.*

## Résumé

Les données de haute dimension ont été regroupées dans des zones que nous voulons découvrir, et chaque jour, la perception du cerveau de l'être humain est confrontée au problème de la réduction de la dimensionnalité pour trouver des structures significatives de basse dimension, cachées dans leurs observations de haute dimension. Le cerveau humain réduit la dimensionnalité des problèmes en fractions de seconde (reconnaissance des visages, textures, lettres, mots, scènes, etc ...). L'objectif pour le scientifique est donc similaire : réduire (projeter) des problèmes de haute dimensionnalité dans une basse dimensionnalité en gardant le maximum de modes de variabilité dans les données, afin de les traiter en un temps raisonnable avec ces dimensions réduites. Cependant, de nombreux ensembles de données contiennent des structures non linéaires qui sont invisibles pour les techniques classiques telles que l'analyse en composantes principales (ACP) et l'échelonnement multidimensionnel (MDS), ces deux méthodes ne parvenant pas à détecter les véritables degrés de liberté de l'ensemble des données. La structure sous-jacente des données, compte tenu des observations de haute dimension (difficiles à stocker, à traiter, à analyser et à visualiser sur le plan informatique), doit être réduite. Dans cette situation, le problème de la DR, connu sous le nom de Manifold learning, devient encore plus difficile, surtout lorsque la dimension d'entrée est très élevée.

Ce travail étudie une approche pour modéliser le problème de l'assemblage d'images afin de combiner plusieurs ensembles d'images des scènes naturelles, où chaque ensemble contient des images appartenant à la même représentation de la scène par des techniques de réduction de la dimensionnalité non linéaire. Notre objectif est d'identifier la suite d'images voisines parmi tous les ensembles d'images.

L'approche proposée s'appuie sur les progrès récents de l'apprentissage par les variétés qui permettent de découvrir des tendances non linéaires dans les données, mais notre cas est différent des approches NLDR utilisées, qui supposent que la correspondance est connue avant de calculer la distance du plus court chemin. Nous avons besoin d'un moyen de trouver la correspondance d'abord, avant de calculer la distance du chemin le plus court et la correspondance du collecteur. Pour ce faire, il est crucial d'employer une représentation d'ensemble d'images adéquatement flexible pour traiter la non-linéarité. {Vous parlez peut-être d'expériences }

### Mots clés :

Apprentissage de variété, Réduction de dimensionnalité, apprentissage profond, données massives, mosaïquage, Isomap, LLE, transformation non linéaire, distance géodésique, k-plus proche voisin, norme MPEG 7.

## **Abstract**

High dimensional data have been grouped into areas that we want to discover, and every day the perception of the human brain is faced with the problem of reducing dimensionality to find meaningful low dimensional structures, hidden in their high dimensional observations. The human brain reduces the dimensionality of problems in fractions of a second (recognition of faces, textures, letters, words, scenes, etc ...). The objective for the scientist is therefore similar: to reduce (project) high dimensionality problems into low dimensionality while keeping the maximum of variability modes in the data, in order to process them in a reasonable time with these reduced dimensions. However, many data sets contain nonlinear structures that are invisible to conventional techniques such as principal component analysis (PCA) and multidimensional scaling (MDS), both of which fail to detect the true degrees of freedom in the data set. The underlying structure of the data, given the high-dimensional observations (difficult to store, process, analyze and visualize computationally), must be reduced. In this situation, the DR problem, known as Manifold learning, becomes even more challenging, especially when the input dimension is very high.

This work studies an approach to model the image stitching problem to combine multiple sets of images of natural scenes, where each set contains images belonging to the same representation of the scene by nonlinear dimensionality reduction techniques. Our objective is to identify the sequence of neighboring images among all the image sets.

The proposed approach builds on recent advances in variety learning that can discover nonlinear patterns in the data, but our case is different from used NLDR approaches, which assume that the match is known before computing the shortest path distance. We need a way to find the match first, before computing the shortest path distance and the manifold match. To do this, it is crucial to use an adequately flexible image set representation to deal with nonlinearity. {You maybetalking about experiments}

### **Keywords:**

Variety learning, Dimensionality reduction, deep learning, massive data, mosaicking, Isomap, LLE, nonlinear transformation, geodesic distance, k-nearest neighbor, MPEG 7 standard.

## Table des matières

<i>Introduction générale</i> .....	2
------------------------------------	---

### **CHAPITRE I ETAT DE L'ART**

<i>Introduction</i> .....	4
<i>Réduction de dimensionnalité</i> .....	4
<i>1.Travaux connexes</i> .....	7
<i>2. Apprentissage de variété</i> .....	8
Méthodes Manifold Learning.....	9
A - Méthodes de projection :.....	9
B- Méthodes génératives :.....	9
C- Méthodes d'intégration :.....	9
D- Méthodes d'information mutuelle.....	9
Apprentissage non supervise de Variétés.....	10

### **Chapitre 2 Conception**

<i>1.Introduction</i> .....	14
<i>2.APPROCHE PROPOSEE</i> .....	15
<i>3.MPEG-7 Standard</i> .....	11
<i>4.Correspondance de similarité en combinant EHD et CLD</i> .....	15
4.1 Descripteurs des données générales.....	15
4.2 Démarche proposée pour la combinaison des deux valeurs IDCT de CLD et EHD :.....	18
4.3. Pourquoi la DCT ?.....	19
<i>5.Algorithme de calcul pour la DCT</i> .....	20
<i>6.LLE-Algorithmes</i> :.....	22
6.1. Zoom sur LLE :.....	22
6.2. Démarche proposée.....	24
<i>7.Conclusion</i> :.....	24

### **Chapitre 3 Implémentation**

<i>1.Introduction</i> .....	26
<i>2.Données expérimentales et mesures de qualité</i> .....	26

<i>3.Résultats expérimentaux</i> .....	27
<i>4.Conclusion</i> .....	30
<i>Conclusion Générale</i> .....	32
<i>Références Bibliographiques</i> .....	33

## Liste des figures

Figure 1 : une illustration de la procédure de mise en correspondance d'images à l'aide de NLDR .....	2
<b>Chapitre 1</b>	
Figure 1.1. Taxonomie des techniques de réduction de dimensionnalité.....	5
<b>Chapitre 2</b>	
Figure 2.1. : Schéma général de l'approche proposée .....	15
Figure 2.2. Schéma fonctionnel du processus d'extraction du descripteur de disposition des couleurs. ....	16
Figure 2.3. Definition of Sub-image and Image-block in the EHD .....	17
Figure 2.4. Five Types of Edges in the EHD .....	18
Figure 2.5. Le processus d'analyse des images de la base de données en combinant les deux valeurs IDCT de CLD et EHD .....	21
<b>Chapitre 3</b>	
Figure 3.1. LLE avec hybridation=31 et k=3 voisins les plus proches .....	27
Figure 3.2. LLE appliqué à n=195 image k=7 sans hybridation EHD.....	27
Figure 3.3. LLE appliqué à n=15 image k=7 .....	28
Figure 3.4. Incorporation bidimensionnelle de N=217 images de scènes de paysages obtenue par isomap en utilisant k=7 plus proches voisins. ....	28
Figure 3.5. Montrée l'intégration LLE 2-D de N=217 images de scènes de paysage en utilisant k=6 voisins les plus proches.....	29

# **Introduction Générale**



## Introduction générale

Les données du monde réel produisent des quantités massives de jeux de données tels que des images, des signaux vocaux, des photographies numériques, des caractères et des vidéos

Le traitement d'ensembles de données massives à haute dimension est un défi actuel. Lorsqu'on traite de telles données, la haute dimensionnalité est un obstacle à tout traitement efficace des données. En effet, de nombreux algorithmes classiques de traitement de données ont une complexité de calcul qui croît exponentiellement avec la dimension [1]. En revanche, la source peut ne bénéficier que d'un nombre limité de degrés de liberté. Cela signifie que la plupart des variables qui décrivent chaque point de données sont fortement corrélées, du moins localement, ou, de manière équivalente, que l'ensemble de données a une faible dimensionnalité intrinsèque [8,]. Il devrait donc être possible d'obtenir des représentations à faible dimension des échantillons. Les techniques classiques de réduction de la dimensionnalité sont simples et garantissent la découverte de la véritable structure des données [18].

Les techniques de réduction de dimension non linéaire (NLDR) ont toujours été l'un des objectifs de la recherche sur l'apprentissage informatique, et elles se sont avérées utiles dans de nombreux problèmes de vision par ordinateur [9, 10, 11, 18].

Dans le but est de traiter adéquatement les données non linéaires complexes, de nombreuses recherches scientifiques ont montré que les techniques non linéaires sont plus performantes que leurs homologues linéaires dans les tâches artificielles complexes.

Dans ce qui suit suivant, nous avons un ensemble d'images, et nous voudrions identifier les paires d'images qui partagent le plus de voisins, ici la correspondance n'est pas connue avant de calculer la distance du chemin le plus court, et nous avons besoin d'un moyen de trouver la correspondance d'abord, avant de calculer la distance du chemin le plus court basée sur l'information de correspondance et la correspondance de manifold.

Le contenu principal de ce document est le suivant : Section 2. nous décrivons quelques travaux antérieurs liés à la comparaison d'images ; Section 3. Nous présentons deux méthodes d'intégration comme technique de réduction de la dimensionnalité non linéaire. Sect. 4. nous utilisons l'apprentissage par le manifold dans la mise en correspondance d'images ; la section 5 présente une hybridation des valeurs IDCT du CLD et de l'EHD comme indicateur de distance robuste de la similarité dans l'ensemble de données ; la section 6 démontre et discute les résultats des expériences ; la section 7 donne les conclusions.

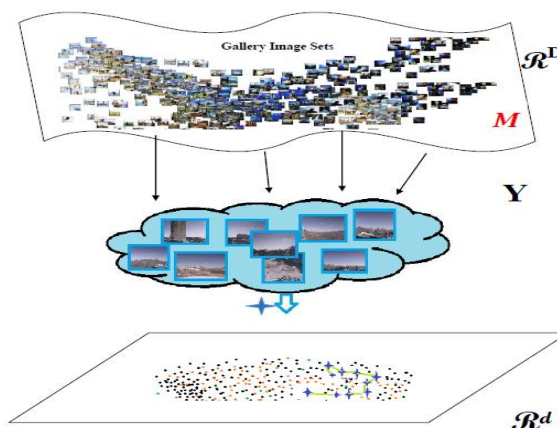


Figure 1 : une illustration de la procédure de mise en correspondance d'images à l'aide de NLDR

***CHAPITRE I***  
***ETAT DE L'ART***

### Introduction

Le cerveau humain réduit instantanément et sans s'en rendre compte la taille des problèmes quotidiens tels que la reconnaissance faciale, les textures, les lettres et les mots.

En neurobiologie, 30 000 nerfs auditifs et 1 000 000 de nerfs visuels, mais nous n'extrayons qu'un sous-ensemble relativement petit d'informations importantes sur le plan de la perception.

Le but pour le scientifique est donc similaire de réduire (projeter) des problèmes à haute dimensionnalité dans une dimensionnalité minimale tout en gardant le maximum de modes de variabilité dans les données, afin de les traiter avec des algorithmes qui fonctionnent en temps raisonnable avec ces dimensions réduites (SVM, réseaux multicouches, etc...)

### Réduction de dimensionnalité

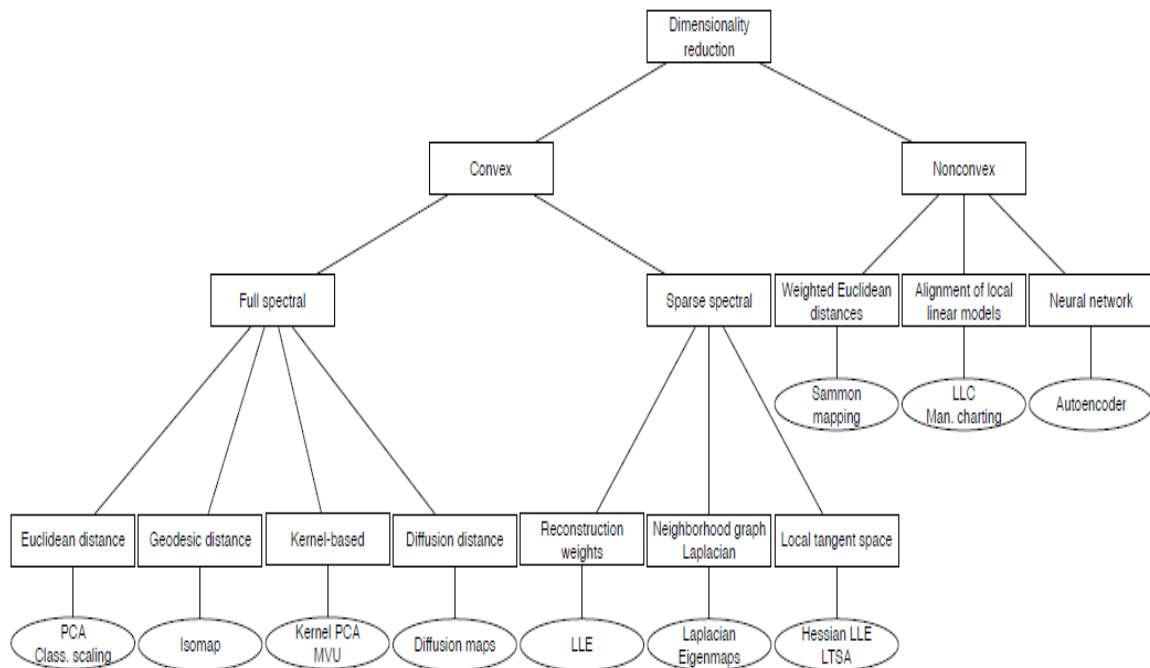
Comme le choix des variables, cette stratégie vise à exprimer des exemples à travers un plus petit nombre d'attributs. Cependant, chacun de ces attributs peut être une fonction non linéaire de toutes les variables de l'exemple d'origine ; par exemple, le produit des variables. De plus, contrairement aux techniques de sélection de variables, les algorithmes de réduction de dimensionnalité ne sont généralement pas non supervisés; ils se basent simplement sur la distribution spatiale des exemples pour déterminer quels attributs sont pertinents.

Ils supposent que les données sont sur une variété (pas nécessairement linéaire) de dimensions plus petites. Le but est de trouver le changement de coordonnées qui permet d'obtenir les coordonnées le long de la variété. Nous verrons ces techniques plus loin dans ce document. Algorithmes de réduction dimensionnalité, que nous décrirons **au section X**, permettra de découvrir un tel système coordonnées en effectuant une décomposition en vecteurs propres et valeurs propres une matrice de similarité sur les données. Le résultat ainsi obtenu est généralement optimal dans la mesure où les vecteurs propres sont des vecteurs qui, projetés sur le sous-espace qu'ils génèrent, donnent la meilleure approximation possible en termes d'erreur quadratique moyenne (DIAMANTRAS et KUNG 1996).

Notons que ce résultat est optimal pour les données d'apprentissage et ne garantit pas la généralisabilité. Cependant, nous savons que nos estimations ont tendance à croître de manière asymptotique dans le sens de l'estimation des problèmes de continuum par rapport à la densité de données réelle (SHAW ETAYLOR et W ; LL ; AMs2003).

L'objectif de réduire la dimensionnalité des données (c'est-à-dire le nombre de variables utilisées pour les représenter) est :

- ✓ Faciliter la visualisation des données ;
- ✓ Réduire les coûts de calcul, de stockage et d'acquisition des données ;
- ✓ Améliorer l'apprentissage en construisant des modèles moins complexes, en éliminant les variables non pertinentes qui pourraient fausser les prédictions et enfin en réduisant le problème du fléau de la dimensionnalité.



**Figure 1.1. Taxonomie des techniques de réduction de dimensionnalité**

La figure 1.1 présente une taxonomie des techniques de réduction de la dimensionnalité. Dans [] les techniques de réduction de la dimensionnalité sont subdivisées en techniques convexes et non convexes. Les techniques convexes optimisent une fonction objective qui ne contient pas d'optima locaux, tandis que les techniques non convexes optimisent des fonctions objectives qui contiennent des optima locaux. Pour plus de détails sur les autres subdivisions de la taxonomie.

### **a- Réduction linéaire de dimensionnalité**

Supposons que les données se retrouvent sur une variété linéaire de plus faible dimension. Pour obtenir le nouveau système de coordonnées, il suffirait de translater la variété de sorte qu'elle passe par l'origine et de prendre comme nouvel espace le sous-espace généré par toutes les données. Cette stratégie ne fonctionne évidemment pas si les points sont bruités, ce qui est le cas pour pratiquement toutes les vraies données.

Pour réduire la dimension, on peut agir sur deux leviers :

- Supprimer des dimensions
- Combiner les variables afin d'obtenir un plus petit nombre de nouvelles variables plus expressives et/ou moins redondantes.

Les méthodes les plus connues sont :

- Analyse des composants principaux (ACP) :
- **Multi-DimensionalScaling (mds) :**
- Recherche une représentation en basse dimension des données dans laquelle les distances respectent bien les distances dans l'espace original de grande dimension, MDS part d'une notion de distance ou d'affinité  $K$  qui est calculée entre chaque paire d'exemples d'apprentissage.

- **ISOMAP** *Isometric Mapping* :

- Isomap recherche un encastrement de dimension inférieure qui maintient les distances géodésiques entre tous les points, généralise MDS aux variétés non linéaires.

- **Qu'est-ce que ISOMAP ?**

- Une technique de réduction de dimensionnalité non-linéaire qui tient compte de l'ensemble des données, calcule la distance à des voisins (K-plus proche ou radius) et calcule par la suite les plus courts chemins de tout les points à tout les points; une extension de MDS, mais linéarité locale seulement

### ISOMAP: Algorithme

1-Construit un graphe de voisinage:

- On définit un Graph G sur tout les points en connectant les points i et j si (mesuré par  $d_x(i,j)$ ) ils sont plus proche que  $\epsilon$  ( $\epsilon$ -Isomap) ou si i est un k plus proche voisin (k-Isomap). On la longueur des arc est  $d_x(i,j)$

2-Calcule les plus courts chemins:

- Initialise  $d_G(i,j)=d_x(i,j)$  si i et j partagent un arc,  $\infty$  sinon. Pour tout k à n remplace tout les  $d_G(i,j)$  par  $\min\{d_G(i,j), d_G(i,k)+d_G(k,j)\}$ , G contient les plus courts chemins point à point.

3-Construit la réduction en d-dimension

- Soit  $l_p$  est le p-ième eigenvalue (en ordre décroissant)  $t(D_G)$  et  $v_p^i$  le i-ème élément du p-ième eigenvector; On donne au p-ème composante du vecteur  $y_i$  de coordonnées en (petit)d-dimension la valeur:  $\sqrt{l_p v_p^i}$ . (MDS)
- Fonction de coût, En reconstruisant on minimise:

$$E = \|\tau(D_G) - \tau(D_Y)\|_{L^2}$$

### plongement localement linéaire (locally linear embedding) :

Récemment, nous avons introduit une méthode de vecteur propre plus puissante, appelée plongement localement linéaire. (LLE)—pour le problème de la réduction de dimensionnalité non linéaire (Roweis et Saul, 2000). Ce problème est illustrée par les variétés de la **figure 1**. Dans ces exemples, la réduction de dimensionnalité par LLE réussit à récupérer les variétés sous-jacentes, alors que les plongements linéaires par carte ACP ou MDS les données éloignées pointent vers des points proches dans le plan, créant des distorsions à la fois locales et globales géométrie. Comme ACP et MDS, notre algorithme est simple à mettre en œuvre, et ses optimisations ne impliquent des minima locaux. Contrairement à ces méthodes, cependant, il est capable de générer des données hautement non linéaires. plongements, et sa principale optimisation implique un problème de valeurs propres éparées qui s'adapte bien à de grands ensembles de données de grande dimension.

### ISOMAP vs LLE :

- Utilisent tous deux en eigensolver pour la matrice calculé
- ISOMAP calcule toutes les distances

- LLE ne regarde que localement plus rapide mais moins précis
- ISOMAP plus lent mais converge dans plus de cas pcqprecedent
- Approches similaires LLE local, ISOMAP global

### 1. Travaux connexes

Depuis le début de la photographie, la création de photos ayant un champ de vision plus large, en combinant plusieurs images est couramment utilisée pour augmenter le champ de vision, car le champ de vision de la caméra est assez limité et plus petit que le champ de vision humain. L'utilisation d'une caméra omnidirectionnelle peut fournir une solution partielle avec une résolution limitée. De plus, il est impossible de capturer une grande scène telle qu'une photographie aérienne dans une seule image. La solution est de fusionner les images. Un mosaïquage d'images est une collection d'images prises sous différents angles et déplacées vers le même point de référence. L'objectif est d'élargir le champ de vision limité du système d'imagerie. La construction de la mosaïque est un sujet important dans le domaine de la vision par ordinateur. Cette importance est évidente dans de nombreux cas où une image panoramique doit être composée de plusieurs images différentes [12].

L'idée de base de la plupart des algorithmes décrits dans la littérature est de faire correspondre les images en fonction de leurs propriétés radiométriques et géométriques en utilisant une fonction spécifique pour évaluer la qualité de la correspondance [13].

A l'ère du multimédia, l'image est de plus en plus présente dans divers domaines. Ces dernières années, la construction de bases de données est un domaine où les images sont largement utilisées. À mesure que l'utilisation des données d'image augmente, le besoin de rechercher et de faire correspondre des images de manière efficace et précise à partir de grandes bases de données d'images ou de groupes de bases de données a augmenté. [13].

Pour de nombreux problèmes de vision par ordinateur et d'apprentissage automatique, un bon nombre d'ensembles d'apprentissage est la clé d'une bonne performance[13].

Cependant, la partie la plus coûteuse en termes de calcul de nombreux algorithmes de vision par ordinateur et d'apprentissage automatique consiste à trouver les correspondances les plus proches des vecteurs de haute dimension qui représentent les données d'apprentissage. De toute évidence, la meilleure solution est de dériver une fonction de mappage de paramètres explicite entre l'espace d'échantillon de haute dimension et l'espace de coordonnées de basse dimension [1], [13].

Au début, toutes les méthodes de mosaïquage d'images supposaient l'existence d'un plan de référence commun pour afficher le panorama et d'une étape de recalage géométrique efficace pour assembler les patches d'images[23].

Le développement le plus important dans le domaine de la mosaïque d'images ces dernières années est la technologie du polymorphisme et ses variations [23]. La méthode de mosaïque de collection de base a été proposée et développée par Rousso, Peleg et al [22]. Dans [21], Hongyan Zhang et al ont introduit un cadre théorique pour modéliser certains puzzles de scènes avec des structures topologiques arbitraires.

Notre objectif est de trouver toutes les images correspondantes. Les ensembles connectés d'images correspondantes deviendront plus tard des panoramas. Puisque chaque image peut potentiellement correspondre à toutes les autres, ce problème semble d'abord être quadratique par rapport au nombre d'images.

Cependant, il est seulement nécessaire de faire correspondre chaque image à un petit nombre d'images se chevauchant afin d'obtenir une bonne solution pour la géométrie de l'image. À partir de l'étape de correspondance des caractéristiques, nous avons identifié les images qui ont un grand nombre de correspondances entre elles. Notre objectif est de créer à partir de la collection d'images, une série d'images qui contiennent une meilleure description de la scène.

Les techniques de sous-espace d'image est une classe spéciale de transformations d'image qui projette l'image d'entrée dans un sous-espace plus petit. L'idée de la transformation du sous-espace est de projeter une image d'entrée de haute dimension dans un espace de petite dimension.

Considérons une image d'entrée  $A$  de  $M \times N$  dimension. L'image peut être considérée comme un vecteur unidimensionnel  $a$  de dimension  $M \times N$ , ou de manière équivalente, un point dans un espace de dimension  $M \times N$ . Cependant, la plupart des images naturelles n'occupent qu'une petite partie de cet espace. Dans ce cas, nous devrions être en mesure de trouver un sous-espace de faible dimension dans lequel réside une image naturelle donnée.

Les principales raisons d'utiliser un sous-espace de faible dimensionnalité sont :

- **Visualisation** : Une réduction à un sous-espace de faible dimension permet de comprendre la structure intrinsèque des données d'entrée.
- La **généralisation** : Une représentation à faible dimension permet une meilleure généralisation.
- **Calcul** : en général, la manipulation de données à faible dimension est à la fois plus rapide et nécessite moins de mémoire que les manipulations correspondantes de données à haute dimension.
- **Modèle** : La représentation à basse dimension peut être utilisée comme un modèle à part entière.

## 2. Apprentissage de variété

Les ensembles de données à haute dimension peuvent être très difficiles à visualiser. Alors que les données en deux ou trois dimensions peuvent être tracées pour montrer la structure inhérente des données, les tracés équivalents en haute dimension sont beaucoup moins intuitifs. Pour faciliter la visualisation de la structure d'un ensemble de données, la dimension doit être réduite d'une manière ou d'une autre.

La façon la plus simple d'accomplir cette réduction de la dimensionnalité est de faire une projection aléatoire des données. Bien que cela permette un certain degré de visualisation de la structure des données, le caractère aléatoire du choix laisse beaucoup à désirer. Dans une projection aléatoire, il est probable que la structure la plus intéressante des données soit perdue.

Apprentissage de variété est une approche récente et populaire de la réduction non linéaire de la dimensionnalité. Les algorithmes pour cette tâche sont basés sur l'idée que la dimensionnalité de nombreux ensembles de données n'est qu'artificiellement élevée; bien que chaque point de données soit constitué de milliers de caractéristiques, il peut être décrit comme une fonction de quelques paramètres sous-jacents seulement.

Pour répondre à cette préoccupation, un certain nombre de systèmes de réduction de la dimensionnalité linéaire supervisés et non supervisés ont été illustrés

### Méthodes Manifold Learning

Les méthodes d'apprentissage de variétés sont l'une des évolutions les plus intéressantes de ces dernières années dans le domaine de l'apprentissage automatique. L'idée principale qui sous-tend ces méthodes est que, bien que les données réelles soient généralement représentées dans des espaces à très haute dimension, le processus générant les données est souvent considéré comme ayant relativement peu de degrés de liberté. Une caractérisation mathématique de cette intuition consiste à modéliser les données comme étant situées sur ou près d'un collecteur de faible dimension.

1. Méthodes de projection,
2. Méthodes génératives,
3. Méthodes d'intégration,
4. Méthodes d'information mutuelle.

#### **A - Méthodes de projection :**

- La première consiste à trouver les surfaces principales passant par le milieu des données, telles que les courbes principales [25]. Bien que géométriquement intuitif, le premier a du mal à généraliser la variable globale - paramètre de longueur d'arc - à une surface de plus grande dimension.

#### **B- Méthodes génératives :**

- adopte des modèles de topologie générative [26] [24], et émet l'hypothèse que les données observées sont générées à partir des nœuds latents de faible dimension régulièrement espacés. Et puis la relation de cartographie entre l'espace d'observation et l'espace latent peut être modélisée. Résultat de l'insuffisance inhérente des algorithmes EM (Expectation Maximisation) adoptés, les modèles génératifs tombent néanmoins facilement dans le minimum local et ont des taux de convergence lents.

#### **C- Méthodes d'intégration :**

- Divisé en algorithmes d'intégration globaux et locaux. ISOMAP [29], en tant qu'algorithme global, suppose que les propriétés isométriques doivent être préservées à la fois dans l'espace d'observation et dans l'espace d'enfouissement intrinsèque au sens affine. Et les extensions aux mappages conformes sont également discutées dans [32] [27]. D'autre part, l'inclusion localement linéaire (LLE) [35] [30] [28] et LaplacianEigensamp [31] se concentrent sur la préservation de la structure du voisin local.

#### **D- Méthodes d'information mutuelle**

- on suppose que l'information mutuelle est une mesure des différences de distribution de probabilité entre l'espace observé et l'espace embarqué, comme dans le voisinage stochastique le plus proche (désormais SNE) [33] et la cartographie des variétés [34].



Dans cette thèse, nous avons proposé une démarche basée d'un algorithme d'apprentissage de type variété utilisant la distance du plus court chemin et la sélection du voisinage conjoint.

Plus précisément, un graphe conjoint des plus proches voisins est construit, puis la distance du plus court chemin dans chaque modalité est calculée à partir du graphe conjoint de voisinage, suivi de l'intégration spatiale et la mise en correspondance à faible dimension.

### Apprentissage non supervisé de Variétés

Le défi des prochaines années est de permettre aux machines d'apprendre à partir de données brutes, non étiquetées, telles que des images, des vidéos ou des textes. C'est ce qu'on appelle l'apprentissage non supervisé, qui caractérise la distribution des données et les relations entre les variables, sans faire de distinction entre les variables observées et les variables prédites [18,4]. Cette présentation suppose une certaine familiarité avec les concepts de base de l'apprentissage profond. Les principales formes d'apprentissage non supervisé sont :

- La fonction d'estimation de la densité ou fonction de probabilité. Il s'agit de la forme la plus générale d'apprentissage non supervisé.
- La découverte de classes naturelles, ou le clustering (par exemple, l'algorithme K-means) qui cherche à découvrir les principaux modes de distribution, les "prototypes" les principales catégories, etc. Cela donne une forme de réduction de la dimensionnalité
- Trouver une base à faible dimension pour décrire des données à forte dimension. Cela fournit une représentation de faible dimension des données, qui peut être une étape importante pour visualiser les données et / ou comme un prétraitement avant l'apprentissage supervisé.

L'objectif de l'apprentissage non supervisé de caractéristiques est souvent de découvrir des caractéristiques à faible dimension qui capturent une certaine structure sous-jacente aux données d'entrée à haute dimension. L'extraction d'informations pertinentes à partir d'une grande quantité de données est un défi qui est à l'origine d'importants et récents développements scientifiques en statistique et en apprentissage automatique, et de l'application de ces nouveaux outils d'analyse de données [4]. Au cours des dernières années, la communauté de l'apprentissage automatique a développé de nombreux outils de réduction de la dimensionnalité non linéaire. Cependant, la diversité des méthodes reflète la diversité des critères de qualité utilisés à la fois pour optimiser les algorithmes et pour évaluer leurs performances.

### 1.1 Les méthodes de plongement

Pour effectuer une réduction de dimensionnalité des données, on va supposer que les exemples se trouvent sur une variété de plus faible dimension.

On va chercher à caractériser cette variété

Ensuite, on va essayer de caractériser cette variété. Plus précisément, on cherche de trouver un nouveau système de coordonnées, où chaque axe correspond à un déplacement le long d'une des dimensions de la variété.

Les méthodes **de plongement** sont divisées en algorithmes **de plongement** globaux et locaux. ISOMAP [1], en tant qu'algorithme global, présume que les propriétés isométriques doivent être

préservées à la fois dans l'espace d'observation et dans l'espace d'incorporation intrinsèque au sens affine. Les extensions aux mappings conformes sont également discutées dans [5][4]. D'autre part, l'incorporation localement linéaire (LLE) [6][3][2]. Nous allons ici présenter le concept de base d'isomap et de LLE.

### **3.1.1 Isomap**

La technique de réduction non linéaire de la dimensionnalité Isomap donne, pour de nombreuses scènes intéressantes, une représentation de très faible dimensionnalité de l'espace des images possibles. L'essence de l'approche isomap de la NLDR réside dans la nouvelle formulation du problème : rechercher un encastrement euclidien de faible dimension d'un ensemble d'observations qui capture leurs similarités intrinsèques, telles que mesurées le long du chemin géodésique du collecteur d'observation [1,5].

Isomap apprend une représentation des observations perceptives dans laquelle il est facile d'effectuer des interpolations et d'autres informations complexes en faisant correspondre des géodésiques à des lignes droites dans l'espace des caractéristiques.

L'algorithme ISOMAP tente de récupérer l'intégration originale des données cachées  $\{y_i\}$ . Nous décrivons brièvement la procédure standard de l'algorithme Isomap [7].

1. Déterminer un graphe de voisinage  $G$  des données observées  $\{X_i\}$
2. Calculer les plus courts chemins dans le graphe de toutes les paires de points de données.
3. Appliquer le MDS à la matrice de distance de plus court chemin  $D$  résultante, pour trouver le point de données reconstruit en  $\{y_i\} R^d$

## **1. MPEG-7 Standard**

MPEG-7 est une norme ISO/CEI développée par MPEG (Moving Picture Experts Group), et elle est conçue pour décrire le contenu multimédia. Elle est formellement appelée "Multimedia Content Description Interface". Elle intègre des représentations de contenu efficaces pour la recherche par similarité [15].

Les outils de description visuelle MPEG-7 se composent de structures et de descripteurs de base qui couvrent les caractéristiques visuelles de base suivantes pour le contenu multimédia: couleur, texture, forme, mouvement, localisation et reconnaissance des visages. Parmi ces descripteurs, on peut citer :

- Descripteurs de couleur : Le descripteur de structure de couleur (CSD) représente une image à la fois par la distribution des couleurs et par la structure spatiale de la couleur.
- Le descripteur de couleur échelonnable (SCD) est un codage basé sur la transformée de Haar d'un histogramme de couleur dans l'espace couleur HSV.
- Dominant ColorDescriptor (DCD) spécifie jusqu'à huit couleurs représentatives (dominantes) dans une image ou une région d'image.

## *Chapitre 1 Etat de l'art*

---

- ColorLayoutDescriptor (CLD) est un descripteur de couleur compact, invariant en termes de résolution, qui représente efficacement la distribution spatiale des couleurs.
  - Le descripteur Group of Frames or Group of Pictures (GoF/GoP) est utilisé pour les caractéristiques basées sur la couleur de plusieurs images ou de plusieurs images dans un segment vidéo. Il s'agit d'une alternative à la représentation des segments vidéo basée sur une seule image clé. Le descripteur est obtenu en agrégeant les histogrammes de plusieurs images ou champs et en représentant l'histogramme final avec un algorithme de couleur évolutif. l'histogramme final avec un descripteur de couleur évolutif.
  - Descripteurs de texture : Le descripteur d'histogramme de bord (EHD) décrit la distribution spatiale des bords dans une image.
  - Le descripteur de texture homogène (HTD) caractérise la texture d'une région en utilisant région en utilisant l'énergie moyenne et l'écart d'énergie d'un ensemble de canaux de fréquence, qui sont modélisés par des fonctions de Gabor.
  - Le descripteur de navigation de texture (TBD) caractérise les textures perceptivement en termes de régularité, de finesse et de directionnalité
- Dans cette mémoire, nous utilisons une description efficace des caractéristiques locales de la couleur appelée ColorLayout et le descripteur d'histogramme de bord (EHD), qui ont été acceptés dans le projet final du comité MPEG-7[37].

*Chapitre 2*  
*Conception*

## **2. Introduction**

L'être humain reçoit la grande majorité des informations sur son environnement par la vision. La vision est également la clé de la construction de systèmes artificiels capables de percevoir et de comprendre leur environnement. En raison de ses nombreuses applications et de ses grands défis de recherche, la vision par ordinateur est l'un des domaines de recherche les plus actifs dans le domaine des technologies de l'information. Considérant le problème à affronter. Une grande quantité d'informations qui doivent être traitées dans un temps raisonnable et avec précision, ce processus nécessite de fournir une base de données structurée, donc notre première étape est de concevoir et de créer une telle BD images qui est basée sur la meilleure façon de la décrire.

Ces dernières années, les approches permettant de décrire efficacement le contenu des images ont suscité un grand intérêt dans la recherche sur la vision par ordinateur. Les descripteurs d'images jouent un rôle crucial dans la plupart des systèmes et applications de vision par ordinateur.

La fonction des descripteurs est de convertir les informations au niveau du pixel en une forme utile qui capture les facteurs les plus importants de la scène imagée, mais qui est insensible aux aspects non pertinents causés par l'environnement variable. Un descripteur efficace est capable d'ignorer les aspects non pertinents causés par les changements de l'environnement. Les descripteurs MPEG-7 de bas niveau se sont avérés efficaces, cohérents avec ce que nous avons observé, et donc largement utilisés par les chercheurs en vision par ordinateur, en reconnaissance des formes et en multimédia. Nous allons nous intéresser à des mesures de distance autres que celles recommandées par MPEG-7[37].

### 3. APPROCHE PROPOSEE

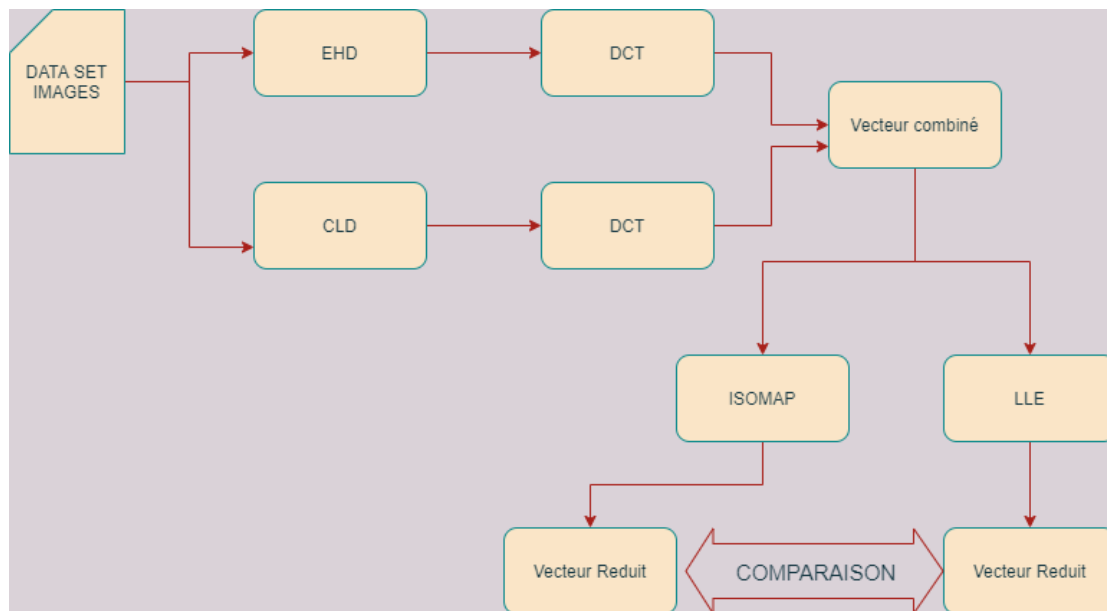


Figure 2.1. : Schéma général de l'approche proposée

### 4. Correspondance de similarité en combinant EHD et CLD

Le processus de mise en correspondance de deux ou plusieurs images capturées à partir de points de vue différents, à des moments différents et dans des conditions environnementales différentes (éclairage, capteurs différents) est appelé recalage d'images. Une fois que la correspondance des caractéristiques a été établie et que l'alignement géométrique a été effectué, les images sont combinées pour fournir une représentation de la scène qui est à la fois géométriquement et photométriquement cohérente. Le recalage d'images est un problème clé de la vision par ordinateur qui apparaît dans une grande variété d'applications telles que les mosaïques d'images, l'analyse d'images médicales, le suivi des visages, la reconnaissance de l'écriture manuscrite, la correspondance stéréo et l'analyse du mouvement. (Taylor et Bhusnurmath, 2008) [14].

#### 4.1 Descripteurs des données générales

##### a) Descripteur de disposition des couleurs (CLD) :

- Ce descripteur spécifie la distribution spatiale des couleurs, dans une image ou une région arbitraire ; il recommande d'utiliser l'espace couleur YCbCr pour le CLD.
- Le processus d'extraction des descripteurs de disposition des couleurs comprend quatre étapes : Le partitionnement de l'image, la sélection de la couleur dominante, la transformation DCT et la quantification non linéaire des coefficients DCT balayés en zigzag (voir figure).

Dans la première étape, une image d'entrée est divisée en 64 blocs. La taille de chaque bloc est  $W/8 \times H/8$ , où  $W$  et  $H$  désignent respectivement la largeur et la hauteur de l'image d'entrée.

Dans la deuxième étape, une seule couleur dominante est sélectionnée dans chaque bloc pour construire une petite image de taille  $8 \times 8$ . Toute méthode de sélection de couleur dominante peut être appliquée. Nous utilisons des couleurs moyennes simples comme couleurs dominantes.

Dans la troisième étape, chacune des trois composantes ( $Y$ ,  $Cb$  et  $Cr$ ) est transformée par DCT  $8 \times 8$ , et nous obtenons trois ensembles de coefficients DCT. Certains coefficients de basse fréquence sont extraits à l'aide d'un balayage en zigzag et quantifiés pour former le descripteur de disposition des couleurs d'une image fixe [24].

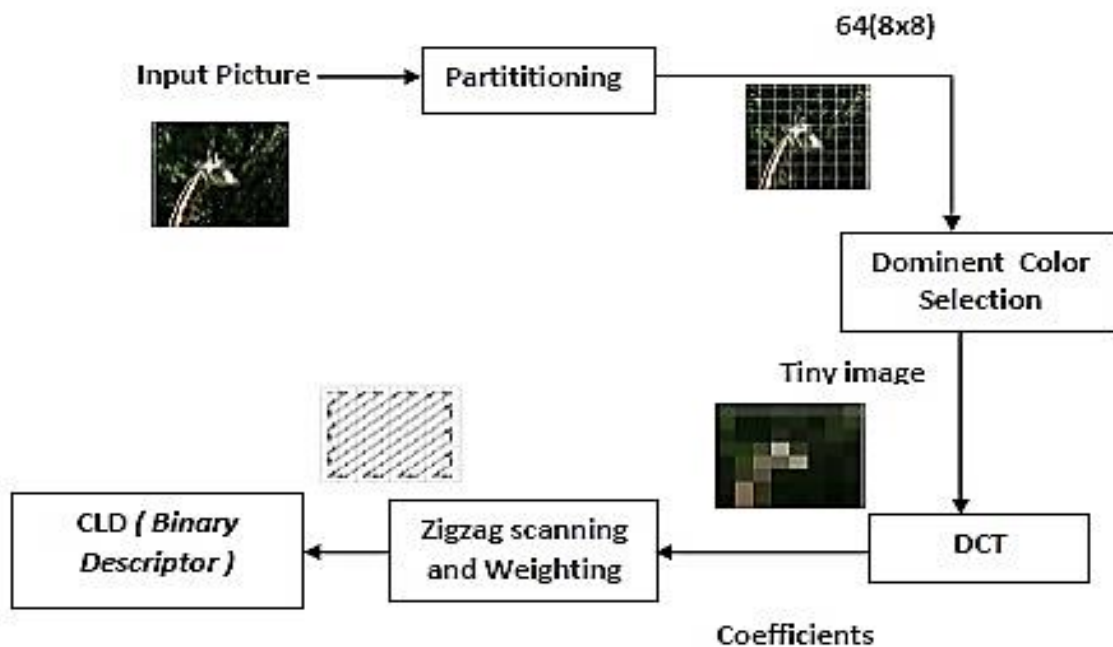


Figure 2.2. Schéma fonctionnel du processus d'extraction du descripteur de disposition des couleurs.

**Processus d'appariement (The matching process)**

Le processus de mise en correspondance évalue si deux éléments sont égaux en comparant les deux éléments et en calculant la distance qui les sépare. Dans le cas des descripteurs de couleur, le processus de mise en correspondance évalue si deux images sont similaires.

Pour faire correspondre deux CLD,  $\{ DY, DCr, DCb \}$  et  $\{ DY', DCr', DCb' \}$ , la mesure de distance suivante peut être calculée comme suit :

$$D = \sqrt{\sum_i w_{yi} (DY_i - DY'_i)^2} + \sqrt{\sum_i w_{bi} (DCb_i - DCb'_i)^2} + \sqrt{\sum_i w_{ri} (DCr_i - DCr'_i)^2}$$

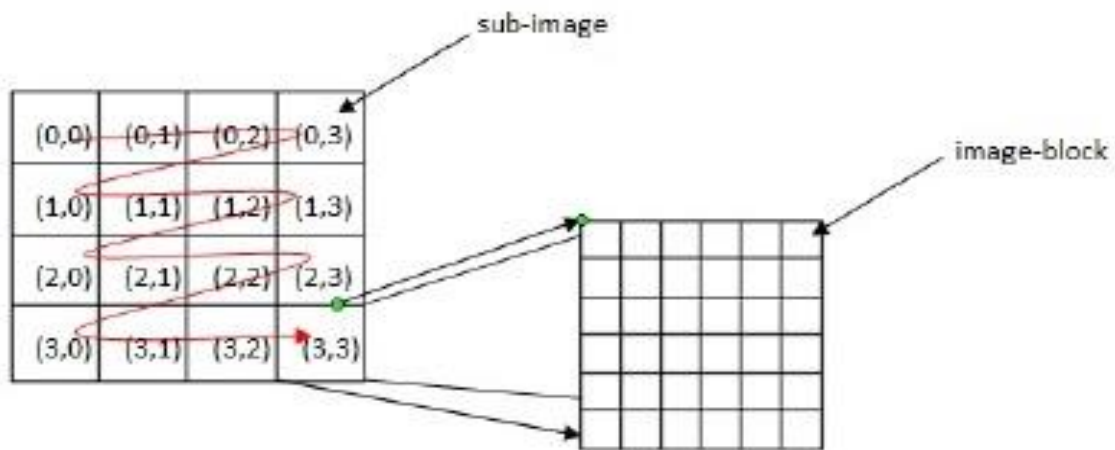
/Où (  $w_{yi}$  ,  $w_{bi}$  ,  $w_{ri}$  ) représentent les ième coefficients DCT des composantes de couleur représentatives. Les distances sont pondérées de manière appropriée, des poids plus importants étant attribués aux composantes de fréquence inférieure.

- 2 images sont identiques si la distance est de 0
- 2 images sont similaires si la distance est proche de 0

Par conséquent, ce processus de mise en correspondance permettra d'identifier les images dont les descripteurs de couleur sont similaires. Comme la complexité du processus de mise en correspondance par similarité présenté ci-dessus est peu élevée, il est possible d'effectuer une mise en correspondance d'images à grande vitesse.

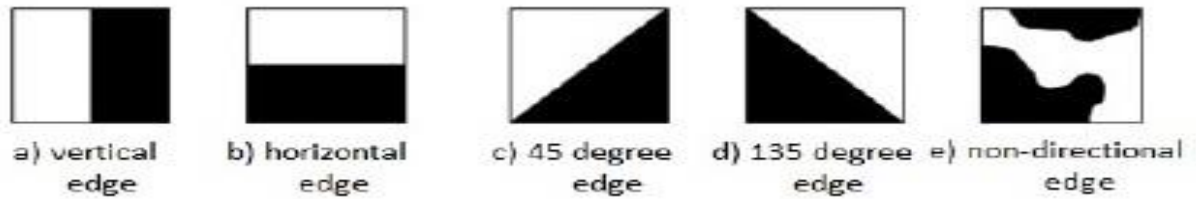
**b) Descripteur d'histogramme de bord**

- Le descripteur d'histogramme de bord (EHD) est l'une des méthodes largement utilisées pour la détection de forme. Il représente essentiellement la fréquence relative d'apparition de 5 types de bords dans chaque zone locale appelée sous-image ou bloc d'image. La sous-image est définie en partitionnant l'espace image en blocs 4x4 ne se chevauchant pas, comme le montre la figure 3. Ainsi, la partition de l'image crée définitivement 16 blocs de taille égale, quelle que soit la taille de l'image d'origine. Pour définir les caractéristiques du bloc image, nous générons ensuite un histogramme de distribution des contours pour chaque bloc image. Les bords du bloc d'image sont classés en 5 types : verticaux, horizontaux, diagonale à 45 degrés, diagonale à 135 degrés et bords non directionnels, comme le montre la figure 4. Ainsi, l'histogramme de chaque bloc d'image représente la distribution relative de les 5 types de bords dans la sous-image correspondante [16, 17] .



**Figure 2.3. Definition of Sub-image and Image-block in the EHD**





**Figure 2.4. Five Types of Edges in the EHD**

## 4.2 Démarche proposée pour la combinaison des deux valeurs IDCT de CLD et EHD :

### Rappels mathématiques :

Le passage par la DCT a été l'idée majeure pour la compression JPEG. En effet ce processus appartient à une classe d'opérations mathématiques, tout comme la Transformée de Fourier. Elle permet un changement de domaine d'étude, tout en gardant exactement la même fonction étudiée. Dans notre cas, on étudie une image, c'est à dire une fonction à 3 dimensions : X et Y, indiquant le pixel, et Z avec la valeur du pixel en ce point. Dans le cas d'une image couleur, il faut donc considérer indépendamment 3 fonctions, pour chacun des canaux RGB.

L'application de la DCT, ou d'une Transformée de Fourier fait passer l'information de l'image du domaine spatial en une représentation identique dans le domaine fréquentiel. Pourquoi ce changement de domaine est-il si intéressant? Justement parce qu'une image classique admet une grande continuité entre les valeurs des pixels. Les hautes fréquences étant réservées à des changements rapides d'intensité du pixel, ceux-ci sont en général minimales dans une image. Ainsi on parvient à représenter l'intégralité de l'information de l'image sur très peu de coefficients, correspondant à des fréquences plutôt basses. La composante continue (valeur moyenne de l'image traitée) ayant une grande importance pour l'oeil.

La DCT s'applique à une matrice carrée. Le résultat fournit est représenté dans une matrice de même dimension. Les basses fréquences se trouvant en haut à gauche de la matrice, et les hautes fréquences en bas à droite.

La transformation de la matrice DCT est orthogonale, elle est accompagnée d'une méthode d'inversion pour revenir au domaine spatial. Ainsi après avoir effectué des modifications dans le domaine fréquentiel, éliminant les variations de l'image presque invisibles à l'œil humain, on revient à une représentation sous forme de pixels.

Formule pour calculer la DCT sur une matrice NxN

$$DCT(i, j) = \frac{1}{\sqrt{2}} C(i) C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \text{pixel}(x, y) \cos\left(\frac{(2x+1)i\pi}{2N}\right) \cos\left(\frac{(2y+1)j\pi}{2N}\right)$$

$$C(x) = \frac{1}{\sqrt{2}}$$

si x vaut 0, et 1 si x > 0.

Formule pour calculer la IDCT sur une matrice NxN

$$\text{Pixel}(x, y) = \frac{1}{\sqrt{2N}} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} c(2i) c(2j) DCT(i, j) \cos\left(\frac{(2x+1)i\pi}{2N}\right) \cos\left(\frac{(2y+1)j\pi}{2N}\right)$$

$$C(x) = \frac{1}{\sqrt{2}}$$

si x vaut 0, et 1 si x > 0.

### 4.3. Pourquoi la DCT ?

Nous avons vu que la DCT fait partie de la même classe d'outils mathématiques que la transformée de Fourier. Alors pourquoi les membres du groupe JPEG ont-ils choisi la DCT ? Ces deux méthodes permettent une décomposition de l'information dans une autre base : Une base cosinus, ou la base de Fourier.

Cependant, cette décomposition en base de Fourier pose plusieurs problèmes : si l'image présente des discontinuités, alors la décroissance des coefficients de la transformée de Fourier n'est que de  $1/k$ ,  $k$  se référant à l'indice du coefficient..

Le fait de décomposer la fonction sur une base de cosinus fait que la fonction sera symétrique par rapport à  $(-\frac{1}{2})$ . Cependant la DCT pose un problème d'optimisation. En effet le calcul d'un coefficient nécessite  $N^2$  multiplications, or il y a  $N^2$  coefficients à calculer. Le coût d'une telle décomposition devient alors démesuré si notre image est de taille 512x512. Ainsi, au lieu de traiter toute l'image, on découpe celle-ci en blocs 8x8. Ce choix représente un compromis performance qualité : en effet, en augmentant la taille de ces blocs, la compression serait meilleure, mais le coût en temps a été jugé trop grand. Sur chacun de ces blocs, on procède ensuite à une DCT.

Cela permet d'avoir un algorithme rapide. Or le DCT par blocs 8x8 est justement un des facteurs limitants de la compression JPEG : en effet lorsqu'on augmente la compression, on voit apparaître ces blocs.

### 5. Algorithme de calcul pour la DCT

Une bonne façon d'implémenter la DCT en bloc serait de créer une matrice de la transformée en cosinus, C.

$$C_{i,j} = \begin{cases} \sqrt{\frac{2}{N}} \cos\left(\frac{(2j+1)i\pi}{2N}\right) & \text{si } i > 0 \\ \frac{1}{\sqrt{N}} & \text{si } i = 0 \end{cases}$$

Une fois cette matrice créée, on crée la matrice Ct qui correspond à la transposée de C. Ensuite la transformée par bloc 8x8 se réduit à deux multiplications **matricielles** :

$$\mathbf{DCT} = \mathbf{C} * \mathbf{PIXELS} * \mathbf{Ct}$$

Donc le coût pour calculer la DCT sur un bloc se réduira à  $2N$  multiplications sur des entiers, et  $2N$  additions sur des entiers : cela représente un gain important par rapport à une simple double boucle sur les indices comme pourrait le laisser suggérer la formule initiale de la DCT. On passe d'un coût de  $O(N^2)$  à un coût très intéressant de  $O(N)$  pour calculer UN bloc.

Spécifiquement, le descripteur de disposition de couleur et le descripteur d'histogramme de bord dans MPEG-7 sont utilisés pour extraire des caractéristiques d'image dans les sous-images choisies. Puisque les coefficients IDCT et EHD du CLD sont basés sur des grilles 8x8 et 4x4, respectivement

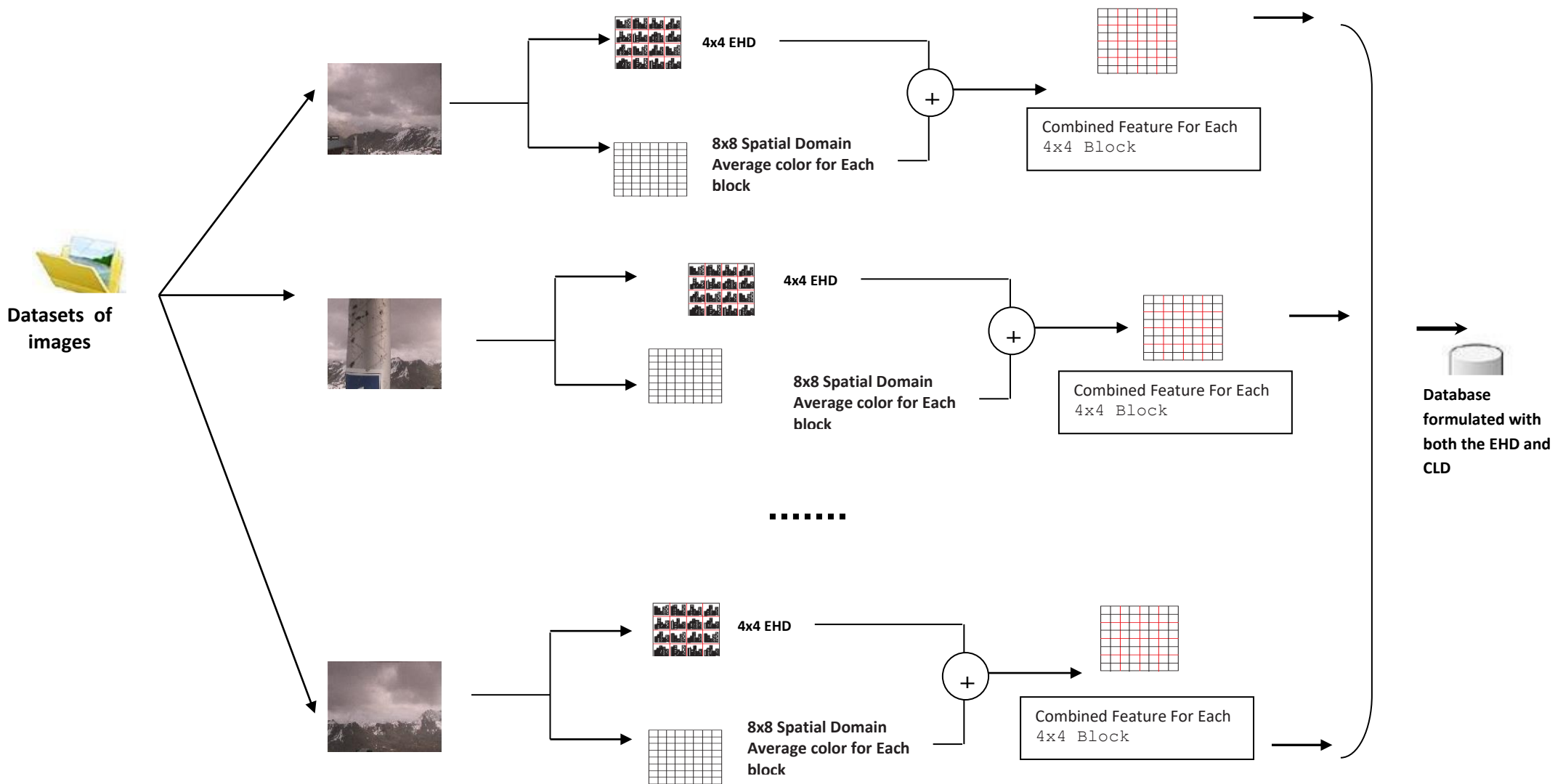


Figure 2.5. Le processus d'analyse des images de la base de données en combinant les deux valeurs IDCT de CLD et EHD

- La figure 5 montre qu'il existe une correspondance de localisation entre la sous-image de l'EHD et les valeurs IDCT 2x2 du CLD, et les valeurs IDCT 2x2 voisines correspondent à une sous-image de l'EHD.

C'est-à-dire,  $S_{ij}^{EHD}$ , la mesure de similarité de l'EHD entre la sous-image de la requête et Celle de l'image de test à  $(i, j) \in \Omega$ , est la suivante

$$S_{ij}^{EHD}(L) = \frac{1}{S_{max}^{EHD}} \sum_{k=1}^5 \left| E^{Q(k)} - E_{ij}^{T_l}(k) \right| \quad (1)$$

-Où les exposants  $Q$  et  $T_i$  représentent respectivement la requête  $L^{th}$  et les images de test.

De plus,  $S_{max}^{EHD}$  est une constante de normalisation telle que

$$S_{max}^{EHD} = \max_{\substack{l=1, \dots, L \\ 0 \leq i \leq 3, 0 \leq j \leq 3}} S_{ij}^{EHD}(l) \quad (2)$$

Où  $L$  est le nombre total d'images de test. Cela implique que  $S_{ij}^{EHD}(L)$  est normalisée telle que  $0 \leq S_{ij}^{EHD}(L) \leq 1$ . De même, nous avons la similitude de CLD comme suit

$$S_{ij}^{EHD}(L) = \frac{1}{3S_{ij}^{CLD}(L)} \left( \left| Y^Q - Y_{ij}^{T_i} \right| + \left| Cb^Q - Cb_{ij}^{T_i} \right| + \left| Cr^Q - Cr_{ij}^{T_i} \right| \right) \quad (3)$$

Enfin, en combinant les équations (1) et (3), nous avons une similarité combinée comme suit

$$S_{ij}(L) = 0.25 * S_{ij}^{CLD}(L) + 0.75 * S_{ij}^{EHD}(L) \quad (4)$$

Or, une valeur  $S_{ij}(L)$  plus petite signifie une plus grande similarité entre les deux régions.

Calculatrice (4) pour toutes les images de l'ensemble de données  $\Omega$ .

Après avoir combiné les valeurs IDCT CLD et EHD, nous mettrons en œuvre deux algorithmes de réduction de la dimension, ISOMAP et LLE, pour faire en sorte de réduire les vecteurs et d'assurer une bonne compression des ensembles de données ; à la fin, nous comparerons et choisirons le plus efficace en termes de temps d'exécution et de mémoire.

## 6. LLE-Algorithmme :

On explique cette algorithme comme suite :

### 6.1. Zoom sur LLE :

L'intégration linéaire locale (LLE) est une méthode associée conçue pour préserver la structure locale des données et résoudre le problème de la réduction de la dimensionnalité non linéaire. Le plongement résultant est optimisé pour préserver les configurations locales des plus proches voisins, tout en supposant une dépendance linéaire locale entre eux. Les définitions de « voisinage » de chaque point sont définies par l'utilisateur et peuvent inclure tous les points dont la distance du point spécifique est inférieure à une certaine valeur, un nombre fixe de points les plus proches ou toute autre décision raisonnable.

L'entrée de l'algorithme LLE classique est un ensemble de signaux multidimensionnels. L'algorithme tente de calculer un plongement de faible dimension avec la propriété que les points voisins dans l'espace de grande dimension restent proches et co-localisés de manière similaire les

uns par rapport aux autres dans l'espace de faible dimension. C'est-à-dire que l'intégration est optimisée pour préserver les configurations locales des voisins les plus proches. L'algorithme fonctionne sans recourir à des mesures de relation entre des points de données éloignés.

**Entrée X** : matrice D par N composée de N éléments de données dans D dimensions.

**Sortie Y** : matrice d par N constituée de  $d < D$  coordonnées d'insertion dimensionnelles pour les points d'entrée.

1. Trouver des voisins dans l'espace X [b,c].

pour  $i=1:N$

calculer la distance de  $X_i$  à tout autre point  $X_j$

trouver les K plus petites distances

assigner les points correspondants comme voisins de  $X_i$

fin.

2. Résoudre les poids de reconstruction W.

pour  $i=1:N$

créer une matrice Z composée de tous les voisins de  $X_i$  [d]

soustraire  $X_i$  de chaque colonne de Z

calculer la covariance locale  $C=Z'*Z$  [e]

résoudre le système linéaire  $C*w = 1$  pour w [f]

mettre  $W_{ij}=0$  si j n'est pas voisin de i

définir les éléments restants dans la ième ligne de W égaux à  $w/\text{sum}(w)$ ;

fin.

3. Calculez les coordonnées de plongement Y à l'aide des poids W.

créer une matrice creuse  $M = (I-W)*(I-W)$

trouver les vecteurs propres inférieurs d+1 de M

(correspondant aux d+1 plus petites valeurs propres)

définir la qième LIGNE de Y comme le q+1 plus petit vecteur propre

(rejeter le vecteur propre inférieur [1,1,1,1...] avec la valeur propre zéro)

## 6.2. Démarche proposée

L'algorithme LLE s'attend à ce que chaque point de données et ses voisins se trouvent sur un patch localement linéaire de la variété initiale ou proche de celui-ci. Ensuite, la caractérisation de la géométrie locale au voisinage de chaque point de données est effectuée par des coefficients linéaires qui reconstruisent le point de données à partir de ses voisins. Enfin, un encastrement dans un espace de faible dimension qui préserve cette caractérisation géométrique locale est recherché. Il y a les 3 étapes de mise en œuvre dans l'algorithme :

1. Recherche du voisin le plus proche. Les éléments non nuls de la matrice de poids sont identifiés.
2. Calcul des poids. Les poids qui reconstruisent chaque point de données à partir de ses voisins sont calculés à l'aide des moindres carrés contraints.
3. Calcul du plongement de faible dimension. Enfin, les plongements de faible dimension qui préservent le mieux les poids sont trouvés en résolvant un problème aux valeurs propres.

[a] Cela peut être fait de différentes manières, par exemple ci-dessus, nous calculons les  $K$  voisins les plus proches en utilisant la distance euclidienne. D'autres méthodes telles que epsilon-ball incluent tous les points dans un certain rayon ou des métriques de distance locales adaptatives et/ou spécifiques à un domaine plus sophistiquées.

[b] Même pour des règles de voisinage simples comme K-NN ou epsilon-ball utilisant la distance euclidienne, il existe des techniques très efficaces pour calculer les voisins de chaque point, comme les arbres KD.

[c]  $Z$  est constitué de toutes les colonnes de  $X$  correspondant aux voisins de  $X_i$  mais pas de  $X_i$  lui-même

## 7. Conclusion :

Dans ce chapitre, nous proposons notre approche pour résoudre le problème posé par les algorithmes d'apprentissage de variétés non linéaires tels que les algorithmes LLE et Isomap. Ces derniers présentent un intérêt particulier pour notre travail qui se fonde essentiellement sur l'analyse et la représentation de la série de données d'images de différentes catégories de scènes de paysage en utilisant une hybridation des valeurs CLD et EHD IDCT comme indicateur de distance de similarité robuste dans la base de données.

*Chapitre 3*  
*Implémentation*



### 1. Introduction

Dans ce chapitre, nous présenterons également les outils de développement que nous avons utilisés, l'architecture matérielle que nous avons mise en place, et enfin les tests effectués et les résultats obtenus.

### 2. Données expérimentales et mesures de qualité

Pour évaluer la performance de notre démarche, nous présentons l'utilisation d'algorithmes d'apprentissage de variétés non linéaire (LLE et Isomap) qui sont d'un intérêt particulier dans notre travail, ce qui nous permet d'analyser des ensembles de données d'images provenant de différentes catégories de scènes de paysage en utilisant une hybridation des valeurs IDCT CLD et EHD comme descripteur robuste de la similarité entre ces données.

Le système qui nous avons implémenté est exécuté dans MATLAB sur un ordinateur doté d'un processeur Intel Core 2.60 GHz et d'une capacité de mémoire de 4 Go. Le langage de programmation MatLAB a été utilisé pour le développement, car il est plus efficace en termes de temps de calcul et peut donc traiter de grandes quantités de données matricielles plus rapidement que d'autres langages. MATLAB est le langage le plus approprié lorsqu'il s'agit d'exprimer les mathématiques computationnelles [37].

La base de données est une collection d'images RVB 32 bits [0-255], provenant de plusieurs catégories de paysages différents. 217 images au total, de taille 640x480. Certaines d'entre elles sont des séquences vidéo.

#### a) **Bibliothèque d'extraction de caractéristiques MPEG-7 de bas niveau (statique/dynamique, C++) :**

Extraction de descripteurs MPEG-7 de bas niveau à partir d'images/cadres, de régions, de plusieurs cadres le cas échéant. Il compile à la fois sur Linux et Windows avec peu ou pas de changement. Descripteurs pris en charge : descripteur de structure de couleur (CSD), descripteur de couleur évolutif (SCD), descripteur de groupe de cadres d'images (GoF/GoP), descripteur de disposition des couleurs (CLD), descripteur de couleur dominante (DCD), descripteur de reconnaissance faciale (FRD) , Descripteur de texture homogène (HTD), Descripteur d'histogramme de bord (EHD), Descripteur de forme de région (RSD), Descripteur de forme de contour (CShD). [36]

#### b) **Bibliothèque d'extraction de caractéristiques MPEG-7 : [36]**

Cette bibliothèque est adaptée du logiciel de référence MPEG-7 XM pour la faire fonctionner avec les structures de données de la bibliothèque Open Source Computer Vision (OpenCV) (par exemple, IplImage, Mat). Il a une API très facile à utiliser. De plus, les bogues existants (entraînant des valeurs de descripteur incorrectes) dans le logiciel XM sont corrigés.

3. Résultats expérimentaux

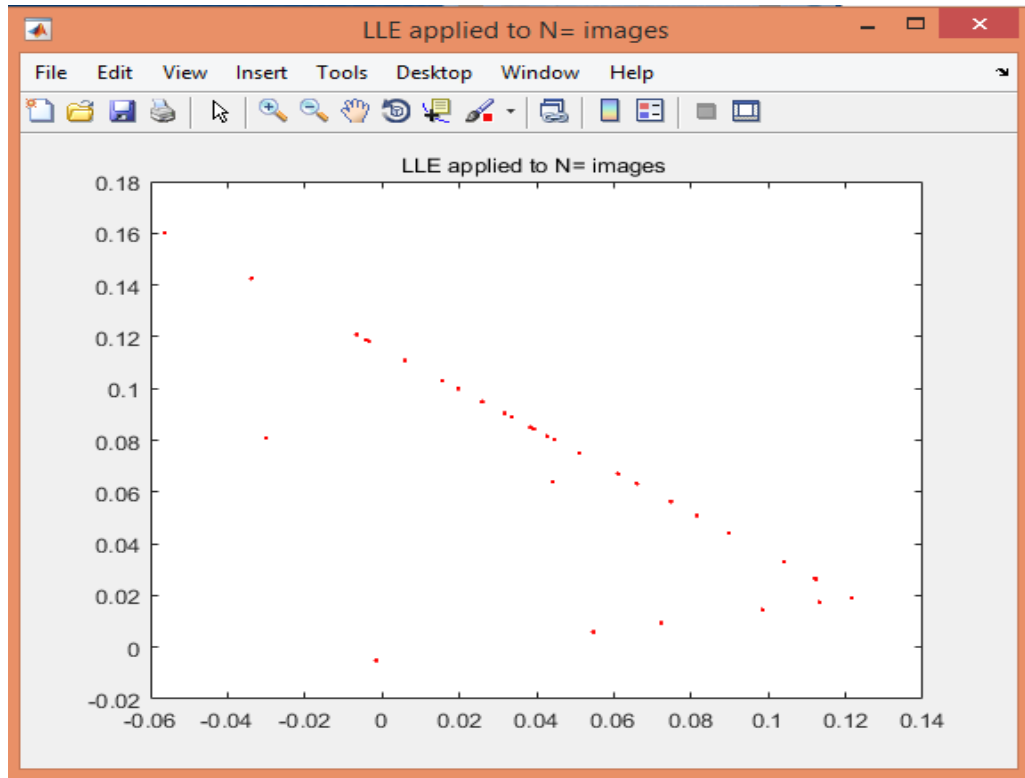


Figure 3.1. LLE avec hybridation=31 et k=3 voisins les plus proches

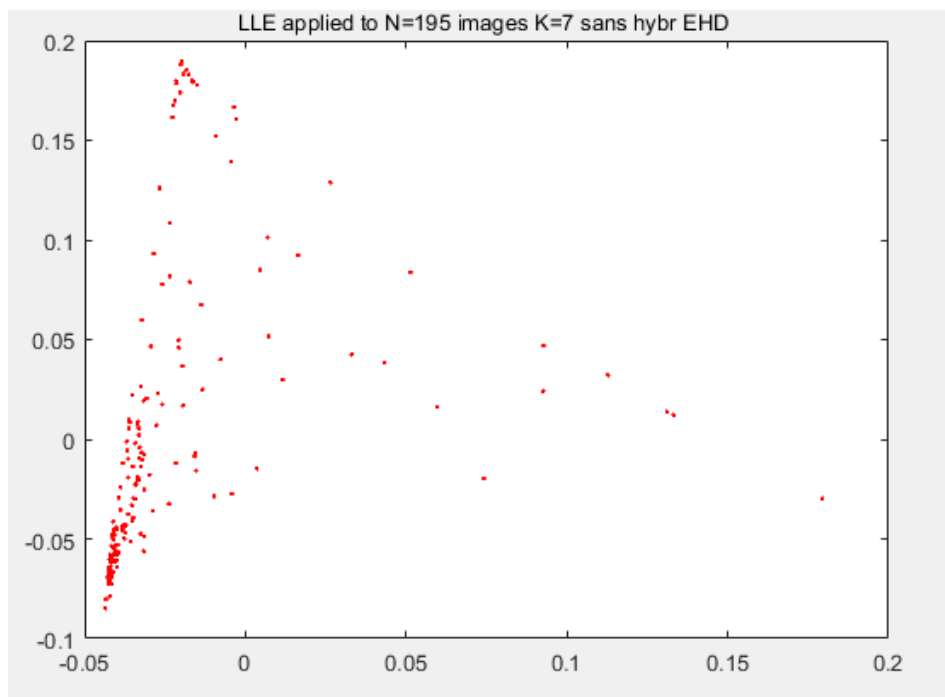


Figure 3.2. LLE appliqué à n=195 image k=7 sans hybridation EHD

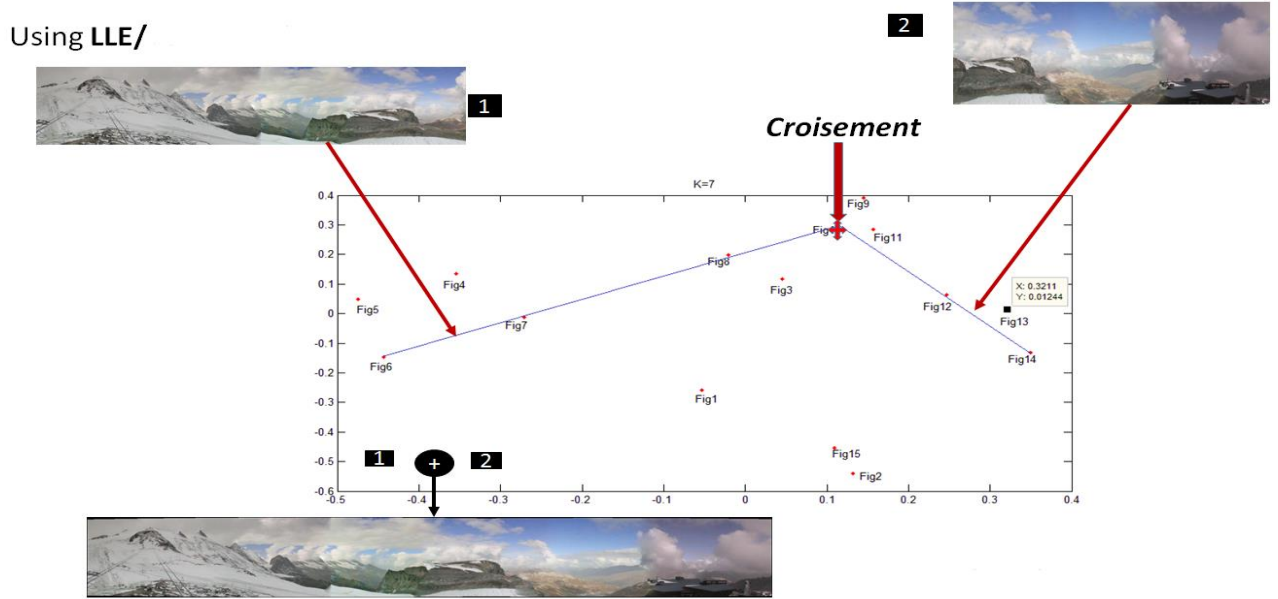


Figure 3.3. LLE appliqué à n=15 image k=7

Figure 3 : montre les résultats de LLE sur l'ensemble de données sur n = 195 images. Les images contiennent différentes catégories de paysage en utilisant k = 7 voisins les plus proches.

- Les coordonnées globales de LLE offrent un moyen simple d'analyser et de manipuler des observations de grande dimension en termes de degrés de liberté intrinsèques non linéaires. Pour un ensemble d'images de paysage, LLE détecte correctement

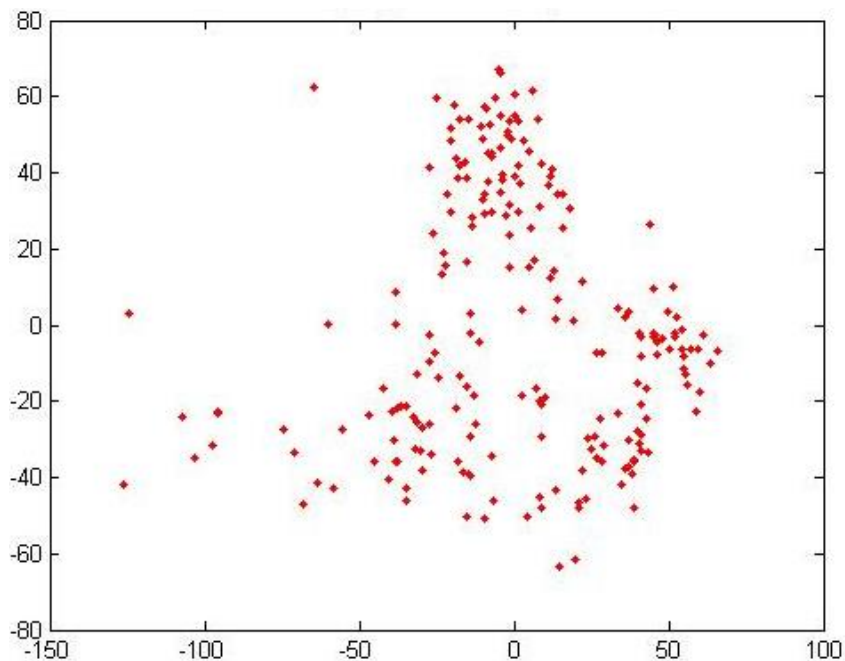
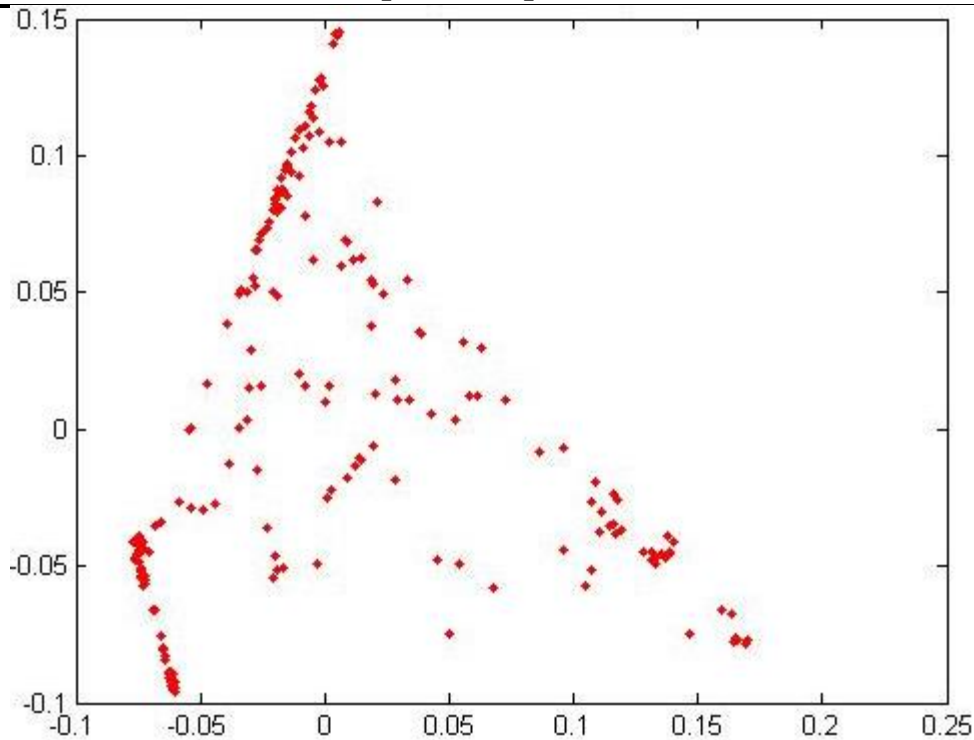


Figure 3.4. Incorporation bidimensionnelle de N=217 images de scènes de paysages obtenue par isomap en utilisant k=7 plus proches voisins.



**Figure 3.5.** Montrée l'intégration LLE 2-D de N=217 images de scènes de paysage en utilisant k=6 voisins les plus proches.

### Isomap vs LLE

Les figures 4 et 5 démontrent que Isomap et LLE ont une forme d'intégration presque différente pour un ensemble de données. Puisque LLE vise à préserver la relation de voisinage local, alors que Isomap vise à préserver la distance géodésique globale. LLE peut être instable, en fonction de la topologie des données et de notre objectif.

Généralement, Isomap est basé sur l'idée d'approximation des distances géodésiques par la construction de chemins linéaires par morceaux entre les échantillons. Les chemins sont construits en reliant les plus proches voisins, et la distance géodésique entre deux points est approximée sur les segments linéaires entre les plus proches voisins. Ainsi, isomap n'a besoin que des distances entre les points de données proches pour découvrir la structure des manifolds dans les ensembles de données. Le recours aux seules distances entre les plus proches voisins est important pour cet mémoire.

La cartographie isométrique des caractéristiques (Isomap) est l'une des techniques représentatives du manifold. Elle est intuitive, bien comprise et produit des résultats de cartographie raisonnables.

#### **4. Conclusion**

L'objectif de l'approche proposée est de fournir un moyen efficace de combiner des informations provenant de plusieurs espaces à haute dimension, et de réduire le coût du calcul des similarités.

Nous tentons de modéliser le problème de l'assemblage d'images pour combiner plusieurs ensembles d'images naturelles, où chaque ensemble contient des images appartenant à la même représentation de la scène par le biais de techniques de réduction de la dimensionnalité non linéaire.

Nous avons examiné deux cadres de réduction de la dimensionnalité non linéaire (Isomap et LLE) et les avons appliqués pour analyser la dimensionnalité intrinsèque des images dans l'ensemble de données en combinant les deux valeurs IDCT de CLD et EHD pour extraire les caractéristiques des images et trouver un indicateur de distance plus robuste de la similarité dans l'ensemble de données.

LLE et Isomap ont des propriétés différentes, comme discuté et comparé dans un grand nombre de travaux précédents. LLE vise à préserver la relation de voisinage local, tandis qu'Isomap vise à préserver la distance géodésique globale. Nous tendons, l'utilisation d'une des deux techniques dépend de la topologie des données et de notre objectif. Pour des recherches ultérieures, la méthode proposée serait généralisée à des cas plus complexes.

# *Conclusion Générale*

## **Conclusion Générale**

Dans le cadre de notre travail, nous avons assuré la bonne réduction de dimensionnalité et Fournit une base de données bien réduite, au domaine d'imagerie le traitement très compliqué c'est pour son obligée d'appliquer la bonne réduction pour faciliter les accès à la base de données (réduire la complexité).

## *Références Bibliographiques*



**References:**

- [1] Joshua B. Tenenbaum, Vin de Silva, John C. Langford,, A Global Geometric Framework for Nonlinear Dimensionality Reduction, R E P O R T S, www.sciencemag.org SCIENCE VOL 290 22 DECEMBER 2000.
- [2] S. T. Roweis, and K. S. Lawrance, (2000), "Nonlinear Dimensionality reduction by locally linear embedding," *Science*, 290, pp. 2323-2326
- [3] Sam T. Roweis and Lawrence K. Saul, Nonlinear Dimensionality Reduction by Locally Linear Embedding, R E P O R T S, www.sciencemag.org SCIENCE VOL 290 22 DECEMBER 2000
- [4] V. S.Silva, J. B. Tenenbaum, (2002), "Unsupervised Learning of curved manifolds", *Nonlinear Estimation and Classification*, Springer-Verlag, New York.
- [5] Y. Bengio, J-F. Paiement, and P. Vincent. Out-of-Sample Extensions for LLE, Isomap, MDS, Eigenmaps, and Spectral Clustering. Technical Report 1238, Departement d'informatique et recherche operationnelle, Universite de Montreal, 2003.
- [6]. Dick de Ridder, Robert P.W. Duin, "Locally Linear Embedding for Classification", technical reports, Delft University of Technology, The Netherlands
- [7] V. de Silva and J. B. Tenenbaum. Global versus local methods in nonlinear dimensionality reduction. *Neural Information Processing Systems 15 (NIPS'2002)*, pp. 705-712, 2003.
- [8] K. Fukunaga. *Introduction to Statistical Pattern Recognition*. Academic Press Professional, Inc., San Diego, CA, USA, 1990.
- [9] J. Venna. *Dimensionality reduction for visual exploration of similarity structures*. PhD thesis, Helsinki University of Technology, 2007.
- [10] L.K. Saul, K.Q. Weinberger, J.H. Ham, F. Sha, and D.D. Lee. Spectral methods for dimensionality reduction. In *Semisupervised Learning*, Cambridge, MA, USA, 2006. The MIT Press.
- [11] C.J.C. Burges. *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*, chapter *Geometric Methods for Feature Selection and Dimensional Reduction: A Guided Tour*. Kluwer Academic Publishers, 2005.
- [12] R.Szeliski, "Image alignment and stitching: A tutorial". *Foundations and Trends in Computer Graphics and Computer Vision*,2:1-104, Dec 2006.
- [13] J Zhang, S Li, J Wang, manifold learning and applications in recognition, *Intelligent Multimedia Processing with Soft Computing*, 281-300
- [14] Camillo J. Taylor, Arvind Bhusnurmath: Solving Image Registration Problems Using Interior Point Methods. *ECCV (4) 2008*: 638-651.
- [15] P. Ndjiki-Nya, J. Restat, T. Meiers, J.-R. Ohm, A. Seyferth, and R. Sniehotta, "Subjective Evaluation of the MPEG-7 Retrieval Accuracy Measure (ANMRR)," *ISO/WG11 MPEG Meeting*, Geneva, Switzerland, Doc. M6029, May 2000.
- [16] S.-J. Park, C. S. Won, and D. K. Park, "Core Experiments on MPEG-7 Edge Histogram Descriptor," *ISO/WG11 MPEG Meeting*, Beijing, China, MPEG Document M6174, July 2000.
- [17] NeeteshPrajapati, Amit Kumar Nandanwar, G.S. Prajapati ; Edge Histogram Descriptor, Geometric Moment and Sobel Edge Detector Combined Features Based Object Recognition and Retrieval System, NeeteshPrajapati et al, / (IJCSIT) *International Journal of Computer Science and Information Technologies*, Vol. 7 (1) , 2016, 407-412
- [18] LJP Van der Maaten, EO Postma, HJ Van den Herik, Dimensionality reduction: A comparative review, *Technical Report TiCC TR 2009-005*
- [19] Joshua M. Lewis, Laurens van der Maaten, Virginia de Sa (2012). A Behavioral Investigation of Dimensionality Reduction. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, pp. 671-676.

- [20] Cencheng Shen, Joshua T. Vogelstein, Carey E. Priebe, *Manifold Matching using Shortest-Path Distance and Joint Neighborhood Selection*, *Pattern Recognition Letters*, vol. 92, pp. 41-48, 2017
- [21] Hongyan Zhang, Tuotuo Li, Jason Geng; *Manifold Modeling and Its Application to Tubular Scene Manifold Mosaicing Algorithm*, *Journal of Mathematical Imaging and Vision*, Volume 44 Issue 1, September 2012 Pages 80-98.
- [22] Peleg, S., Rousso, B.: *Generalized panoramic mosaic*. Patent US7006124, Feb. 28, 2006
- [23] Fang, X.Y., Zhang, M.M., Pan, Z.G., Wang, P.: *A new method of manifold mosaic for large displacement images*. *J. Comput. Sci. Technol.* 21(2), 218–223 (2006). 13th Pacific Conference on Computer Graphics and Applications, Macao, China, Oct. 12–14, 2005.
- [24] A. Yamada, and E. Kasutani, “*The MPEG-7 Color Layout Descriptor: A Compact Image Feature Description for Highspeed Image/Video Segment Retrieval*,” *ICIP*, vol. 1, Oct. 2001, pp. 7-10.
- [25] C.M. Bishop, M. Sevensen, and C.K.I. Williams, (1998), “*GTM: The generative topographic mapping*,” *Neural Computation*, 10, pp.215-234. [26] B. Kégl, A. Krzyzak, T. Linder, and K. Zeger, (2000), “*Learning and design of principal curves*”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 3, pp. 281-297.
- [27] A.J. Smola, S.Mika, et al, (1999), “*Regularized Principal Manifolds*,” *In Computational Learning Theory: 4th European Conference*, Vol 1572 of *Lecture Notes in Artificial Intelligence*, New York: Springer, pp. 251-256
- [28] V. S.Silva, J. B. Tenenbaum, (2002), “*Unsupervised Learning of curved manifolds*”, *Nonlinear Estimation and Classification*, Springer-Verlag, New York.
- [29] S. T. Roweis, and K. S. Lawrence, (2000), “*Nonlinear Dimensionality reduction by locally linear embedding*,” *Science*, 290, pp. 2323-2326
- [30] Joshua B. Tenenbaum, Vin de Silva, John C. Langford,, *A Global Geometric Framework for Nonlinear Dimensionality Reduction*, *R E P O R T S*, [www.sciencemag.org](http://www.sciencemag.org) *SCIENCE VOL 290 22 DECEMBER 2000*.
- [31] Sam T. Roweis and Lawrence K. Saul, *Nonlinear Dimensionality Reduction by Locally Linear Embedding*, *R E P O R T S*, [www.sciencemag.org](http://www.sciencemag.org) *SCIENCE VOL 290 22 DECEMBER 2000*.
- [32] Y. Bengio, P. Vincent, J-F. Paiement, O. Delalleau, M. Ouimet, and N. Le Roux. *Spectral Clustering and Kernel PCA are Learning Eigenfunctions*. Technical Report 1239, Département d’informatique et recherche opérationnelle, Université de Montreal, 2003.
- [33] Y. Bengio, J-F. Paiement, and P. Vincent. *Out-of-Sample Extensions for LLE, Isomap, MDS, Eigenmaps, and Spectral Clustering*. Technical Report 1238, Département d’informatique et recherche opérationnelle, Université de Montreal, 2003.
- [34] G. Hinton and S. Roweis, (2002), “*Stochastic Neighbor Embedding*,” *Neural Information Proceeding Systems: Natural and Synthetic*, Vancouver, Canada, December 9-14.
- [35]. M. Brand, MERL, (2002), “*Charting a manifold*,” *Neural Information Proceeding Systems: Natural and Synthetic*, Vancouver, Canada, December 9-14.
- [36]. Dick de Ridder, Robert P.W. Duin, “*Locally Linear Embedding for Classification*”, *technical reports*, Delft University of Technology, The Netherlands .
- [37] B. S. Manjunath, P. Salembier, and T. Sikora, Eds., *Introduction to MPEG-7: Multimedia Content Description Interface*. England: WILEY, 2002.
- [38] [https://ch.mathworks.com/fr/products/get-matlab.html?s\\_tid=gn\\_getml](https://ch.mathworks.com/fr/products/get-matlab.html?s_tid=gn_getml). Date de consultation 08/09/2021.