

الجمهورية الجزائرية الديمقراطية الشعبية
République algérienne démocratique et populaire
وزارة البحث العلمي والتعليم العالي
Le Ministère de la Recherche Scientifique et de l'Enseignement Supérieur

Université Mohamed Khider Biskra
Faculté des Sciences et de la Technologie
Département : Génie Electrique
Ref :.....



جامعة محمد خيضر بسكرة
كلية العلوم و التكنولوجيا
قسم: الهندسة الكهربائية
المرجع:.....

Thèse présentée en vue de l'obtention
du diplôme de

Doctorat en sciences en Electronique

Option : Electronique

Titre

Reconnaissance Biométrique par Fusion Multimodale de Visages

Présentée par :

ADEL SAOUD

Soutenue publiquement le 30/06/2022

Devant le jury composé de :

BAARIR Zine-Eddine	Professeur	Président	Université de Biskra
OUAFI Abdelkrim	Professeur	Rapporteur	Université de Biskra
TALEB-AHMED Abdelmalik	Professeur	Co-rapporteur	Université de Valenciennes France
OUAMANE Abdelmalik	Professeur	Examineur	Université de Biskra
AJGOU Riadh	MCA	Examineur	Université d'El Oued
HETTIRI Messaoud	MCA	Examineur	Université d'El Oued

Remerciements

La soutenance d'une thèse marque assurément la fin d'une étape dans la vie d'un chercheur. Je voudrais saisir cette occasion pour exprimer ma gratitude à tous ceux qui m'ont soutenu, encouragé à poursuivre ce travail et à le finir.

Tout d'abord, je tiens à remercier Monsieur **OUAFI Abdelkrim**, Professeur à l'Université de Biskra, qui m'a encadré tout au long de cette thèse et qui m'a fait partager ses brillantes intuitions. Qu'il soit aussi remercié pour sa gentillesse, sa disponibilité permanente et pour les nombreux encouragements qu'il m'a prodigués.

J'adresse tous mes remerciements à Monsieur **TALEB-AHMED Abdelmalik**, Professeur à l'Université de Valenciennes en France, de l'honneur qu'il m'a fait en acceptant d'être Co-rapporteur de cette thèse.

J'exprime ma gratitude à Monsieur **BAARIR Zine-Eddine**, Professeur de l'Université de Biskra, qui m'a fait l'honneur de présider le Jury de cette thèse.

Je tiens à remercier chaleureusement Mr **OUAMANE Abdelmalik**, Professeur à l'Université de Biskra, pour sa contribution, son aide et ses encouragements qu'il m'a prodigués durant la période de la thèse.

Je voudrais, également, exprimer ma gratitude à Monsieur **AJGOU Riadh** et à Monsieur **HETTIRI Messaoud**, Maitres de Conférences Assistants à l'Université d'El Oued, membres du jury, qui ont eu l'amabilité de lire ce travail, de l'enrichir et de le discuter.

Je tiens aussi à remercier Monsieur **Chouchane Ammar** Maitre-assistant (A) à l'Université de Médéa pour l'aide et la correction de mon article.

Finalement je remercie **mes parents** et **mes amis** pour leurs encouragements et leurs soutiens qui m'a été bien utile durant ma thèse.

Résumé

Les systèmes biométriques sont considérés comme l'une des méthodes les plus efficaces pour protéger et sécuriser la vie privée ou publique contre tous les types de vol. La reconnaissance faciale est l'une des méthodes les plus utilisées, non pas parce qu'elle est la plus efficace et la plus fiable, mais plutôt parce qu'elle est naturelle et non intrusive et relativement acceptée par rapport aux autres biométries telles que les empreintes digitales et l'iris. L'objectif de développement d'applications biométriques, comme la reconnaissance faciale, est récemment devenu important dans les villes intelligentes. Au cours des dernières décennies, de nombreuses techniques, dont les applications incluent les systèmes de visioconférence, la reconstruction faciale, la sécurité, etc. proposées de reconnaître un visage dans une image 2D ou 3D. généralement, le changement d'éclairage, les variations de pose et les expressions faciale rendent la reconnaissance faciale 2D loin d'être fiable. Cependant, les modèles 3D peuvent être capable de surmonter ces contraintes, sauf que la plupart des méthodes de reconnaissance faciale 3D traitent toujours le visage humain comme un objet rigide. Cela signifie que ces méthodes ne sont pas capables de gérer les expressions faciales.

Dans cette thèse, nous proposons une nouvelle approche pour la vérification automatique de visage en codant les informations locales des images faciales 2D et 3D comme un tenseur d'ordre élevé. D'abord, les histogrammes de deux descripteurs locaux multi-échelles (LPQ et BSIF) sont utilisés pour caractériser à la fois l'images faciale en 2D et 3D. Ensuite, une représentation du visage basée sur tenseur est conçu pour combiner toutes les caractéristiques extraites à partir de visages 2D et 3D. De plus, pour améliorer la discrimination de la représentation de visage tenseur proposée, nous avons utilisé deux méthodes de sous-espace multilinéaire (MWPCA et MDA combinée avec WCCN). De plus, la technique WCCN est appliquée aux tenseurs de visage pour réduire l'effet des directions intra-classes en utilisant une transformée de normalisation, ainsi que pour améliorer le pouvoir de discrimination de la MDA. Nos expériences ont été réalisées sur les trois plus grandes bases de données : FRGC v2.0, Bosphorus et CASIA 3D sous différentes expressions faciale, variations de pose et occlusions. Les résultats expérimentaux ont montré la supériorité de l'approche proposée en termes de taux de vérification par rapport à la méthode récentes de l'état de l'art.

Mots clés : vérification de visage 2D et 3D, tenseur, transformation de sous espace multilinéaire, la fusion multimodale 2D+3D. MWPCA, MDA, WCCN.

المخلص

تعتبر أنظمة القياسات الحيوية من أكثر الطرق فعالية لحماية وتأمين الحياة الخاصة والعامّة ضد كل أنواع السرقة. يعد التعرف على الوجه من أكثر الطرق استخدامًا ، ليس لأنه الأكثر كفاءة وموثوقية ، بل لأنه طبيعي وغير تدخلي ومقبول نسبيًا مقارنة بالقياسات الحيوية الأخرى مثل بصمة الإصبع وقزحية العين.

أصبح الهدف من تطوير تطبيقات القياسات الحيوية ، كالاعرف على الوجه ، مهمًا مؤخرًا نظرا لاستعمالاتها في العديد من الأماكن كالمدن الذكية مثلا. وعلى مدى العقود الماضية ، ظهرت العديد من التقنيات التي تستخدم في تطبيقاتها ميزة التعرف على صور الوجه ثنائي (2D) أو ثلاثي (3D) الأبعاد كالمؤتمرات عبر الفيديو ، وأنظمة إعادة بناء الوجه ، والأمن ، وما إلى ذلك.

بشكل عام ، إن التغيير في الإضاءة والاختلافات في وضعيات الرأس وتعابير الوجه تصعب من عملية التعرف على الوجه ثنائي (2D) الأبعاد، في حين قد لا تتأثر النماذج ثلاثية (3D) الأبعاد بهذه العوامل. بالرغم من أن معظم طرق التعرف على الوجه ثلاثي (3D) الأبعاد تعطي مردودا جيدا إلا أنها لا تزال تتعامل مع وجه الإنسان ككائن صلب. وهذا ما يجعلها غير قادرة على التعامل مع تعابير الوجه بشكل خاص.

في هذه الأطروحة نقترح نهجًا جديدًا للتحقق التلقائي من الوجه عن طريق تشفير المعلومات المحلية لصور الوجه ثنائية (2D) وثلاثية (3D) الأبعاد كموتر عالي المستوى. أولاً ، يتم استخدام الرسوم البيانية لاثنين من الواصفات المحلية متعددة النطاقات (LPQ و BSIF) لوصف كل من صور الوجه ثنائية (2D) وثلاثية (3D) الأبعاد. بعد ذلك ، يتم تصميم تمثيل الوجه المستند إلى الموتر لدمج جميع الميزات المستخرجة من الوجوه ثنائية وثلاثية الأبعاد. علاوة على ذلك ، لتحسين التمييز في تمثيل وجه الموتر المقترح ، استخدمنا طريقتين متعددتي الخطوط للفضاء الجزئي (MWPCA و MDA جنبًا إلى جنب مع WCCN). بالإضافة إلى ذلك ، يتم تطبيق تقنية WCCN للتقليل من تأثير الاتجاهات داخل الطبقة باستخدام تحويل التطبيق ، وكذلك لتحسين القوة التمييزية لـ MDA. تم إجراء تجاربنا على ثلاثة من أكبر قواعد البيانات: FRGC v2.0 و Bosphorus و CASIA 3D حيث تضمنت تعابير وجه متعددة وضعيات رأس مختلفة. أظهرت النتائج التجريبية تفوق النهج المقترح من حيث معدل التحقق مقارنة بأحدث الأساليب الحديثة.

Abstract

Biometric systems are considered to be one of the most effective methods of protecting and securing private or public life against all types of theft. Facial recognition is one of the most widely used methods, not because it is the most efficient and reliable, but rather because it is natural and non-intrusive and relatively accepted compared to other biometrics such as fingerprint and iris. The goal of developing biometric applications, such as facial recognition, has recently become important in smart cities. Over the past decades, many techniques, the applications of which include videoconferencing systems, facial reconstruction, security, etc. proposed to recognize a face in a 2D or 3D image. Generally, the change in lighting, variations in pose and facial expressions make 2D facial recognition less than reliable. However, 3D models may be able to overcome these constraints, except that most 3D facial recognition methods still treat the human face as a rigid object. This means that these methods are not able to handle facial expressions.

In this thesis, we propose a new approach for automatic face verification by encoding the local information of 2D and 3D facial images as a high order tensor. First, the histograms of two local multiscale descriptors (LPQ and BSIF) are used to characterize both 2D and 3D facial images. Next, a tensor-based facial representation is designed to combine all the features extracted from 2D and 3D faces. Moreover, to improve the discrimination of the proposed tensor face representation, we used two multilinear subspace methods (MWPCA and MDA combined with WCCN). In addition, the WCCN technique is applied to face tensors to reduce the effect of intra-class directions using a normalization transform, as well as to improve the discriminating power of MDA. Our experiments were carried out on the three largest databases: FRGC v2.0, Bosphorus and CASIA 3D under different facial expressions, variations in pose and occlusions. The experimental results have shown the superiority of the proposed approach in terms of verification rate compared to the recent state-of-the-art method.

Keywords: 2D and 3D face verification, tensor, transformation of multilinear subspace, 2D + 3D multimodal fusion. MWPCA, MDA, WCCN.

Table de matières

Remerciements	ii
Résumé	iii
الملخص	iv
Abstract	v
Table de matières	vi
Liste des figures	ix
Liste des tableaux	xii
Abréviations	xiii
Notations mathématiques	xv
Chapitre 1 : Introduction générale	1
1.1 Contexte et motivation	1
1.2 Problématique	3
1.3 Objectifs et contributions	6
1.4 Modes de fonctionnement d'un système biométrique.....	7
1.5 Evaluation des performances d'un système biométrique.....	8
1.6 Organisation de la thèse	10
Chapitre 2 : Etat de l'art sur la reconnaissance faciale.....	12
2.1 Introduction.....	12
2.2 Reconnaissance de visage 2D	12
2.2.1 Approches globales.....	13
2.2.1.1 Les techniques linéaires.....	13
2.2.1.2 Les techniques non linéaires :	19
2.2.2 Approches locales.....	21
2.2.2.1 Techniques basées sur l'apparence locale.....	21

2.2.2.2	Techniques basées sur les points d'intérêt	23
2.2.3	Approches hybrides	24
2.2.4	Comparaison entre les techniques de reconnaissance faciale 2D	25
2.3	La reconnaissance de visage 3D	30
2.3.1	Approches globales.....	30
2.3.1.1	Les méthodes de sous-espace	30
2.3.1.2	Méthodes basées sur l'alignement de surfaces	31
2.3.1.3	Méthodes basées le modèle morphable (déformable)	35
2.3.2	Approches locales	38
2.3.2.1	Méthodes basées sur le calcul des points clés (Keypoint).....	38
2.3.2.2	Méthodes basées sur le calcul de courbure.....	41
2.3.2.3	Méthodes locale basée sur la surface.....	44
2.3.3	Discussion.....	47
2.4	La reconnaissance de visage 2D+3D	49
2.5	Conclusion	52
Chapitre 3 : Apprentissage sous-espace multilinéaire		54
3.1	Introduction.....	54
3.2	Propriétés des tenseurs pour la modélisation de données multidimensionnelles	54
3.2.1	Notations et concepts	54
3.2.1.1	Définition d'un tenseur.....	54
3.2.2	Opérations de bases d'algèbre multilinéaire.....	57
3.2.2.1	Vectorisation : transformation tenseur en vecteur	57
3.2.2.2	Déploiement (Matricisation) : Transformation tenseur en matrice	57
3.2.2.3	Multiplication des tenseurs.....	58
3.2.2.4	Produit scalaire et la norme Frobenius	60
3.2.2.5	Tenseur de rang un (Rank-one tensor)	60
3.2.2.6	La distance euclidienne	61
3.2.2.7	Produit Kronecker	61
3.2.3	Projections multilinéaires	61
3.2.3.1	Projection : vecteur à vecteur (PVV).....	62
3.2.3.2	Projection : tenseur à tenseur (PTT).....	63
3.2.3.3	Projection : tenseur à vecteur (PTV)	64
3.2.4	Relation entre les projections multilinéaires.....	66
3.3	Multilinear Whitening Principal Component Analysis (MWPCA)	69

3.4 Multilinear Discriminant Analysis (MDA).....	69
3.5 Within-Class Covariance Normalization (WCCN).....	70
3.6 L'algorithme MDA+WCCN.....	72
3.7 Conclusion :	73
Chapitre 4 : Résultats et Discussions	75
4.1 Introduction.....	75
4.2 Aperçu de l'approche de vérification de visage proposée.....	75
4.2.1 Pré-traitement de visage 2D et 3D.....	75
4.2.2 Extraction des caractéristiques	75
4.2.2.1 Descripteur LPQ.....	76
4.2.2.2 Descripteur BSIF.....	75
4.2.3 Réduction de la dimensionalité, comparaison et décision	80
4.3 Résultats expérimentaux	80
4.3.1 Les bases de données	81
4.3.1.1 FRGC v2.0	81
4.3.1.2 Bosphorus.....	82
4.3.1.3 CASIA 3D.....	83
4.4 Configuration de paramètres.....	83
4.5 Discussion et comparaison.....	83
4.5.1 Descripteurs locaux LPQ et BSIF.....	84
4.5.2 L'approche proposée par rapport à la méthode MDA	86
4.5.3 Les performances de la modalité 2D et 3D.....	88
4.5.4 Les performances de la vérification multimodale de visage 2D + 3D.....	89
4.5.5 L'approche proposée VS l'état de l'art	90
4.6 Conclusion	93
Chapitre 5 : Conclusion générale et perspectives.....	94
5.1 Conclusion générale.....	94
5.2 Perspectives.....	99
Bibliographie.....	101

Liste des figures

Chapitre 1 : Introduction générale

Figure 1.1 différents types de modalités biométriques (les images sont recueillies via l'internet).....	2
Figure 1.2 Comparaison de diverses caractéristiques biométriques basées sur la compatibilité	3
Figure 1.3 : système biométrique de reconnaissance faciale en mode identification	7
Figure 1.4 : système biométrique de reconnaissance faciale en mode vérification.	8
Figure 1.5 : illustration du taux FRR et taux FAR.	9
Figure 1.6 : la courbe ROC	9

Chapitre 2 : Etat de l'art sur la reconnaissance faciale

Figure 2.1 Vecteurs propres (Eigenfaces) des 12 plus grandes valeurs propres obtenus à partir de la base de données ORL [14].....	14
Figure 2.2 Exemple d'un vecteur de caractéristiques réduit par PCA.	15
Figure 2.3 Images reconstruites par la méthode 2D-PCA [15].	15
Figure 2.4 Les 12 premiers vecteurs (Fisherfaces) obtenus à partir de la base de données ORL [14].	16
Figure 2.5 Exemple de vecteurs de base ICA provenant de la base de données ORL [14]. ...	17
Figure 2.6 Exemple d'un code binaire LBP.	22
Figure 2.7 Classification des algorithmes principaux de reconnaissance faciale 2D.....	29
Figure 2.8 Exemple 1 recalage et calcul de similarité entre un modèle 3D et un modèle requête de la même personne avec des variations d'expressions faciales [110].....	33
Figure 2.9 Exemple 2 recalage entre un modèle 3D et un modèle requête de la même personne avec des variations d'expressions faciales plus significatives [110].	34
Figure 2.10 Exemple 3: recalage entre un modèle 3D et un modèle requête de deux personnes différentes [110].	35
Figure 2.11 Modélisation des déformations pour l'appariement des visages 3D [112].....	36
Figure 2.12 Exemples de visages d'ajustement du modèle morphable ont différentes expressions [115].....	38
Figure 2.13 Un point clé sur une visage 3D et sa texture correspondante [118].....	39
Figure 2.14 Points saillants détectées par K_{max} (à gauche) et K_{min} (à droite) [124].	40

Figure 2.15 Orientations canoniques (flèche), point saillant et ses 8 sommets de voisinage (+) affectés de 9 cercles [124].	41
Figure 2.16 Courbes de niveaux a) de la fonction géodésique [126], b) de la fonction de profondeur [125] pour plusieurs niveaux.	42
Figure 2.17 Procédure d'extraction des courbes radiales [135].	43
Figure 2.18 Le masque binaire et 17 SRAs [137].	44
Figure 2.19 Illustration de la représentation normale du visage : histogramme de la LNP [144].	46

Chapitre 3 : Apprentissage sous-espace multilinéaire

Figure 3.1 Exemples de la représentation tensorielle : (a) image en niveau de gris (mode 2) ; (b) image couleur (mode 3) ; (c) séquence vidéo en couleur (mode 4).	55
Figure 3.2 Passage des scalaires aux tenseurs.	56
Figure 3.3 Exemples d'un tenseur du troisième ordre sous forme des fibres et des tranches.	56
Figure 3.4 Exemple de déploiement d'un tenseur d'ordre 3 dans mode1, 2 et 3 [192].	58
Figure 3.5 Illustration visuelle de la multiplication tenseur-matrice dans le mode 1.	59
Figure 3.6 Exemple d'un tenseur de rang 1 du troisième ordre 3.	61
Figure 3.7 Illustration de la projection : vecteur à vecteur (PVV) [195].	63
Figure 3.8 Projection Tenseur à tenseur (PTT) [195].	63
Figure 3.9 Projection tenseur à vecteur (PTV) [195].	64
Figure 3.10 Illustration de la projection d'un tenseur à vecteur à travers la projection multilinéaire élémentaire [195].	65
Figure 3.11 Comparaison du nombre de paramètres à estimer par PVV, PTV et PTT normalisés par rapport au nombre VVP pour visualisation [195].	68

Chapitre 4 : Résultats et Discussions

Figure 4.1 Système proposé pour la vérification multimodale de visage 2D+3D.	76
Figure 4.2 Procédures de générer le code LPQ.	78
Figure 4.3 Extraction des caractéristiques pour l'image de profondeur et d'intensité en utilisant LPQ multi-échelles.	78
Figure 4.4 Extraction des caractéristiques par le descripteur BSIF.	80
Figure 4.5 Ensemble d'images faciales de la base de données FRGC v2.0.	82
Figure 4.6 Ensemble des échantillons d'images faciales extraites de base de données Bosphorus [206].	83

Figure 4.7 Exemples des images faciales extraites de base de données CASIA 3D.....	83
Figure 4.8 Courbes ROC1 de MDA et MDA + WCCN pour les images faciales 2D, 3D et 2D + 3D sur bases de données FRGC v2.0.	88
Figure 4.9 Courbes ROC de la vérification de visage 2D, 3D et 2D + 3D sur la base de données FRGC v2.0.	89

Liste des tableaux

Chapitre 2 : Etat de l'art sur la reconnaissance faciale

Tableau 2.1 Avantages et les inconvénients des approches de la reconnaissance faciale 2D [98].	28
--	----

Chapitre 3 : Apprentissage sous-espace multilinéaire

Tableau 3.1 Nombre de paramètre à estimer par les trois projections multilinéaires [196]. ..	68
Tableau 3.2 L'algorithme de l'approche proposée.....	73

Chapitre 4 : Résultats et Discussions

Tableau 4.1 Taux de vérification 2D, 3D et 2D + 3D à 0,001 FAR (%) sur la base de données FRGC v2.0.	85
Tableau 4.2 Taux d'erreur égal (%) 2D, 3D et 2D + 3D sur la base de données FRGC v2.0.	86
Tableau 4.3 Taux de vérification 2D, 3D et 2D + 3D à 0,001 FAR (%) sur la base de données CASIA 3D.	87
Tableau 4.4 Taux de vérification 2D, 3D et 2D + 3D à 0,001 FAR (%) sur la base de données Bosphorus.....	87
Tableau 4.5 Taux de vérification (%) de l'approche proposée et de l'état de l'art sur la base de données FRGC v2.0.	91
Tableau 4.6 Taux de vérification (%) de l'approche proposée et l'état de l'art sur la base de données CASIA 3D.....	92
Tableau 4.7 Taux de vérification (%) de l'approche proposée et l'état de l'art sur la base de données Bosphorus.....	92

Abréviations

TFR : Taux de Faux de Rejet (FRR : False Rejection Rate).

TFA : Taux de Fausse Acceptation (FAR : False Accept Rate)

EER: Equal Error Rate.

ROC: Receiver Operating Characteristic.

CMC: Cumulative Match Characteristic.

LDA: Local Discriminant Analysis.

ICA: Independent Component Analysis.

PCA: Principal Component Analysis.

2-D PCA: Two-Dimensional Principal Component Analysis.

NMF: non-Negative Matrix Factorization.

DWT: Discrete Wavelet Transform.

RDA: Regularized Discriminant Analysis.

DCT: Discrete Cosine Transform.

EFM: Enhanced Fisher linear discriminant Model.

DCV: Discriminant Common Vector.

BDA: Bilinear Discriminant Analysis.

KPCA: Kernel Principal Component Analysis.

KDA: Kernel Linear Discriminant Analysis.

SVM: Support vector machine.

KICA: kernel independent component analysis.

EDA: Exponential discriminant analysis.

LPP: Locality preserving projection.

LLE: Local linear embedding.

KFD: Kernel Fisher discriminant analysis.

LTSA: Local tangent space analysis.

EWPCA: Evolutionary weighted principal component analysis.

WT: Wavelet transform.

RT: Radon Transform.

CNN: Cellular Neural Networks.

LBP: Local Binary Patterns.

BSIF: Binarized Statistical Image Features.

LPQ: Local Phase Quantization.

HOG: Histogram of oriented gradients.

SIFT: Scale Invariant Feature Transform.

BRIEF: Binary Robust Independent Elementary Features.

SURF: Speeded-Up Robust Features.

EBGM: Elastic bunch graph matching.

LDP: Local directional pattern.

LGBPHS: Local Gabor Binary Pattern Histogram Sequence.

HMM: Hidden Markov model.

AAM: Active Appearance Model.

GMM: Gaussian Mixture Model.

ICP: Iterative Closest Point.

MPCA: Multilinear Principal Component Analysis.

MWPCA: Multilinear Whitening Principal Component Analysis.

MDA: Multilinear Discriminant Analysis.

WCCN: Within-Class Covariance Normalization.

Notations mathématiques

$\mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$: espace tensoriel.

$x \in \mathbb{R}$: Scalaire.

$\mathbf{x} \in \mathbb{R}^{I_1}$: Vecteur de dimension I_1 .

$\mathbf{X} \in \mathbb{R}^{I_1 \times I_2}$: Matrice de dimension $I_1 \times I_2$.

$\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$: Tenseur d'ordre N de dimension $I_1 \times I_2 \times \dots \times I_N$.

$\mathbf{A}_{(n)}$: Matrice dépliant dans le mode n du tenseur \mathcal{A} .

$\|\mathcal{A}\|_F$: La norme Frobenius du tenseur \mathcal{A} .

\circ : le produit extérieur de N vecteurs.

$\|\mathcal{A} - \mathcal{B}\|_F$: La distance entre les tenseurs \mathcal{A} et \mathcal{B} .

$vec(\mathcal{A})$: la représentation vectorielle de \mathcal{A} .

\otimes : Le produit Kronecker.

$\mathbf{U}^{(n)}$: Matrice de projection dans le mode n .

$\bar{\mathcal{B}}$: le moyen du tenseur \mathcal{B} .

$\psi_{\mathcal{B}}$: la dispersion totale du tenseur \mathcal{B} .

\mathbf{S}_b^k : matrices de dispersion inter-classe.

\mathbf{S}_w^k : matrices de dispersion intra-classe.

$\mathbf{S}_{T_B}^{(k)}$: La matrice de dispersion totale en mode k de N échantillons.

$\mathbf{A}_{j,i}^k$: la matrice dépliée en mode k du tenseur \mathcal{A}_i .

\mathbf{Z}_k : la matrice de projection de WCCN.

ε : Petit nombre constant.

M : le nombre d'échantillons d'apprentissage appartenant à la classe C .

itr_{max} : le nombre maximal d'itérations dans les solutions multilinéaires.

Chapitre 1

Introduction générale

1.1 Contexte et motivation	1
1.2 Problématique	3
1.3 Objectifs et contributions	6
1.4 Modes de fonctionnement d'un système biométrique.....	7
1.5 Evaluation des performances d'un système biométrique.....	8
1.6 Organisation de la thèse	10

Chapitre 1 : Introduction générale

1.1 Contexte et motivation

Ces dernières années, les systèmes biométriques apparues comme une solution prometteuse pour divers organismes gouvernementaux. Identifier ou vérifier les gens à partir leurs traits biologiques reste le choix le plus favorable grâce à ses efficacités contre les méthodes frauduleuses. L'utilisation d'une clé, mot de passe, code pin, carte à puce pour identifier ou vérifier un individu présente certains risques. Le mot de passe et le code pin sont difficiles à mémoriser longtemps et peuvent être voler ou deviner ; les cartes à puce et les clés peuvent être aussi égarer, oublier, voler ou dupliquer. Cependant, les traits biologiques d'un individu comme les caractéristiques physiques (tels que visage, iris, rétine, empreintes digitales, géométrie de la main, géométrie de l'oreille) ou comportementales (tels que la voix, la signature, la démarche, la dynamique des frappes au clavier) ne peuvent être volés ou falsifiés [1]. Ces caractéristiques physiques ou comportementales doivent répondre à certaines mesures biologiques de la biométrie, qui sont identifiés par *Jain et al* [2]:

- **Universalité** : indique que chaque personne doit avoir une caractéristique spécifique.
- **Unicité** : deux personnes doivent être suffisamment différent en termes de caractéristique.
- **Permanence** : indique que les traits biométriques doivent être suffisamment invariants sur une période de temps.
- **Enregistrable** : indique que la caractéristique doit pouvoir être acquise.
- **Performance** : indique la précision et la vitesse du système biométrique, ainsi que les ressources nécessaires pour atteindre la reconnaissance désirée sous différentes contraintes.
- **Acceptabilité** : indique la volonté du publique d'utiliser l'identificateur biométrique.
- **Contournement** : indique que le système biométrique doit être difficile à tromper par les méthodes frauduleuses.

La plupart de ces méthodes biométriques dépendent de la volonté de l'utilisateur. Par exemples, la reconnaissance des empreintes digitales ou la géométrie de la main obligent l'usager de placer sa main ou son doigt sur un détecteur de la géométrie. La reconnaissance de l'iris ou la rétine besoins à l'utilisateur à se tenir dans une position fixe et très proche de dispositif pour les détecter. Ces actions sont relativement mal reçues par l'usager. La reconnaissance faciale reste

la biométrie la plus populaire et commune par à rapport aux autres méthodes biométriques. Elle dispose un niveau de sécurité très satisfaisant. En plus, elle est non intrusive parce la capture de visage s'effectue à distance sans contact avec le dispositif. La réalisation d'un système de la reconnaissance faciale demande qu'un simple équipement (une caméra et un ordinateur) contraire aux empreintes digitales et iris qui nécessitent des équipements coûteux et trop sensibles au mouvement du corps. D'autre part, le tissu de l'épiderme de la main ou le doigt peut endommager par le temps et devient inutile grâce à divers accidents. Tandis que les traits de visage tels que les yeux, le nez, les lèvres, la bouche et les oreilles sont invariants avec temps. En outre, l'une des raisons pour lesquelles la reconnaissance de visage a attiré l'attention de la recherche scientifique et le développement durant les 40 années passées, est son succès dans nombreux applications commerciales ou gouvernementaux. Notamment les applications de



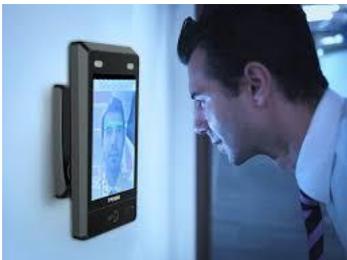
Géométrie de la main



Empreintes digitales



Voix



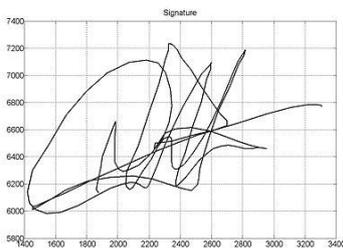
Visage



DNA



Démarche



Signature



Rétine



Dynamique de frappe

Figure 1.1 différents types de modalités biométriques (les images sont recueillies via l'internet).

vidéo surveillance automatique, contrôle d'accès, communication multimédia, ...etc.

D'autre part, Parmi les six attributs biométriques : le visage, la voix, les yeux (iris et rétine), géométrie de la main, la signature, les empreintes digitales (voir la **Figure 1.1**) considérés par Hietmeyer [3]. Comme indique la **Figure 1.2**, les traits du visage marquent le meilleur score de compatibilité dans le système MRTD (Machine-Readable Travel Documents) basé sur plusieurs facteurs d'évaluation, comme l'inscription, renouvellement de données, la redondance, les matériaux requis et la perception du public.

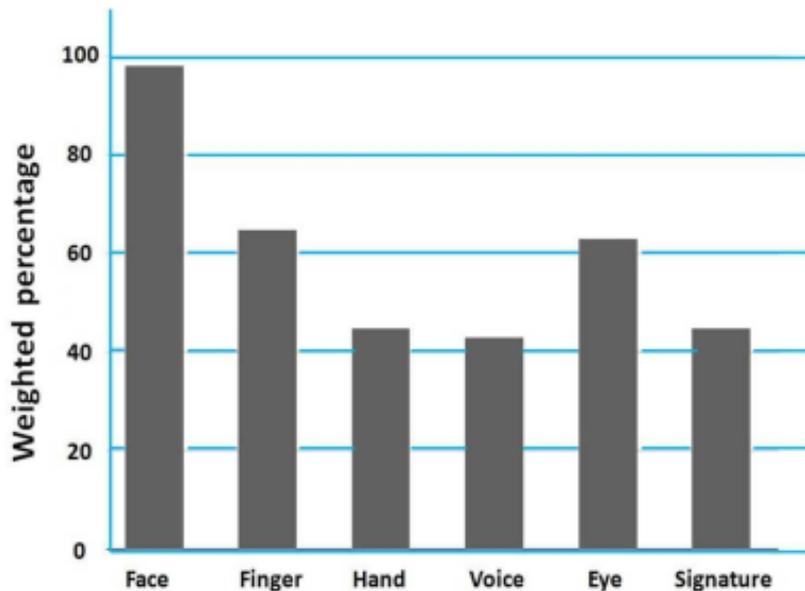


Figure 1.2 Comparaison de diverses caractéristiques biométriques basées sur la compatibilité MRTD (Machine-Readable Travel Documents).

Pour construire un système de reconnaissance faciale fiable, deux exigences critiquables doivent être disponibles : une grande base de données d'images et un protocole de test pour évaluer les systèmes. En septembre 1993, le programme de technologie de la reconnaissance faciale (FERET) a commencé de collecter la base de données FERET, et les premiers tests d'évaluation avec cette base de données FERET sont effectués en août 1994 et mars 1995 [4]. Le programme FERET va pousser les chercheurs à évaluer et améliorer la performance de leurs algorithmes, ce que conduit à enrichir l'état de l'art de la reconnaissance faciale. Couramment, l'aspect évolué des matériels et logiciels a motivé beaucoup l'exploration de la technologie de reconnaissance faciale soit en marché commercial soit en les recherches académiques.

1.2 Problématique

Les images de visages humaines généralement représentées par des milliers de pixels encodées dans un tableau multidimensionnelle. Cependant, ils sont intrinsèquement intégrés dans un sous-espace de très faible dimension. Une image de visage en niveaux de gris 2D,

images couleur ou une séquence vidéo 3D généralement se présentent comme des objets multidimensionnels, qui sont appelés tenseurs. Le terme tenseur a différentes significations en mathématiques et en physique [5]. L'usage de ce terme dans notre thèse fait la référence à sa signification en mathématiques, en particulier en algèbre multilinéaire. Le tenseur est défini par deux indices, qui sont l'ordre du tenseur et les modes de représentations. De nombreuses images et séquences vidéo sont naturellement des objets tensoriels. Par exemples, les images de niveau de gris sont des tenseurs d'ordre deux avec modes de colonnes et lignes, les images couleurs sont des tenseurs d'ordre trois avec modes de colonnes, lignes et couleur. Les séquences vidéo en niveaux de gris peuvent être visualisés comme des tenseurs de troisième ordre avec modes de colonnes, lignes et temps. Naturellement, les séquences vidéo couleur sont des tenseurs du 4ème ordre avec l'ajout d'un mode couleur.

Généralement, Le système de reconnaissance faciale se compose de trois étapes essentielles, qui sont : la détection de visages, extraction des caractéristiques et la classification (reconnaissance). Les caractéristiques faciales comportent des données naturellement multidimensionnelles inclue divers facteurs tels que l'éclairage de l'environnement, les variations de pose de la tête humaine, les différentes expressions faciales...etc. Les méthodes d'analyses des caractéristiques de visages qui agissent directement sur un espace de grande dimension souffrent généralement de problème appelé malédiction de la dimensionnalité. Ce problème se produit quand le nombre d'échantillons d'apprentissage disponibles par sujet est généralement beaucoup plus petit que la dimensionnalité de l'espace de l'échantillon.

La réduction de dimensionnalité via l'apprentissage du sous espace est une tentative de transformer un ensemble de données de grande dimension à un sous espace de dimension inférieur tout en préservant la plus part d'informations concernant la structure sous adjacente dans les données [6]. La majorité des méthodes de réduction de dimensionnalité sont linéaires. PCA (Principal Component Analysis) [7] et LDA (Local Discriminant Analysis) [8] sont des méthodes de sous espace les plus utilisées pour la réduction de dimensionnalité. Les deux méthodes PCA et LDA sont linéaires, faciles à comprendre et très simples à mis en œuvre, mais elles ne conduisent pas de bons résultats dans de nombreux applications en monde réel grâce à leurs linéarités ou bi-dimensionnalité. Elles projettent les données multidimensionnelles de visage dans un sous espace bi-dimensionnelle (vecteur). Par conséquent, le processus de vectorisation détruit la structure naturelle et la corrélation dans les données originales. En plus, la vectorisation réduit les redondances et/ou les dépendances d'ordre supérieur présentes dans

les ensembles de données originales, et perd des représentations potentiellement plus compactes ou utiles qui peuvent être obtenues dans les formes de tenseur d'origine.

Par conséquent, les méthodes travaillant directement sur les représentations des tenseurs sont apparues comme des approches prometteuses. Récemment, il y a eu plusieurs tentatives pour développer des algorithmes subspatiaux multilinéaires pour l'extraction et la classification des caractéristiques des objets tensoriels. MPCA (Multilinear Principal Component Analysis) [9] et MDA (Multilinear Discriminant Analysis) [10] sont les générations des algorithmes linéaires PCA et LDA respectivement. La structure des algorithmes MPCA et MDA se base sur l'algèbre multilinéaire qui a l'avantage de pratiquer directement sur les données tensorielles sans besoins de les remodeler en vecteur 1D. En plus, la structure géométrique du données tensorielles reste naturelle sans modifiées ou détruites.

En particulier, l'analyse des caractéristiques faciales par les méthodes multilinéaires améliore généralement le taux de reconnaissance faciale et élimine plusieurs contraintes telles que : variations d'illumination, expressions faciales et variations de pose.

D'autre part, la plupart des techniques de reconnaissance faciale proposés dans la littérature sont basées sur des images faciales 2D ou 3D. Généralement, les applications de reconnaissances faciales 2D ou 3D souffrent de divers facteurs qui rendent ces applications loin d'être fiables en conditions réelles. Dans les conditions non contrôlées, les applications faciales 2D atteignent un faible taux de reconnaissance grâce à l'influence de changement d'éclairage et les variations de pose. Les application faciales 3D sont plus robustes que les application faciales 2D, mais elles aussi sont influencées par l'expressions faciale. Donc, la fusion multimodale des images faciales 2D+3D rend le système de reconnaissance faciale plus robuste et plus performant que le système uni-modale 2D ou 3D car les images faciales 2D et 3D sont complémentaires.

L'extraction des caractéristiques faciales est une étape principale dans le système de reconnaissance automatique de visages. Les descripteurs locaux jouent efficacement sur les petites textures et motifs des images faciales. Contrairement, aux descripteurs holistiques qui permettent la description globale et ignore les petites détails de l'image faciale.

A travers de notre vision pour les problèmes des applications de reconnaissances faciales, nous essayerons dans les chapitres de notre thèse de prouver l'intérêt des techniques de réduction de dimensionalités basées sur les tenseurs, l'intérêt de multimodalité et également l'intérêt des descripteurs locaux pour l'extraction des caractéristiques de visages.

1.3 Objectifs et contributions

La recherche des solutions optimales et efficaces pour les contraintes des systèmes de reconnaissances faciales est le but principal pour la plupart des chercheurs. Notre travail vise à développer et améliorer les performances d'un système de reconnaissance automatique de visage unimodale 2D et 3D ou multimodale 2D+3D. d'abord, l'utilisation des techniques d'apprentissage de sous espace multilinéaires (tenseurs) sont mieux adapter pour représenter les données faciales multidimensionnelles. D'autre part, on a choisi une stratégie efficace d'analyse multi-échelles pour présenter les caractéristiques faciales par des descripteurs locaux. La fusion multimodale des caractéristiques faciales 2D et 3D améliore efficacement le taux de reconnaissance de notre système. Enfin, l'évaluation de notre système testé sur trois bases de données populaires collectées dans diverses conditions tels que : les variations d'éclairage, les expressions faciales et les variations de poses. La comparaison avec l'état de l'art va juger la performance du système proposé.

Les étapes principales de notre travail sont :

- Etude de l'état de l'art de diverses techniques de reconnaissance faciale 2D, 3D et 2D+3D basées sur l'analyse multidimensionnelles.
- Etude et conception d'un nouveau système de vérification automatique de visage unimodale 2D, 3D et multimodale 2D+3D basé sur la représentation tensorielle. Les images faciales d'entrée 2D et 3D sont disposées et organisés sous forme d'un tenseur avec trois modes. Chaque mode est un sous-espace linéaire contenant des informations telles que : les caractéristiques faciales extraites, les histogrammes de différents descripteurs locaux, les échantillons ou personnes de la base de données, ...etc.
- Proposition d'une approche multilinéaire conçue pour la réduction multidimensionnelle, appelée MDA (Multilinear Discriminant Analysis) couplée avec la technique de la normalisation de covariance intra-classe WCCN [11]. La technique WCCN est proposée pour la reconnaissance de locuteur basée sur les machines à vecteurs de support SVM [12] où les vecteurs de caractéristiques d'entrées sont représentés dans un espace de grande dimension. Nous avons inspiré cette idée pour appliquer la technique de WCCN dans le sous espace multilinéaire pour traiter le problème de classification de l'approche MDA. L'intégration de deux techniques MDA et WCCN minimiserait effectivement les erreurs de classification de l'approche MDA et également augmenterait le taux de vérification du système proposé.

- Etude de l'efficacité des descripteurs locaux LPQ (Local Phase Quantizer) et BSIF (Binirized). Les caractéristiques faciales sont bien extraites en utilisant l'analyse multi-échelles et la concaténation des histogrammes qui sont moduler en tenseur d'ordre trois.
- Une phase d'évaluation permet de tester notre algorithme avec trois grandes bases de données FRGC V2.0, Bosphorus et CASIA 3D. Notre algorithme MDA+WCCN est comparé avec l'algorithme MDA pour prouver l'avantage d'intégralité de MDA et WCCN, ainsi que l'amélioration de taux de vérification. La comparaison finale avec les méthodes similaires permet le jugement de l'efficacité de notre l'algorithme.

1.4 Modes de fonctionnement d'un système biométrique

Le système biométrique peut être fonctionner en deux modes, identification (reconnaissance) ou vérification (authentification).

- **Identification (1 : N) :** consiste à comparer une donnée biométrique d'un individu avec celle d'un grand nombre de personnes enregistrées dans une base de données. L'identification permet à répondre à la question « *Qui est cette personne ?* ».

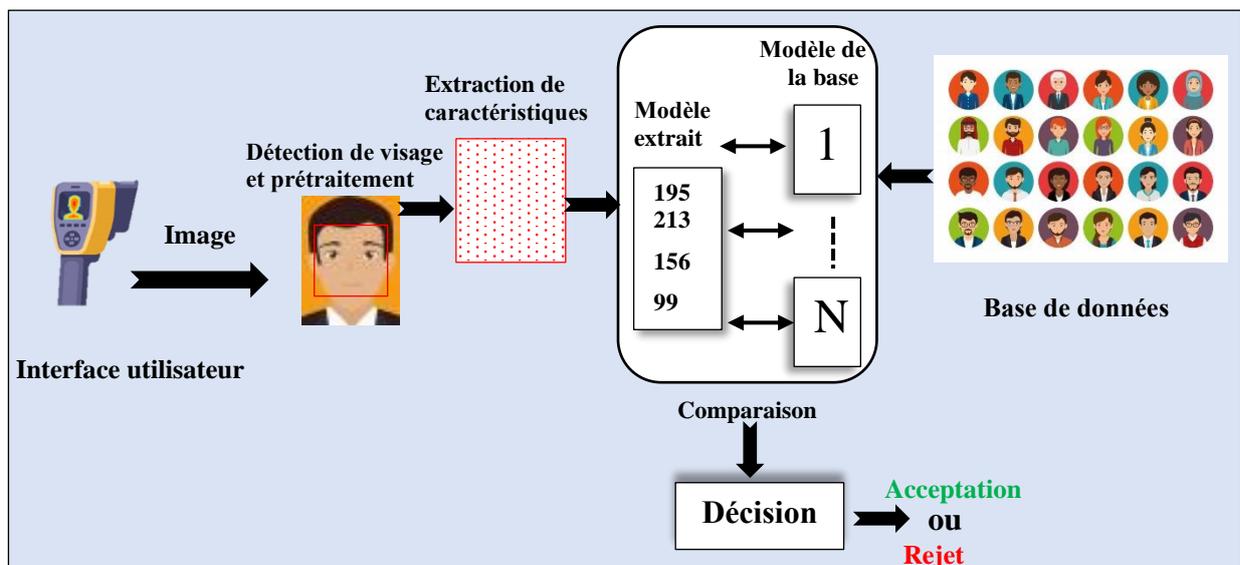


Figure 1.3 : système biométrique de reconnaissance faciale en mode identification

L'identification est généralement utilisée par les organisations de sécurité afin d'identifier les criminels et ce à leur insu. La **Figure 1.3** représente un système biométrique de reconnaissance faciale qui opère en mode identification.

- **Vérification (1 :1) :** consiste à vérifier une donnée biométrique d'un individu enregistré déjà dans une base de données (voir la **Figure 1.4**). La vérification permet de répondre à la question « *Est-ce bien Mr. X ?* ». La vérification est généralement

utilisée par les entreprises et compagnies commerciales afin de vérifier l'identité de leurs clients.

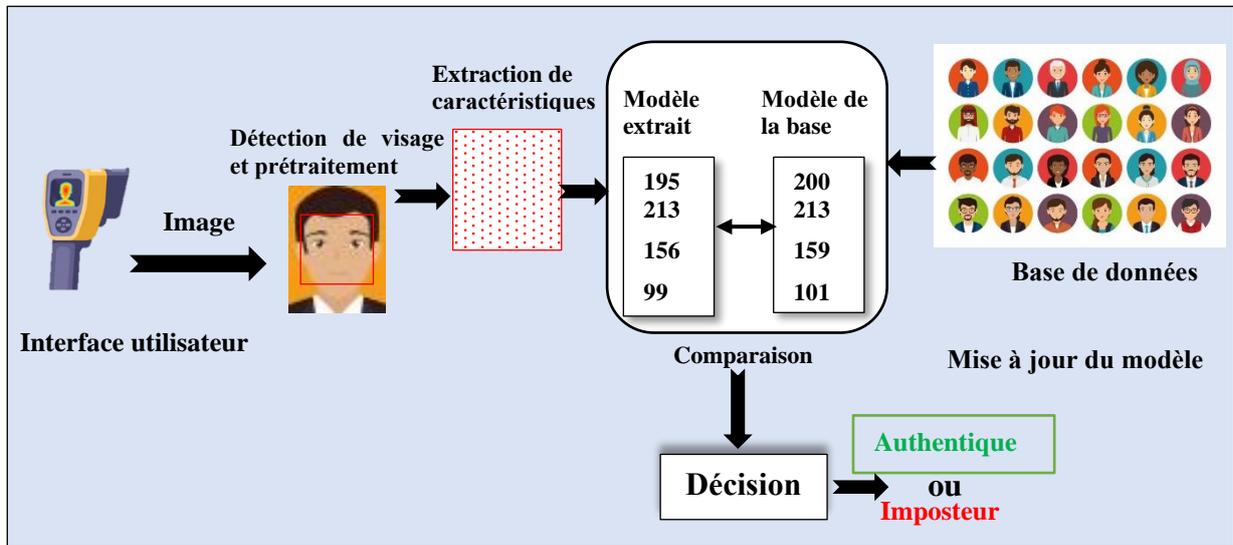


Figure 1.4 : système biométrique de reconnaissance faciale en mode vérification.

1.5 Evaluation des performances d'un système biométrique

Pratiquement, on ne trouve pas un système biométrique parfaitement idéal. Il y a des erreurs qui apparaissent durant la phase de l'identification ou la vérification. A partir de la mesure de ces erreurs, on peut évaluer notre système biométrique. Les principaux critères utilisés pour évaluer un système biométrique sont définis ci-dessous :

1. **Taux de Faux de Rejet (False Rejection Rate ou FRR)** : Indique la probabilité qu'un utilisateur connu soit rejeté par le système biométrique.
2. **Taux de Fausse Acceptation (False Accept Rate ou FAR)** : Indique la probabilité qu'un utilisateur inconnu soit identifié comme étant un utilisateur connu. Une illustration du **FRR** et du **FAR** est représentée par la **Figure 1.5**.

3. **Taux d'Égale Erreur (Equal Error Rate ou EER)** : Donne un point pour lequel le FAR est égal au FRR, c'est-à-dire le meilleur compromis entre les faux rejets et les fausses acceptations.

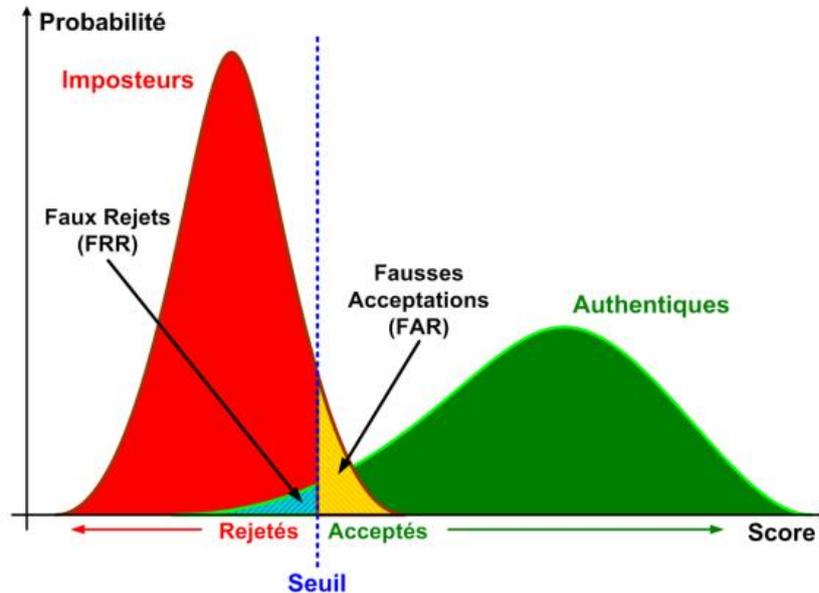


Figure 1.5 : illustration du taux FRR et taux FAR.

4. **La courbe ROC (Receiver Operating Characteristic)** : la courbe ROC représente l'évolution du FRR en fonction du FAR (voir la **Figure 1.6**). L'étude de cette courbe permet de déterminer les performances d'un système biométrique en mode vérification ou authentification.

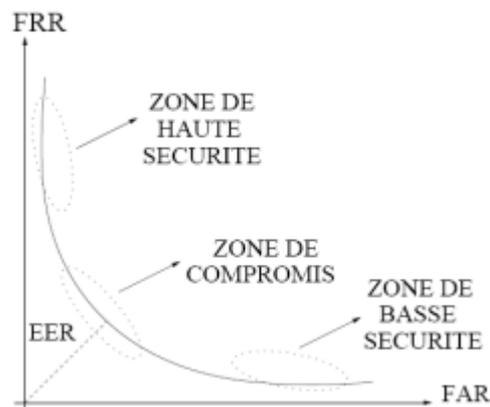


Figure 1.6 : la courbe ROC

5. **La courbe CMC (Cumulative Match Characteristic)** : exprime le pourcentage de personnes reconnues en fonction d'une variable que l'on appelle le rang. Plus le rang

augmente, plus le taux de reconnaissance correspondant diminue. La courbe CMC est utilisée lorsque le système biométrique opère en mode identification.

1.6 Organisation de la thèse

Notre thèse concerne la reconnaissance faciale par la fusion multimodale de visage. Pour rédiger cette thèse, nous avons suivi le plan suivant :

Le premier chapitre est une partie introductive, présente le contexte et la motivation qui nous permettent de choisir le thème de cette thèse. Ainsi que les objectifs et les contributions de notre travail. Dans ce chapitre, nous avons donné une vision globale sur les méthodes biométriques actuellement utilisées, ainsi que leurs intérêts dans le domaine de sécurité par rapport aux méthodes classiques. En particulier, nous présentons la reconnaissance faciale comme la biométrie la plus favorable parmi les autres biométries grâce à ses avantages. Ensuite, nous avons présenté le principe de fonctionnement d'un système biométrique de la reconnaissance faciale pour identifier ou authentifier un individu. Enfin, les critères d'évaluation de la performance d'un système de la reconnaissance faciale sont bien définis et expliqués.

Le deuxième chapitre représente une vision détaillée sur les techniques proposées et développées dans la reconnaissance faciale 2D, 3D et la fusion multimodale 2D+3D au cours de ces dernières années. A travers ce chapitre, nous essayons de classer ces méthodes de manière plus claire et utile afin de présenter les caractéristiques de chaque catégorie, ainsi que les avantages et les inconvénients de chacune d'entre elles. Alors, l'étude critique de ces méthodes nous permettra par la suite, de mettre en évidence les critères d'évaluation qui vont guider le choix des solutions retenues dans le cadre de notre travail.

Le troisième chapitre est consacré à la représentation de l'apprentissage de tenseur. En premier temps, nous donnons quelques notations et concepts du tenseur. Ensuite, nous présentons les outils essentiels de l'algèbre multilinéaire de base. En outre, nous démontrons les différentes projections possibles de tenseur : projection de vecteur à vecteur, projection tenseur à vecteur et projection tenseur à tenseur. Enfin, nous présentons en détail les méthodes multilinéaires utilisées par notre algorithme de reconnaissance. Les méthodes sont MWPCA (Multilinear Whitening Principal Component Analysis), MDA (Multilinear Discriminant Analysis) et WCCN (Within Class Covariance Normalization).

Le quatrième chapitre est dédié à notre approche proposée pour la vérification multimodale de visage 2D et 3D. En premier temps, nous présentons le diagramme du système de vérification

et algorithme de notre approche. Ensuite, nous expliquons la phase de vérification étape par étape, à savoir : le prétraitement de visages 2D et 3D, extraction des caractéristiques locaux de visages 2D et 3D, la réduction de dimensionalité et la classification. Ainsi que, nous décrivons comment ajuster les différents paramètres nécessaires de chaque étape. Enfin, les résultats obtenus par notre approche sont discutés et comparés avec les résultats des approches de l'état de l'art.

Dans **le cinquième chapitre**, nous terminons la rédaction de cette thèse par une conclusion générale, résumant nos contributions principales et les perspectives envisageables à la future.

Chapitre 2

Etat de l'art sur la reconnaissance faciale

2.1 Introduction.....	12
2.2 Reconnaissance de visage 2D	12
2.2.1 Approches globales.....	13
2.2.1.1 Les techniques linéaires.....	13
2.2.1.2 Les techniques non linéaires :	19
2.2.2 Approches locales	21
2.2.2.1 Techniques basées sur l'apparence locale.....	21
2.2.2.2 Techniques basées sur les points d'intérêt	23
2.2.3 Approches hybrides	24
2.2.4 Comparaison entre les techniques de reconnaissance faciale 2D	25
2.3 La reconnaissance de visage 3D	30
2.3.1 Approches globales.....	30
2.3.1.1 Les méthodes de sous-espace	30
2.3.1.2 Méthodes basées sur l'alignement de surfaces	31
2.3.1.3 Méthodes basées le modèle morphable (déformable)	35
2.3.2 Approches locales.....	38
2.3.2.1 Méthodes basées sur le calcul des points clés (Keypoint).....	38
2.3.2.2 Méthodes basées sur le calcul de courbure.....	41
2.3.2.3 Méthodes locale basée sur la surface.....	44
2.3.3 Discussion.....	47
2.4 La reconnaissance de visage 2D+3D	49
2.5 Conclusion	52

Chapitre 2 : Etat de l'art sur la reconnaissance faciale

2.1 Introduction

La reconnaissance de visage est une biométrie très intéressante, elle a attiré l'attention de plusieurs chercheurs depuis dizaines d'années. Récemment, la biométrie de la reconnaissance faciale a connu un grand développement soit en termes de matériels ou en termes d'algorithmique.

Ce chapitre représente brièvement un état de l'art des techniques utilisées dans la reconnaissance de visage. Le but de ce chapitre est de donner un panorama des méthodes les plus significatives en reconnaissance de visages 2D, 3D et 2D+3D, et aussi afin d'identifier les problèmes et les orientations pour les recherches futures dans ce domaine.

2.2 Reconnaissance de visage 2D

Historiquement, Les systèmes de reconnaissance faciale 2D sont les premiers systèmes qui sont proposés. Ils sont reconnus par leurs succès dans plusieurs applications qui s'installent dans des endroits contrôlés. Toutefois, il existe de nombreux algorithmes capables d'atteindre un taux de reconnaissance supérieure à 90%. Cependant, il y a des contraintes qui permettent la détection de visages en temps réels très difficiles. Nous pouvons limiter ces obstacles à quatre facteurs les plus influents [13]

- *Les variations d'éclairage(illumination)* : l'apparence d'un visage dans une image d'intensité ou couleur varie énormément en fonction de l'illumination de la scène lors de la prise de vue. En effet, elles entraînent parfois une différence d'image plus grande que le changement d'identité.
- *Les expressions faciales* : Le comportement d'un individu s'exprime généralement par des émotions (joie, peur, colère, tristesse, dégoût, surprise, honte) qui apparaissent sur le visage. En effet, le mouvement des caractères faciaux implique un changement dans un ou plusieurs zones du visage (déformation). Ces déformations faciales permettent la reconnaissance d'un individu probablement difficile.
- *La variation de pose de la tête* : La variation de pose se réfère principalement à la rotation hors du plan. C'est un défi pour les systèmes de reconnaissance faciale en raison de la nature d'une image faciale 2D ou 3D. Parfois, Les différences dans les images causées par le changement de pose sont plus grandes que les différences d'image entre

les personnes. Pour le moment les approches locales restent les plus robustes contre les variations de poses par rapport aux approches holistiques.

- *Occlusions* : Le visage peut être partiellement masqué par des objets dans la scène, ou par le port d'accessoires tels que lunettes de soleil, chapeaux, foulards...etc. Par conséquent, les occlusions partielles peuvent changer considérablement l'apparence originale d'une image faciale, et aussi elles peuvent également détériorer les performances des systèmes automatiques de reconnaissance de visage.

Dans la littérature, il existe de nombreuses méthodes de reconnaissance faciale 2D. On peut les classer en trois groupes : méthodes globales (ou holistiques), méthodes locales et méthodes hybrides. Chaque groupe de ces méthodes a une stratégie spécifique pour traiter les images faciales 2D comme nous s'expliquons ci-dessous.

2.2.1 Approches globales

Les approches holistiques ou approches de sous-espace sont proposées pour traiter les images faciales d'une manière globale, elles ne nécessitent aucune extraction des caractéristiques des zones dans le visage comme le nez, les yeux, ...etc. En plus, ces approches représentent les images faciales comme des matrices en pixels, généralement transformées en vecteurs plus faciles à manipuler et projeter dans un espace de faible dimension. Toutefois, les approches globales sont rapides à mettre en œuvre mais généralement elles sont sensibles aux conditions de luminosité, de pose et d'expressions faciales.

Les approches globales (ou holistiques) peuvent se subdiviser en deux catégories : les techniques linéaires et les techniques non-linéaires.

2.2.1.1 Les techniques linéaires

Les techniques linéaires permettent la projection linéaire d'une image d'entrée des visages de grand espace à un espace de dimension plus faible. Cependant, deux inconvénients majeurs se produisent durant la projection. D'abord, la sensibilité aux variations de visage non convexes, et aussi la sensibilité aux variations de luminosité. Ainsi, l'utilisation de distances euclidiennes ne sont pas toujours efficaces pour classer les formes et les individus « visages » et « non visages ». Par conséquent, la détection et les taux de reconnaissance de ces méthodes sont généralement insatisfaisants. Plusieurs techniques peuvent être classées comme techniques linéaires :

a) *Eigenfaces* [7] : est une approche populaire plus utilisée pour l'extraction des caractéristiques d'images faciales. Elle se base sur la technique d'Analyse en Composante

Principale (ACP ou PCA en anglais). L'idée principale de PCA (Principal Component Analysis) est de transformer un certain nombre de variables éventuellement corrélés en un petit nombre de variables décorrélés appelées « composantes principales ». La **Figure 2.2** démontre comment représenter le visage par un petit nombre de caractéristiques. Le PCA permet de calculer les vecteurs propres de la matrice de covariance d'une image faciale, et les projette sur un sous espace de faible dimension. Seules les plus grandes valeurs propres associées à vecteurs propres (visages propres) sont réservées. La **Figure 2.1** représente les vecteurs propres des 12 plus grandes valeurs propres pour la base de données ORL.



Figure 2.1 Vecteurs propres (Eigenfaces) des 12 plus grandes valeurs propres obtenus à partir de la base de données ORL [14].

Soit l'ensemble de l'images d'apprentissage $\{X_1, X_2, X_3, \dots, X_n\}$ de dimension $M \times N$, alors la taille de l'image est 4×4 peut également considérer un vecteur de dimension 16. Donc, la moyenne de l'ensemble d'images est calculée par la formule suivante :

$$X = \frac{1}{n} \sum_{i=1}^n X_i \quad (2.1)$$

La matrice de covariance C estimée pour représenter les premiers vecteurs propres est défini par la relation suivante :

$$C = \frac{1}{n} \sum_{i=1}^n (X - X_i) (X - X_i)^T \quad (2.2)$$

Les vecteurs propres et les valeurs propres correspondantes sont calculés en utilisant

$$KV = \lambda V, \quad (V \in R_n, V \neq 0) \quad (2.3)$$

où V est l'ensemble de vecteurs propres de la matrice C associée à sa valeur propre λ . La projection de tous les images d'apprentissage du $n^{\text{ième}}$ personne dans le sous-espace propre correspondant :

$$y_k^i = w^T(x_i), \quad (i = 1, 2, 3, \dots, n) \quad 2.4)$$

où y_k^i sont les composantes principales (eigenfaces).

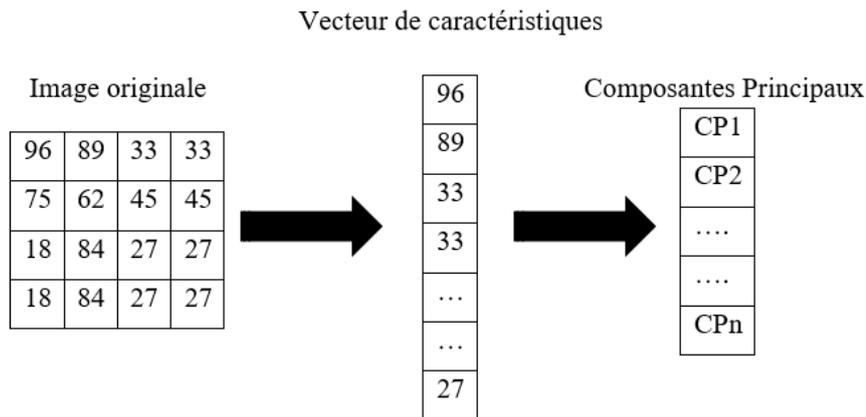


Figure 2.2 Exemple d'un vecteur de caractéristiques réduit par PCA.

La projection d'une image dans un sous espace par l'analyse composantes principales PCA réduit effectivement sa dimension. Cependant, les informations de voisinage peuvent être perdues durant la transformation de l'image en vecteur. Une méthode PCA bi-dimensionnelle (2-D PCA) est proposée pour prendre en entrée des images plutôt que des vecteurs. La **Figure 2.3** illustre certaines images reconstruites d'une image de la base de données ORL en ajoutant au premier nombre des vecteur propres d ($d = 2, 4, 6, 8, 10$) des sous-images en même temps. Les images reconstruites apparaissent plus clairement quand le nombre de sous-images est augmenté. Aussi PCA (eigenfaces) est utilisée pour représenter et reconstruire la même image de visage, où le résultat a été insatisfait pour reconstruire cette image.

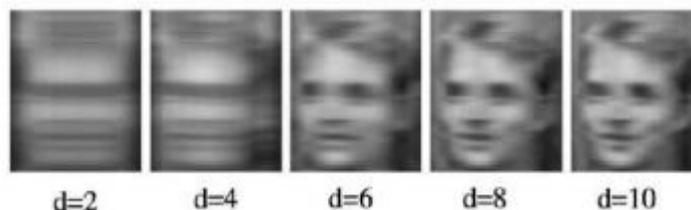


Figure 2.3 Images reconstruites par la méthode 2D-PCA [15].

b) Fisherfaces [8] [16] : est une méthode qui a le même principe que *Eigenfaces*. Elle se base sur la méthode LDA (Linear Discriminant Analysis). Le but de cette méthode est de réduire l'image facial d'un espace de grande dimension à un espace de faible dimension en utilisant la technique LDA. LDA construit un sous espace discriminant pour distinguer de manière optimale les visages de différentes personnes. Soit $X = \{x_1, x_2, \dots, x_N\}$ l'ensemble des échantillons d'apprentissages appartenant aux classe i . Les deux matrices de dispersion inter-classe S_B et intra-classe S_W sont définies comme suit :

$$S_B = \sum_{l=1}^C M_l (x_l - \mu) ((x_l - \mu))^T \quad 2.5$$

$$S_W = \sum_{l=1}^C \sum_{x_k \in X_l} M_l (x_k - \mu) M_l (x_k - \mu)^T \quad 2.6$$

où μ est le vecteur moyen des échantillons appartenant à la classe i , x_k représente le nombre de l'image numérique dans la classe i . C est le nombre de classes distinctes et M_l représente le nombre des échantillons d'apprentissage appartenant aux classe i . pour une bonne classification par la technique LDA, le rapport $\frac{\det\|S_B\|}{\det\|S_W\|}$ doit être au maximum, c'est-à-dire minimiser la matrice intra-classe S_W et maximiser la matrice inter-classe S_B . La **Figure 2.4** représente les 12 premiers vecteurs (Fisherfaces) obtenus à partir de la base de données ORL.

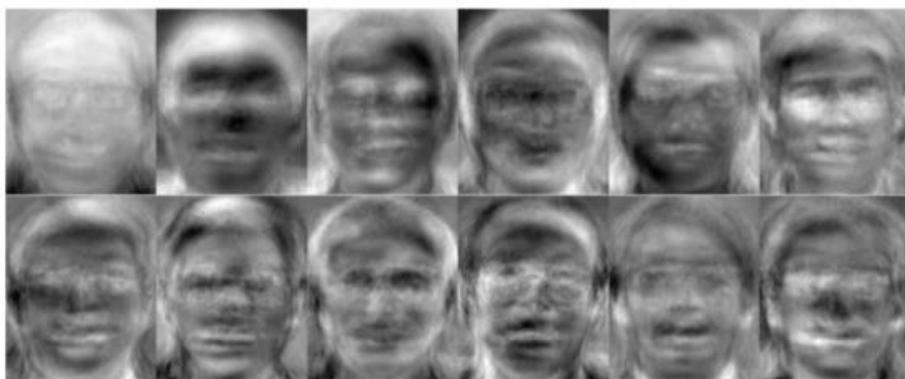


Figure 2.4 Les 12 premiers vecteurs (Fisherfaces) obtenus à partir de la base de données ORL [14].

c) **L'Analyse en Composantes Indépendantes ICA** (independent component analysis) [17]: elle a été introduite par *COMON, Pierre* dans [18]. ICA est une méthode de calcul de statistiques, des réseaux de neurones et du traitement du signal qui est un cas particulier de séparation aveugle de source [19] mais par suite a été appliquée à divers problèmes tel que la reconnaissance de visage [17, 20]. Le but de cette technique est d'effectuer une transformation linéaire afin de réduire la dépendance statistique entre les différents vecteurs de base. Elle consiste à exprimer un ensemble de N variables aléatoires x_1, x_2, \dots, x_n sous forme de combinaison linéaire de N variables aléatoires statistiquement indépendantes s_j , telle que :

$$x_j = a_{j,1}s_1 + a_{j,2}s_2 + \dots + a_{j,n}s_n \quad (2.7)$$

ou se forme matricielle : $x = As$.

La **Figure 2.5** représente des vecteurs de base ICA provenant de la base de données ORL.

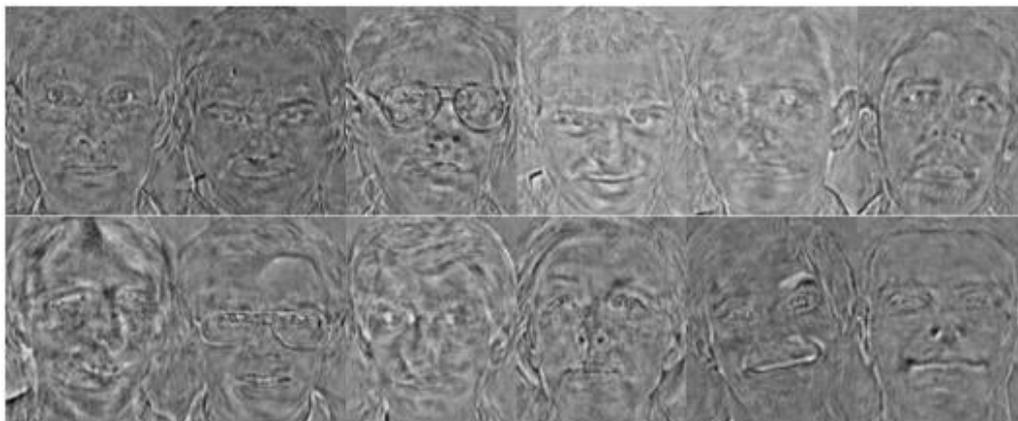


Figure 2.5 Exemple de vecteurs de base ICA provenant de la base de données ORL [14].

d) **Améliorations des techniques PCA, LDA et ICA** : pour améliorer les techniques de sous-espace linéaire, de nombreuses recherches sont développées. *Pentland et al* [21] ont amélioré la technique PCA pour gérer les variations de pose du visage. Le sous-espace probabiliste a été introduit pour fournir une mesure de similarité plus significative dans le cadre probabiliste. En outre, *Lu et al* [22] ont présenté une combinaison entre le D-LDA (LDA directe) et le F-LDA (LDA fractionnel), une variante du LDA dans laquelle les fonctions pondérées sont utilisées pour éviter les erreurs de classification causées par des produits de catégories trop proches. Dans [23], l'auteur a proposé une approche basée sur le tenseur

multilinéaire de décompositions des ensembles pour éviter le problème de la confusion de plusieurs facteurs liées au même système de reconnaissance faciale, tels que l'illumination et la pose.

e) Le Multidimensional Scaling (MDS)[24] : il s'agit d'une autre méthode notable de réduction de la dimension. Plutôt que de conserver la variance de données durant la projection, elle s'efforce de protéger toutes les distances entre chaque paire de modèles $dist(x_i, x_j)$ en cherchant une transformation linéaire qui minimise l'énergie. Ce problème de minimisation peut être résolu par la décroissance des valeurs propres. En utilisant la distance euclidienne entre les données, les sorties de MDS sont les mêmes que les sorties de PCA.

f) la factorisation de matrices non négatives (NMF) [25] : elle est une autre méthode qui représente le visage sans utiliser la notion de classe. L'algorithme de la factorisation matricielle non négative (NMF), tel que PCA traite le visage comme une combinaison linéaire de vecteurs de la base d'espace réduit. La différence est que NMF ne permet pas les éléments négatifs dans les vecteurs de la base dans la combinaison poids. En autres termes, certains vecteurs dans l'espace réduit par PCA (visages propres) ressemblent à des versions distordues du visage entier, tandis que celles réduites par NMF sont des objets localisés qui reflètent mieux à la notion de parties du visage.

g) Transformée discrète en ondelettes (DWT)[26] : Une autre technique linéaire utilisée pour la reconnaissance faciale. Dans les travaux de [26], les auteurs ont utilisé une méthode de transformée en ondelettes discrètes bidimensionnelle (2D-DWT) pour la reconnaissance faciale en utilisant une nouvelle stratégie de patch. Une stratégie de patch non uniforme pour la sous-bande basse fréquence de niveau supérieur est proposée en utilisant une technique de projection intégrale pour deux sous-bandes haute fréquence de niveau supérieur de 2D-DWT basée sur l'image moyenne de tous les échantillons d'apprentissage. Cette stratégie de patch est meilleure pour conserver l'intégrité des informations locales et est plus appropriée pour refléter la caractéristique de structure de l'image du visage. Lors de la construction de la stratégie de correction à l'aide des échantillons de test et d'apprentissage, la décision est prise à l'aide du classificateur voisin. De nombreuses bases de données sont utilisées pour évaluer cette méthode, notamment Labeled Faces in Wild (LFW), Extended Yale B, Face Recognition Technology (FERET) et AR.

De plus, il y a un grand nombre de techniques linéaires qui ont également été utilisées pour calculer les vecteurs de caractéristiques telles que :

- Analyse discriminante régularisée (RDA) [27]
- Régression LDA (RLDA) [28]
- LDA à espace nul (NLDA) [29]
- LDA à double espace [30]
- Transformée en cosinus discrète (DCT) [31]
- Décomposition généralisée des valeurs singulières [32]
- Boosting LDA [33]
- Analyse des caractéristiques locales discriminantes [34]
- Block LDA [35]
- Modèle discriminant linéaire de Fisher amélioré (EFM) [36]
- LDA incrémentale [37]
- Vecteurs communs discriminants (DCV) [38]
- Analyse discriminante bilinéaire (BDA) [39]

2.2.1.2 Les techniques non linéaires :

Les approches linéaires offrent une représentation fidèle aux données d'entrées lorsque les données ont des structures linéaires. Cependant, lorsque les données ne sont pas linéaires, les chercheurs s'adaptent une solution, c'est d'utiliser une fonction appelée « noyau » pour construire un grand espace dans lequel le problème devient linéaire.

Ainsi, des techniques linéaires de réduction de dimensionnalité peuvent être appliquées lorsque la structure intrinsèque des données ne sont pas linéaires. Ces méthodes utilisent typiquement « l'astuce du noyau (Kernel) », qui propose que tout algorithme, formulé avec une fonction noyau, puisse être reformulé avec une autre fonction noyau.

Un processus courant consiste à exprimer la méthode avec un produit scalaire à l'aide d'une fonction noyau. L'astuce du noyau permet de travailler dans l'espace transformé sans avoir calculer explicitement l'image de chaque donnée. Dans ce contexte, plusieurs approches non linéaires ont été proposées :

a) PCA à Noyaux (Kernel Principal Component Analysis, KPCA) [40]: c'est la version améliorée de PCA, qui utilise des fonctions à noyau. KPCA calcule les vecteurs propres (Eigenfaces) de la matrice à noyau, tandis que PCA calcule la matrice de covariance. En

addition, KPCA est la représentation de la technique PCA dans un espace de grande dimension mappé par la fonction à noyau. Un algorithme de KPCA est utilisé pour calculer la fonction de la matrice à noyau K de distribution constituée de n points de données $x_i \in R^d$. Après les points de données sont mappés dans un espace de grand de dimension f comme indiqué l'algorithme 1 en dessous.

Généralement, la performance de la technique KPCA dépend du choix de la matrice du noyau K . Le noyau gaussien ou polynomial sont des noyaux linéaires [41] typiquement utilisés. KPCA a été utilisé avec succès pour résoudre certains problème tels que la détection de nouveaux éléments d'un ensemble [42] et la reconnaissance vocale [43]. L'inconvénient principal de KPCA est que la taille de la matrice du noyau est le carré du nombre d'échantillons de l'ensemble d'entraînement, ce qui peut rapidement être prohibitif.

Algorithme1 : PCA à noyau (kernel PCA)

1. Déterminer le produit scalaire de la matrice K en utilisant la fonction à noyau $K_{i,j} = k(x_i, x_j)$.
 2. Calculer les vecteurs propres à partir de la matrice résultante K et normaliser avec la fonction : $\gamma k(\alpha k \alpha k) = 1$.
 3. Calculer la projection du point de test sur les vecteurs propres Vk à l'aide de la fonction à noyau : $kPCK(x) = (Vk\varphi(x)) = \sum_i^m \alpha k_i k(x_i, x)$.
-

b) Analyse discriminante linéaire à noyau (Kernel Linear Discriminant Analysis, KDA [44]) : la technique KLDA est une extension de noyau de la technique LDA linéaire, dans la même extension de noyau de PCA. L'auteur dans [44] a proposé un classificateur d'analyse discriminante du noyau (CS-KDA) binaire non linéaire spécifique à une classe basée sur l'analyse discriminante du noyau de régression spectrale. Il y a d'autres techniques non linéaires qui ont été également utilisées dans le cadre de la reconnaissance faciale :

- Support vector machine (SVM à noyaux)[45] .
- KICA (kernel independent component analysis) [46].
- Exponential discriminant analysis (EDA) [47].
- Locality preserving projection (LPP) [48].
- Local linear embedding (LLE) [49].

- Kernel Fisher discriminant analysis (KFD) and KPCA [50].
- Nonlinear locality preserving with deep networks [51].
- Local tangent space analysis (LTSA) [52].
- Maximum variance unfolding (MVU) [53].
- Laplacian eigenmaps [54].
- Evolutionary weighted principal component analysis (EWPCA) [55].
- Wavelet transform (WT), radon transform (RT), and cellular neural networks (CNN) [56].
- Nonlinear DCT and kernel discriminative common vector (KDCV) [57].
- Joint transform correlator-based two-layer neural network [58].
- Isomap [59].

Les différentes techniques de sous-espace discutées dans cette section sont introduites pour réduire la dimensionnalité et la complexité des étapes de détection ou de reconnaissance. Les techniques linéaires et non linéaires offrent une reconnaissance robuste dans différentes conditions d'éclairage et expressions faciales. Bien que ces techniques (linéaires et non linéaires) permettent une meilleure réduction des dimensionnalités et améliorent le taux de reconnaissance.

2.2.2 Approches locales

Dans le contexte de la reconnaissance faciale, les approches locales ne traitent que certains traits du visage. Elle sont plus sensible aux expressions faciales, aux occlusions et à la pose [60]. L'objectif principal de ces approches est de découvrir des traits distinctifs. Généralement, Elles peuvent être classer en deux catégories : 1) Méthodes de reconnaissance faciale basées sur l'apparence locale du visage qui sont utilisées pour extraire les caractéristiques locales, tandis que l'image du visage est divisée en petites régions (patches) [61]. 2) Méthodes basées sur les points d'intérêt qui sont utilisées pour détecter les points d'intérêt dans l'image du visage, après que les traits localisés sur ces points sont extraits.

2.2.2.1 Techniques basées sur l'apparence locale

Ces techniques permettent de représenter l'image faciale par un ensemble de vecteurs distincts de faible dimension ou de petites parties. Les techniques basées sur l'apparence locale consistent à localiser sur les points critiques du visage comme : le nez, la bouche, et les yeux pour générer plus de détails. Les caractéristiques couramment utilisées sont : les coefficients de Gabor [62], les ondelettes de Haar [63], la Transformées de Fourier, les caractéristiques basées

sur la méthode du modèle binaire local (LBP) [64], quantification de phase locale (LPQ) [65], WLD (Weber Law Descriptor) [66], BSIF (Binarized Statistical Image Features) [67] et les histogrammes de dégradés orientés (HOG) [68]. Dans cette section, nous avons présenté le descripteur LBP et ses variantes.

a. LBP et ses variantes :

LBP est une très connue pour extraire les caractéristiques locales de n'importe quel objet [69]. La méthode LBP est largement employée dans de nombreuses applications telles que la reconnaissance de visage, la reconnaissance des expressions faciales, la segmentation des textures et classification des textures. La technique LBP divise d'abord l'image faciale en des régions ou sous-blocs. Ensuite, le code LBP est calculé pour chaque sous-bloc. Chaque pixel est comparé à ses huit voisins dans un voisinage 3×3 en soustrayant la valeur du pixel central. Si la valeur d'un pixel voisin est inférieure à la valeur du pixel central, un zéro lui est attribué sinon, on lui en donne un. Le code binaire contient des informations sur la texture locale. Enfin, pour chaque sous block, un histogramme de ces codes est construit. Puis, l'ensemble des histogrammes de tous sous block sont concaténés pour former le vecteur de caractéristiques. Le LBP est défini dans une matrice de taille 3×3 , comme indiqué dans l'équation sous dessous.

$$LBP = \sum_{p=1}^8 2^p (i_0 - i_p), \quad \text{avec } s(x) = \begin{cases} 1 & x \geq 1 \\ 0 & x < 1 \end{cases} \quad (2.8)$$

où i_0 et i_p sont respectivement la valeur d'intensité du pixel central et des pixels de voisinage.

La **Figure 2.6** illustre un exemple pour calculer le code binaire LBP.

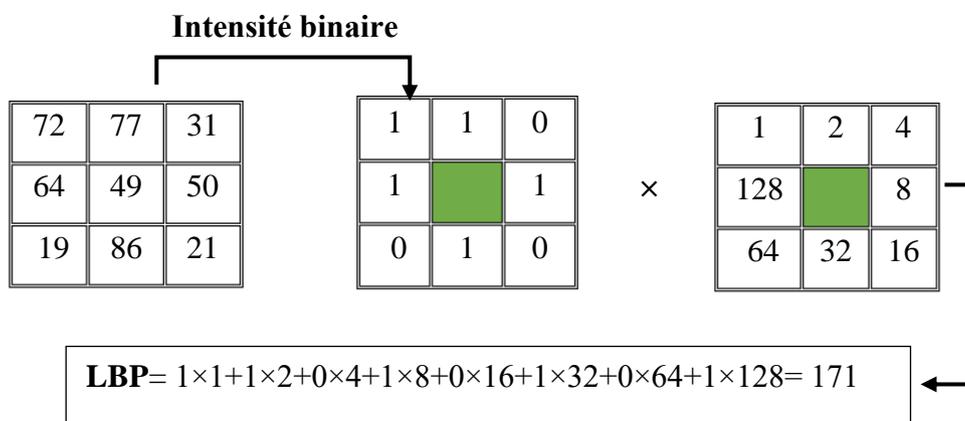


Figure 2.6 Exemple d'un code binaire LBP.

Plusieurs systèmes de reconnaissance faciales basés sur la technique LBP sont proposés. *Khoi et al* [70] ont proposé un système de reconnaissance faciale rapide basé sur la technique LBP et ses variantes : Rotation Invariante de LBP (RILBP) et la pyramide de LBP (PLBP). Une nouvelle technique non supervisée est proposée dans [71], appelée LBP network (LBPNet). Dans [72], l'auteur a proposé une méthode basée sur deux techniques LBP et K-NN (K- Nearest Neighbors). Cette méthode permet de diminuer l'influence des variations d'illumination, expression faciale et pose. *Bonnen et al* [73] ont proposé une variante de la méthode LBP, appelé LBP multi-échelles (MLBP). LTP (local ternary pattern) [74], ILBP amélioré (improved local binary patterns) [75] et LQP (local quantized pattern) [76] techniques sont d'autres variantes de LBP également proposées. La méthodologie de la LBP a été récemment développée avec un grand nombre de variations afin d'améliorer les performances de diverses applications. Ces variations concernent sur différents aspects de l'opérateur LBP d'origine qui sont : la capacité discriminante, la robustesse et la sélection des voisinages.

2.2.2.2 Techniques basées sur les points d'intérêt

Ces méthodes sont utilisées pour détecter des caractéristiques géométriques spécifiques, telles que la distance entre les yeux, la largeur de la tête, etc. Ces techniques sont définies par deux classes selon le point d'intérêt : la première classe concentre sur les performances des détecteurs des points caractéristiques du visage ; tandis que la deuxième classe concentre sur la représentation des informations incluses dans les caractéristiques du visage. Les techniques SIFT (Scale Invariant Feature Transform), BRIEF (Binary Robust Independent Elementary Features) et SURF (Speeded-Up Robust Features) sont largement utilisées pour détecter et représenter les caractéristiques d'image faciale.

Lenc et Kral [77] ont proposé un système de reconnaissance automatique de visage basé sur l'algorithme SIFT afin de détecter et présenter les caractéristiques locales du visage. L'idée principale de descripteur SIFT est de convertir l'image à une représentation composée des points d'intérêt. Ces points comportent les informations sur les caractéristiques du visage. Le descripteur SIFT est invarié par l'échelle et la rotation. Il est largement utilisé surtout dans les applications en temps réel, mais le temps d'appariement des points critiques est l'un de ses inconvénients.

Un autre descripteur et détecteur des caractéristiques du visage, appelé SURF est employé dans [78] . Le descripteur SURF est inspiré par l'idée de SIFT, mais l'évaluation

comparative effectuée dans [79] indique que le descripteur SURF est plus performant que SIFT en termes de répétabilité, distinction, temps d'exécution et robustesse.

Michael et al [80] ont proposé un descripteur, nommé BRIEF. Il est simple et rapide à calculer. Ce descripteur est basé sur les différences entre l'intensité des pixels qui sont similaires à la famille des descripteurs binaires tels que BRISK (Binary Robust Invariant Scalable) proposé dans [81] et le FREAK (Fast Retina keypoint) proposé dans [82] en termes d'évaluation. Le descripteur BRIEF a atteint une excellente performance et précision dans la reconnaissance de formes [79].

D'autres techniques basées sur les points d'intérêt sont également proposées :

- Architecture de liaison dynamique (DLA) [83].
- Le Elastic bunch graph matching (EBGM) [84].
- Vecteur de caractéristiques géométriques [62].
- Modèle statistique de visage [85].
- Extraction des caractéristiques par filtre de Gabor [86].

2.2.3 Approches hybrides

Les approches hybrides consistent à combiner les approches globales et les approches locales afin d'exploiter les avantages de deux approches. De nombreux chercheurs ont pris de fusionner les caractéristiques locales et globales des images faciales en choisissant différentes stratégies de combinaison.

Fathima et al [87] ont proposé une approche hybride pour la reconnaissance faciale. Cette méthode résulte de l'association des ondelettes de Gabor et la technique d'analyse discriminante linéaire (HGWLDA). D'abord, l'image en niveau de gris est normalisée, ensuite approximée et réduite en dimension. Cette image est convoluée avec une banque de filtres de Gabor avec des échelles et des orientations variables. Après, la technique de sous-espace 2D-LDA est utilisée pour maximiser l'espace inter-classe et minimiser l'espace intra-classe. Le classificateur k-nearest neighbour (k-NN) est employé pour reconnaître le visage de l'image de test. Les résultats expérimentaux montrent la robustesse de cette approche dans différentes conditions d'éclairage. *Barkan et al* [88] ont introduit une nouvelle représentation d'image faciale basée sur la technique over-complete LBP (OCLBP). Cette représentation est une version modifiée à plusieurs échelles de l'opérateur LBP. Ensuite, le système a employé l'analyse discriminante linéaire (LDA) couplée à la normalisation de la covariance intra-classe

(WCCN) afin d'améliorer la discrimination de LDA entre les classes. *Kim et al* [89] ont proposé un système de reconnaissance faciale robuste à la variation d'éclairage via le modèle directionnel local (LDP pour Local directional pattern) et l'analyse principale bidimensionnelle 2D-PCA. En particulier, cette approche utilise l'image transformée obtenue à partir du descripteur LDP comme image d'entrée directe de l'algorithme 2D-PCA, contrairement aux approches qui utilisent les modèles de descripteurs locaux pour acquérir les caractéristiques de l'histogramme. L'évaluation des performances du système proposé a été réalisée en utilisant deux bases de données Yale B et CMU-PIE. *Cho et al* [90] ont proposé une approche hybride qui utilise à la fois les caractéristiques locales et globales de l'image faciale. D'abord, La technique PCA est utilisée pour réduire la dimensionnalité. Puis, le descripteur LGBPHS (en anglais, Local Gabor Binary Pattern Histogram Sequence) est appliqué pour réaliser l'étape de reconnaissance. Enfin, Les résultats expérimentaux montrent un meilleur taux de reconnaissance par rapport aux techniques PCA et les ondelettes de Gabor sous des variations d'éclairage.

D'autres approches basées sur des modèles statistiques ont également été largement explorées. Ce modèle statistique formalise les relations entre les variables dans la forme d'équations mathématiques qui décrivent comment une ou plusieurs variables aléatoires sont liées à une ou plus de variables aléatoires. Ce modèle est considéré comme statistique lorsque les variables ne sont pas déterministes, mais liés de manière stochastique. Les techniques Hidden Markov model (HMM) [91], Active Appearance Model (AAM) [92], Gaussian Mixture Model (GMM) [93] et pseudo-2D HMM [94] sont les plus connues dans ce modèle.

2.2.4 Comparaison entre les techniques de reconnaissance faciale 2D

Dans cette section, nous avons présentés les différentes techniques proposées dans la reconnaissance de visage en 2D. généralement, les techniques existes peuvent être classées en trois catégories, les approches globales, les approches locales et les approches hybrides (voir la **Figure 2.7**). Le **Tableau 2.1** résume quelques avantages et inconvénients des approches de reconnaissance faciale 2D.

En particulier, les méthodes de reconnaissance effectuées sur des images statiques donnent de bons résultats sous différentes conditions d'éclairage et d'expression. Cependant, plusieurs méthodes nécessitent de nombreuses images d'apprentissage, ce qui limite leur utilisation dans les systèmes en temps réel, où le temps de réponse est un aspect important.

Le but principal des techniques telles que HOG, LBP, filtre de Gabor, BRIEF, SURF et SIFT est de découvrir les caractéristiques distinctives, qui peuvent être divisées en deux groupes : (1) les techniques basées sur l'apparence locale qui sont utilisées pour extraire les caractéristiques locales lorsque l'image faciale est divisée en petites régions ; et (2) les techniques basées sur des points clés, qui sont utilisées pour détecter les points d'intérêts dans l'image faciale. Dans le cadre de la reconnaissance faciale, les techniques locales ne traitent que certains traits du visage, ce qui les rend très sensibles aux expressions faciales et aux occlusions [95, 96]. La relative robustesse est l'avantage principale de ces techniques locales. De plus, ils prennent en compte la particularité du visage en tant que la forme naturelle pour reconnaître un nombre réduit de paramètres. Un autre avantage est qu'ils ont une capacité de compactage et une vitesse de comparaison élevées. Les principaux inconvénients de ces méthodes sont la difficulté d'automatiser la détection des traits de visage et le fait que la personne responsable de la mise en œuvre de ces systèmes doit prendre une décision arbitraire sur des points vraiment importants.

Contrairement aux approches locales, les approches globales aussi appelées méthodes basées sur l'apparence traitent l'image faciale globalement, c'est-à-dire il n'est pas nécessaire d'extraction des régions du visage ou des points caractéristiques (les yeux, la bouche, le nez, ...etc). Ainsi, une image de visage est représentée par une matrice de pixels, et cette matrice est souvent transformée en vecteurs de caractéristiques pour faciliter leur manipulation. Bien que ces approches soient faciles à mettre en œuvre, elles sont sensibles aux variations (poses, éclairage, expressions faciales et orientation). En effet, tout changement dans l'image du visage se traduit par un changement des valeurs de pixel. Comme mentionné précédemment, dans les méthodes globales, les données d'entrée de visage sont projetées plus tard dans un espace de faible dimension. L'avantage principale des approches globales est qu'elles ne détruisent pas les informations dans l'image faciale en concentrant uniquement sur les régions ou les points d'intérêts. Cependant, cette propriété représente un inconvénient car elle suppose que tous les pixels de l'image ont la même importance. En conséquence, ces techniques sont non seulement coûteuses en calcul, mais elles nécessitent également un degré élevé de corrélation entre les images de test et les images d'apprentissage. De plus, ces approches ignorent généralement les détails locaux, ce qui signifie qu'elles sont rarement utilisées pour identifier les visages.

Les approches hybrides sont basées sur les caractéristiques locales et globales au même temps pour exploiter les avantages des deux techniques. Ces approches combinent les deux

approches locales et globales dans un seul système pour améliorer les performances et la précision de la reconnaissance. Le choix de la méthode à utiliser doit tenir compte l'application dans laquelle elle est appliquée. Par exemple, dans les systèmes de reconnaissance faciale qui utilisent de très petites images, les méthodes basées sur les caractéristiques locales sont un mauvais choix. Une autre considération dans le processus de sélection d'algorithme est le nombre d'exemples d'entraînement. En fin, nous pouvons rappeler que la tendance est de développer des méthodes hybrides, mais ces méthodes sont très complexes et nécessitent plus de temps de traitement.

Tableau 2.1 Avantages et les inconvénients des approches de la reconnaissance faciale 2D [97].

Technique	Les avantages	Les inconvénients
Globales	<ul style="list-style-type: none"> ▪ Rapide à mettre en œuvre. ▪ Complexité moyenne ▪ Réduire la dimension des images. 	<ul style="list-style-type: none"> ▪ Très sensible aux variations d'éclairage, pose et expression faciale. ▪ Elles nécessitent une grande capacité de mémoire. ▪ Aucune conservation des variations de visages non convexes permettant de différencier les individus.
Locales	<ul style="list-style-type: none"> ▪ Fournir des informations supplémentaires basées sur les régions locales. ▪ Moins sensibles aux variations d'éclairage. ▪ Peuvent être utiles et efficaces pour la reconnaissance faciale où une image de référence est disponible. ▪ Capacité à choisir la meilleure façon de représenter les informations de chaque région. 	<ul style="list-style-type: none"> ▪ L'intégration d'informations globales est souvent requise. ▪ La détection et l'extraction des caractéristiques géométriques ne sont pas faciles en particulier lorsqu'il y a des occlusions, ou des variations de pose et des expressions faciales. ▪ Leur performance dépend fortement de l'efficacité de l'algorithme de localisation des points caractéristiques.
Hybrides	<ul style="list-style-type: none"> ▪ La combinaison de l'analyse globale et locale d'un visage peut améliorer la capacité du classificateur. ▪ Permet d'exploiter les complémentarités et offre plus des systèmes efficaces et reconnaissance plus rapide. 	<ul style="list-style-type: none"> ▪ Plus difficile à mettre en œuvre que les approches locales et approches globales.

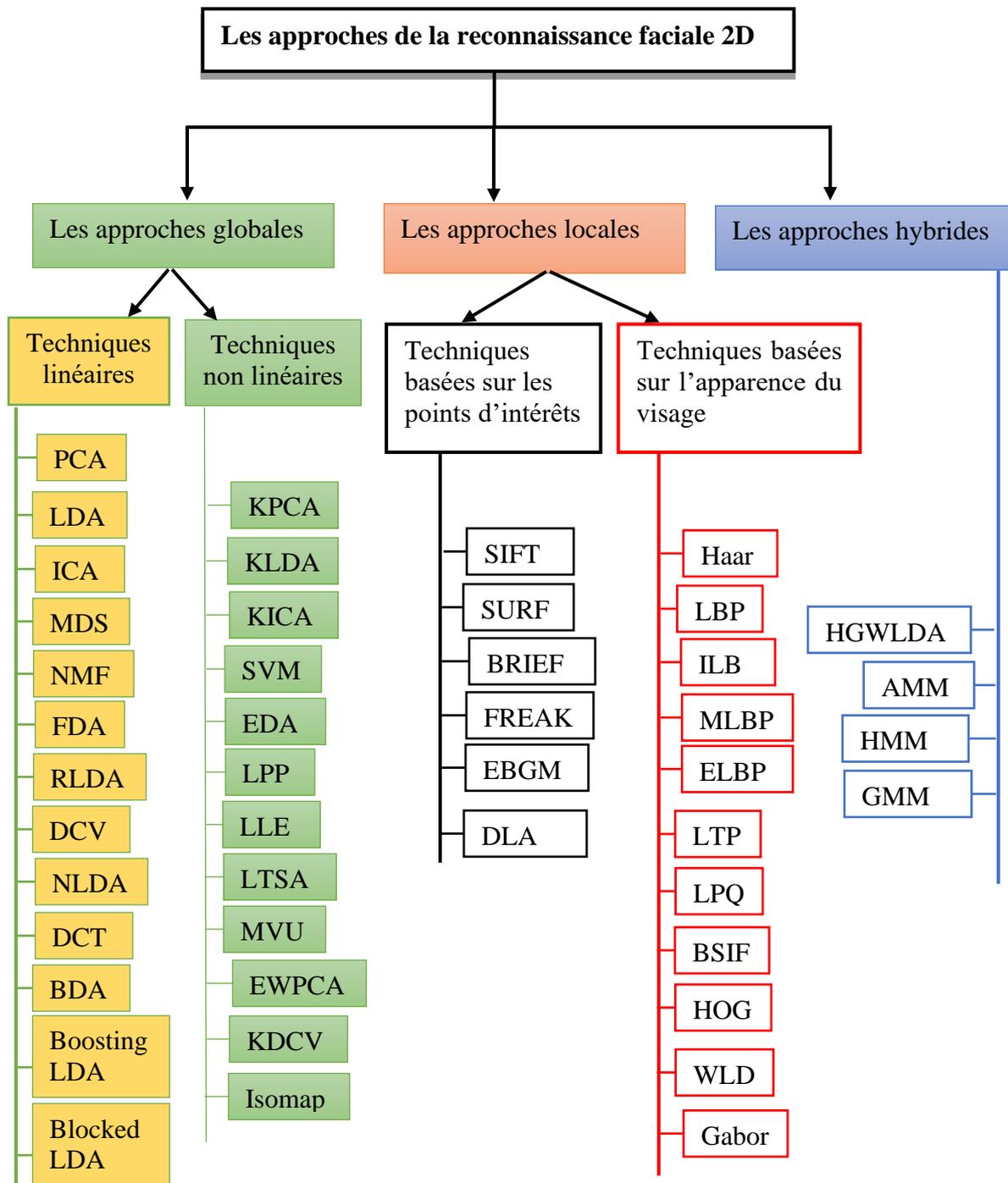


Figure 2.7 Classification des algorithmes principaux de reconnaissance faciale 2D.

2.3 La reconnaissance de visage 3D

Malgré que les techniques de reconnaissances de visage 2D sont robuste aux différents facteurs de variabilités (illumination, pose, occlusion), elles restent loin d'être développées. Avec le développement de caméra 3D, les travaux des recherches se penchent sur l'analyse des images faciale 3D afin de proposer des techniques de reconnaissance faciale 3D plus performantes et invariantes aux variations due à la pose et à l'illumination. Dans cette section, nous présentons les travaux récents sur la reconnaissance de visage 3D. Ces travaux peuvent être classés en deux catégories : les approches globales et approches locales.

2.3.1 Approches globales

Les informations géométriques 3D correspondant aux approches globales travaillent directement sur toute l'image de profondeur ou nuage de points du visage pour calculer la mesure de similarité. De nombreuses méthodes sont proposées.

2.3.1.1 Les méthodes de sous-espace

Une première approche dans cette catégorie a été proposée par *Hesher et al* dans [98]. Ils ont appliqué une méthode de réduction de dimensionnalité PCA sur des images de profondeur du visage. La base de test utilisée contient 222 images de profondeur associées à 37 personnes. Six expressions faciales différentes ont été collectées pour chaque sujet. Un espace propre de dimension réduite est ensuite créé à partir des images de profondeur normalisées et sur lequel sont projetées les images de test. Un taux de reconnaissance de 83% est obtenu.

Achermann et al [99] ont été aussi utilisés une méthode basée sur PCA (eigenfaces) et les modèles de *Markov* cachés. L'approche est évaluée sur une base de données de 24 personnes, chaque personne a 10 images de différentes poses. L'approche est testée sur 5 poses différentes pour chaque sujet. Cette approche a achevé un taux de reconnaissance de 100%.

Heseltine et all [100] ont évalué une nouvelle approche de la reconnaissance faciale en utilisant une variété de représentations de surface de la structure faciale 3D. En appliquant l'analyse en composantes principales (PCA) sur une grande base de données de modèles de visage 3D contenant 330 images de plus de 100 personnes. Le plus faible taux d'erreur égalité (TEE) 17.8% est obtenu par la distance de Mahalanobis, cette dernière est mieux que les distances Euclideennes et Cosinus. *Heseltine et al* [101] ont élargi le test d'un ensemble de données à 1470 images de profondeur de 230 personnes et ils ont exploité LDA dans le même

cadre, affirmant que LDA fonctionne mieux que PCA avec la distance Cosinus, où LDA obtient le plus faible TEE=15.3%.

2.3.1.2 Méthodes basées sur l'alignement de surfaces

Les méthodes appartenant à la première catégorie étendent les techniques de sous-espace, notamment celles de visages propres, sur des images de profondeur. D'autres méthodes basées sur l'appariement des surfaces faciales, initialement utilisées dans des problèmes d'alignement de modèles partiels ou de reconnaissance d'objets 3D, ont été adaptées pour la reconnaissance faciale 3D. L'algorithme généralement utilisé est l'algorithme du plus proche voisin itéré ICP (Iterative Closest Point), qui a été introduit par *Besl et McKay* [102]. ICP tente de manière itérative d'aligner deux surfaces 3D représentées comme points de nuages ou mailles.

Midioni et al [103] ont utilisé l'algorithme ICP pour aligner et calculer la similarité entre les surfaces faciale 3D. l'acquisition de données faciales est effectuée par un capteur stéréo passif afin de gérer 700 modèles 3D. le taux d'identification obtenu est autour de 98%, affirme l'efficacité de l'ICP. *Cook et al* [104] ont proposé une approche basée sur l'algorithme ICP pour établir la correspondance entre les surfaces 3D afin de compenser des erreurs dues à la nature non rigide des visages. En effet, les visages sont comparés en employant un modèle statistique, en l'occurrence un Mélange de Modèles Gaussiens (MMG). La distribution des erreurs est alors paramétrée. *Irfanoglu et al* [105] ont introduit une nouvelle méthode pour représenter les visages sous forme de nuages de points enregistrés en 3D. Un enregistrement fin des surface faciales est effectué en recherchant d'abord automatiquement les repères faciaux importants, puis établissant une correspondance dense entre les points de la surface à l'aide de l'algorithme ICP. Après l'enregistrement des surfaces faciales, la similitude entre deux visages est définie comme une approximation discrète de la différence de volume entre les surfaces faciales.

Bien que ICP soit un outil puissant pour évaluer la similitude entre deux visages, les méthodes basées sur l'algorithme ICP ne sont pas capables de traiter le problème de changement d'expression, car elles considèrent le visage 3D comme un objet rigide, ce qui n'est pas le cas dans la réalité. Plusieurs approches ont été développées pour surmonter ce problème.

Chang et al [106] ont modifié l'algorithme ICP à ICP partiel, qui sélectionne une partie des paires de points les plus proches pour calculer la mesure de dissimilarité pendant l'enregistrement des surfaces faciales, afin de réduire les effets négatifs causés par les variation

d'expression faciale. Le test est effectué sur un ensemble de donnée de 360 modèles de visage 3D, où les résultats ont montré que l'ICP partielle fonctionne mieux que l'ICP originale.

D'autre côté, *Chang et al* [107] ont proposé de segmenter le visage en régions avant d'appliquer l'algorithme de l'ICP sur ces régions. Ces régions sont sélectionnées autour du nez car elles sont les plus statiques de visage. L'approche est évaluée sur la base de données FRGC qui comporte 4000 scans 3D de 400 personnes avec différentes expressions faciales. L'approche proposée a achevé un taux d'identification de 92%.

Dans la plupart des travaux récents, l'algorithme ICP est largement introduit comme une étape clé du prétraitement pour corriger les variations de pose comme dans [108, 109], etc. Les **Figure 2.8**, **Figure 2.9** et **Figure 2.10** illustrent avec trois exemples différents résultats de l'application de l'algorithme ICP pour le recalage des surfaces faciales dans [110].

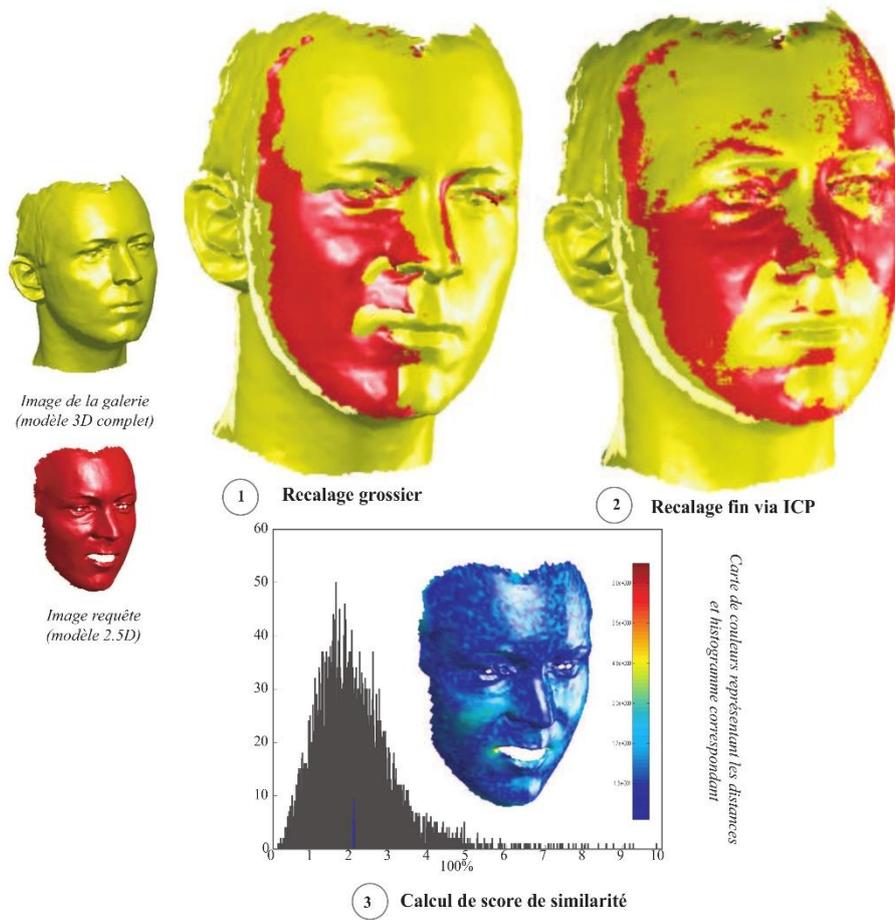


Figure 2.8 Exemple 1 recalage et calcul de similarité entre un modèle 3D et un modèle requête de la même personne avec des variations d'expressions faciales [110].

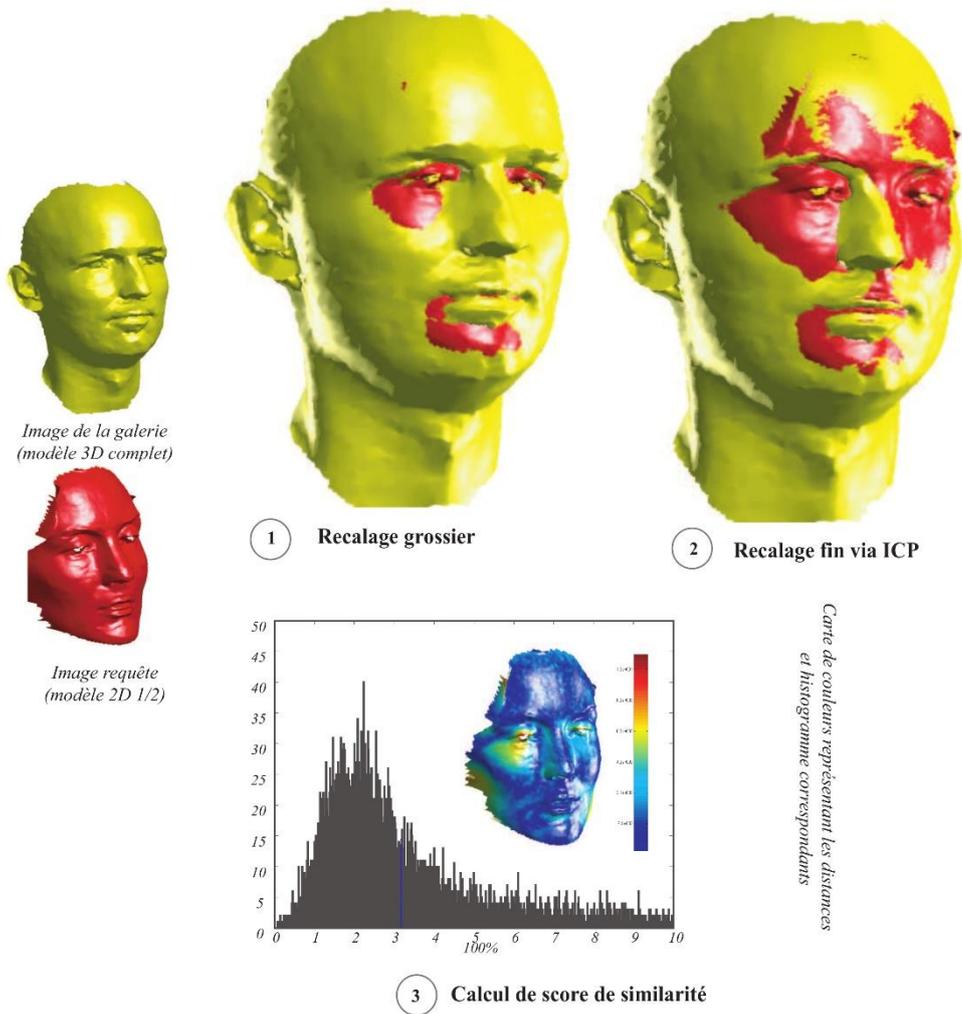


Figure 2.9 Exemple 2 recalage entre un modèle 3D et un modèle requête de la même personne avec des variations d'expressions faciales plus significatives [110].

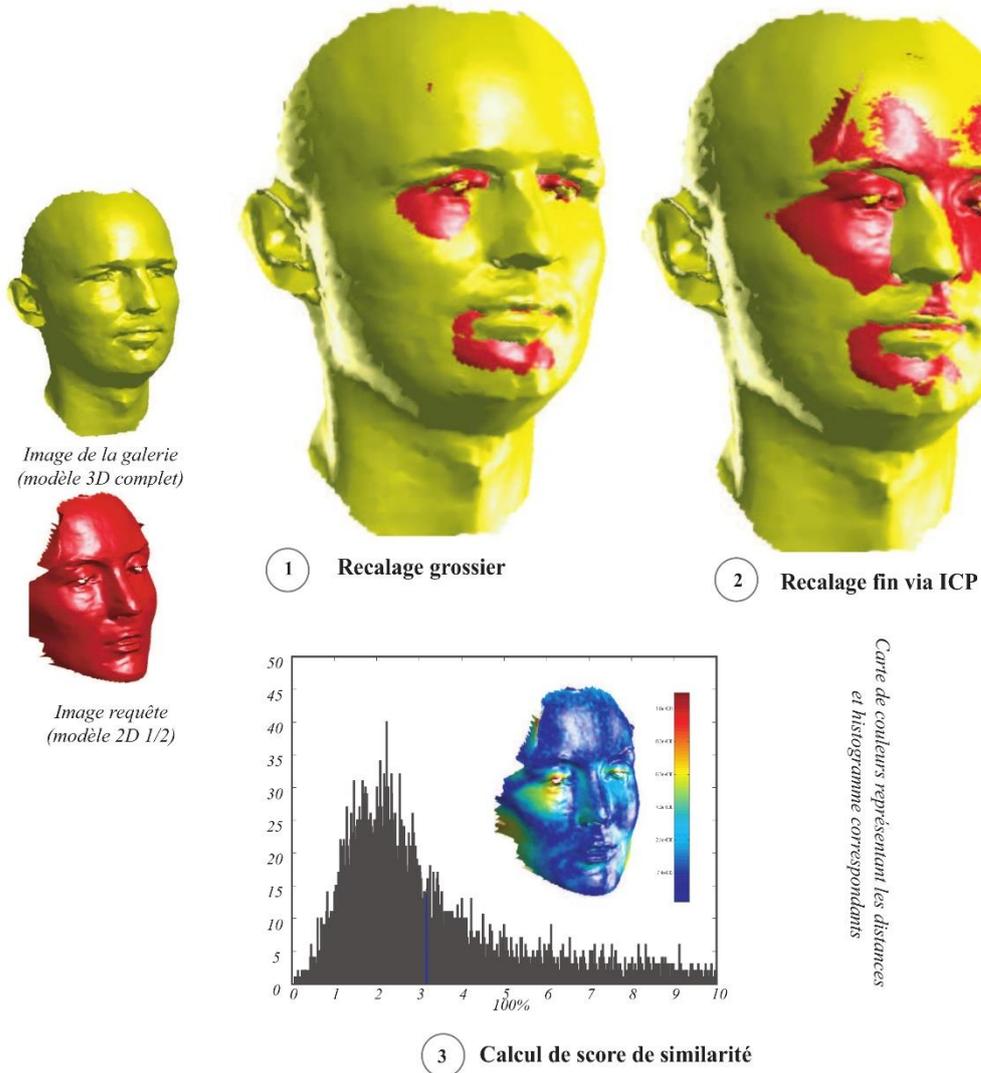


Figure 2.10 Exemple 3: recalage entre un modèle 3D et un modèle requête de deux personnes différentes [110].

2.3.1.3 Méthodes basées le modèle morphable (déformable)

Un modèle de visage morphable 3D, aussi appelé modèle déformable est un modèle génératif pour la forme et l'apparence du visage qui repose sur deux idées. Premièrement, tous les visages sont en correspondance point à point dense, qui est généralement établie sur un ensemble d'exemples de visages dans une procédure d'enregistrement puis maintenu tout au long des étapes de traitement ultérieures. En raison de cette correspondance, des combinaisons linéaires de visages peuvent être définies de manière significative, produisant des visages

morphologiquement réalistes (morphs). La deuxième idée est de séparer la forme et la couleur du visage, et les dissocier des facteurs externes tels que l'éclairage et les paramètres de la caméra. Le modèle morphable peut impliquer un modèle statistique de la distribution des visages, qui était une analyse en composante principale dans le travail originale de *Blanz et Vetter* [111] où ils ont inclus d'autres techniques d'apprentissage dans les travaux ultérieurs.

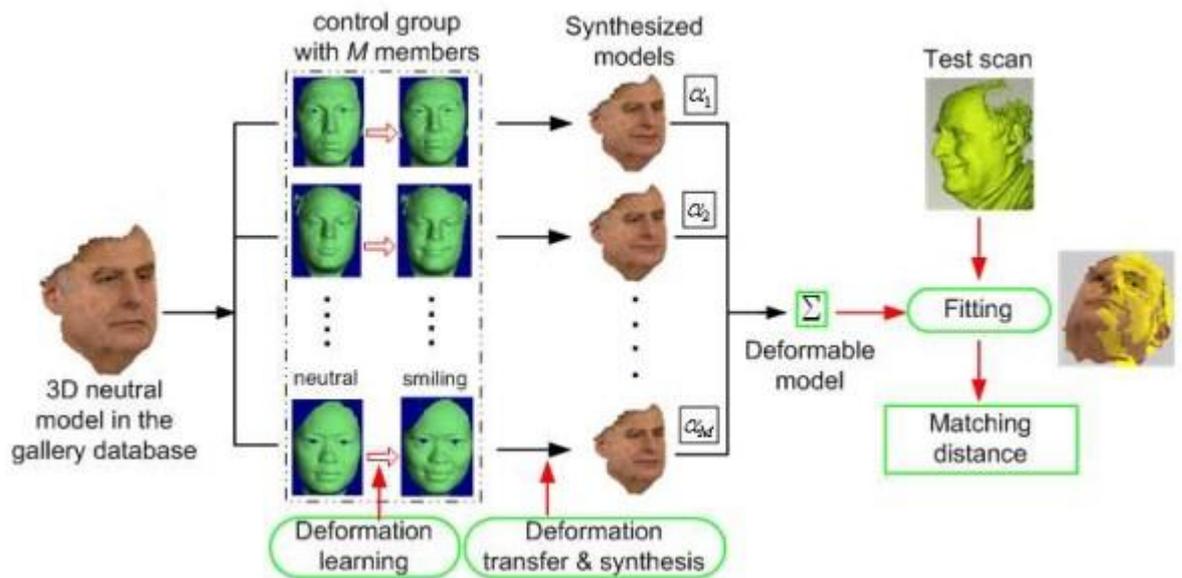


Figure 2.11 Modélisation des déformations pour l'appariement des visages 3D [112].

Lu et Jain [112] ont proposé un algorithme pour faire correspondre des images ayant des variations d'expression en utilisant des modèles de visage en 3D avec une expression neutre (voir la **Figure 2.11**). Une collection de 94 points de repère est utilisée pour étudier la déformation de surface 3D due à des expressions spécifiques. La cartographie est alors établie entre la collection de points de repère ayant une expression neutre et un modèle neutre 3-D. Une déformation comparable est incorporée sur un modèle neutre 3D en appliquant une cartographie basée sur la spline de plaque TPS (Thin Plate Spline). Chaque balayage d'image d'entrée (probe) associé à un modèle déformable est comparé aux visages de la galerie en fonction de l'algorithme de minimisation de la fonction de coût. Cette approche a atteint un taux de reconnaissance de rang 1 égale à 89%.

Amberg et Coll [113] ont introduit une méthode de reconnaissance faciale à expression invariante en adaptant des modèles 3D morphable d'identité et d'expression séparés à une forme de visage d'entrée 3D. Un modèle morphable 3D basé sur la méthode statistique PCA est formé avec une combinaison de 270 composants de visage neutres et 135 composants de visage expressifs. Une variante de l'ICP non rigide est employée comme algorithme

d'ajustement. Un modèle morphable 3D est ajusté aux correspondances de points en utilisant des approximations systématiques de *Jacobian et Gauss-Newton Hessian*. La distance angulaire de *Mahalanobis* est utilisée comme métrique de dissimilarité. Cette approche souffre d'un temps de traitement important (40s par image).

Al-Osaimi et Coll [114] ont utilisé l'ICP pour apprendre et modéliser les déformations d'expressions faciale. La technique de PCA est utilisée car elle est efficace en termes de calcul que d'autres solutions de forme fermée. Le modèle de déformation PCA générique est construit en utilisant des visages non neutres de personnes distinctes. Les modèles de déformation d'expression sont utilisés pour éliminer les expressions du balayage de visage non neutre. L'appariement basé sur l'ICP a augmenté la complexité du temps de calcul.

Haar et Veltkamp [115] ont construit un modèle PCA multi-résolution puissant en utilisant une collection limitée des points de repères faciaux ainsi que des scans neutres et d'expressions. Un seul modèle d'identité morphable et sept modèles d'expression morphable isolés par sujet sont construits. En raison de modèles séparés pour l'identité et les expressions, une expression est neutralisée et les coefficients du modèle sont utilisés pour la reconnaissance faciale. La distance L_1 est employée comme métrique de similarité pour l'appariement de visage. Cette méthode nécessite la localisation manuelle initiale des points de repères fiduciaires pour hisser les modèles morphables. Évidemment, les performances de cette technique sont soumises à une localisation précise des points de repères. La **Figure 2.12** illustre des exemples de visages d'ajustement du modèle morphable ont différentes expressions.

L'approche basée sur un modèle morphable proposée par *Haar et Veltkamp* se révèle être l'approche d'expression invariante la plus robuste car ils ont prouvé sa supériorité en termes de taux d'identification et de vérification sur les bases de données GvabDB, BU-3DFE, UND et FRGC v 2.0.



Figure 2.12 Exemples de visages d'ajustement du modèle morphable ont différentes expressions [115].

2.3.2 Approches locales

Les techniques de correspondance basées sur les caractéristiques locales se concentrent sur des points descriptifs locaux, des courbes ou des régions de scans faciaux 3D et les comparent dans un espace de caractéristiques (les coordonnées originales, les zones, les distances, les angles, les courbures, ...etc) pour prendre la décision. Ces approches ont été largement étudiées dans la reconnaissance faciale 3D, car elles donnent généralement de meilleurs résultats que l'appariement holistiques. Les approches basées sur les caractéristiques locales peuvent classer en cinq catégories : méthodes basées sur des points clés, méthodes basées sur des courbes, méthodes basées sur des régions, méthodes basées sur des descripteurs ainsi que sur une fusion d'entités de forme multiple.

2.3.2.1 Méthodes basées sur le calcul des points clés (Keypoint)

Les points clés sont des points d'intérêt de forme, basées sur la définition de la saillance. Ils sont détectés en fonction de certaines informations géométriques de la surface. Les méthodes impliquent généralement deux étapes principales, la détection des points clés et la description des fonctionnalités [116]. Bien que ces méthodes puissent faire face aux occlusions et aux pièces manquantes, leur coût de calcul est beaucoup plus élevé car elles utilisent un grand nombre de points clés.

a. Transformée de fonction variante d'échelle SIFT

La transformation de fonction variante d'échelle SIFT est un détecteur de points clés proposé par *Lowe* [117] pour extraire des caractéristiques invariantes distinctives à partir d'images qui peuvent être utilisées pour effectuer une correspondance fiable entre différentes vues d'un objet ou d'une scène.

Mian et al [118] ont proposé une nouvelle méthode pour détecter les points clés 3D inspiré du SIFT, où ils utilisent la variation de la forme en combinaison avec des descripteurs

SIFT 2D. pour détecter les points 3D dans une sphère de rayon r et centre p , le vecteur moyen m et la matrice de covariance C sont calculées. Ensuite, la matrice V des vecteurs propres est obtenue en effectuant une analyse en composante principale PCA sur la matrice de covariance C . le point p est défini comme un point clé, si la différence entre les deux premiers axes principaux de la région locale est supérieure à un seuil t . La **Figure 2.13** illustre des exemples de points clés identifiés sur trois ensembles d'images du quatre personnes différentes.

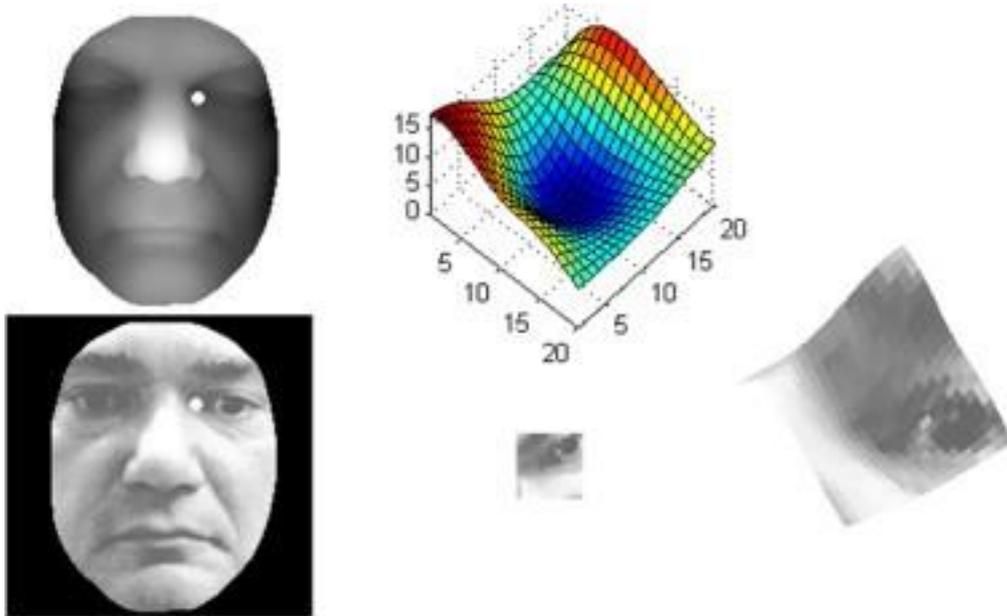


Figure 2.13 Un point clé sur une visage 3D et sa texture correspondante [118].

Inan et Halici [119] ont proposé une nouvelle méthode basée sur des descripteurs de forme locaux pour discriminer les scans faciaux 3D de différents individus. Dans [119] les descripteurs SIFT sont extraits de la matrice 2D des cartes de courbure, où les caractéristiques sont définies à des échelles et des orientations fixes pour des emplacements fixes. *Hang et al* [120] ont proposé une méthode de représentation de la surface faciale 3D basée sur LBP à multi-échelles, qui utilisé avec la carte d'indice de forme pour augmenter les caractères distinctives du visage. Le détecteur SIFT est introduit pour extraire les caractéristiques locales afin d'améliorer la robustesse aux variations de pose. De plus, une correspondance hybride est conçue pour une précision encore améliorée. La méthode proposée a achevé un taux de reconnaissance sur la base de données FRGC v 2.0 égale à 96.1%. La détection de points clés de SIFT est aussi appliquées sur des cartes d'indice de forme pyramidale dans [121].

b. Méthodes basées sur le maillage

Bien que les détecteurs de type SIFT présentent des fonctionnalités informatives sans enregistrement pour les scans frontaux, ils sont sensibles aux grandes variations de pose et occlusions. Pour surmonter ces limitations, le détecteur des points clés SIFT est appliqué directement sur les données du maillage 3D dans les travaux récents. Une extension pour le maillage 3D, appelée meshSIFT a été proposée par *Maes et al* [122], puis étendue par *Smeets et al* [123]. L'approche se compose de quatre étapes majeures : la détection des points clés, l'attribution de l'orientation, la description des caractéristiques locales et la correspondance des caractéristiques. Le détecteur de points clés de type meshSIFT a également été appliqué dans [124] en utilisant les courbures maximales et minimales, estimées dans l'espace d'échelle gaussienne 3D pour détecter les points saillants. La détection des points est illustrée sur la **Figure 2.14**. Pour calculer le descripteur local, un disque géodésique de rayon R est considéré autour de chaque point clé. Ensuite, un cercle de rayon r_1 et huit cercles de rayon r_2 sont extraits, comme le montre la **Figure 2.15**. Histogrammes y compris le gradient de surface (HOG), l'indice de forme (HoS) et le gradient d'indice de forme (HoGS) sont calculés pour chaque cercle. La concaténation de ces trois histogrammes est considérée comme le descripteur local.

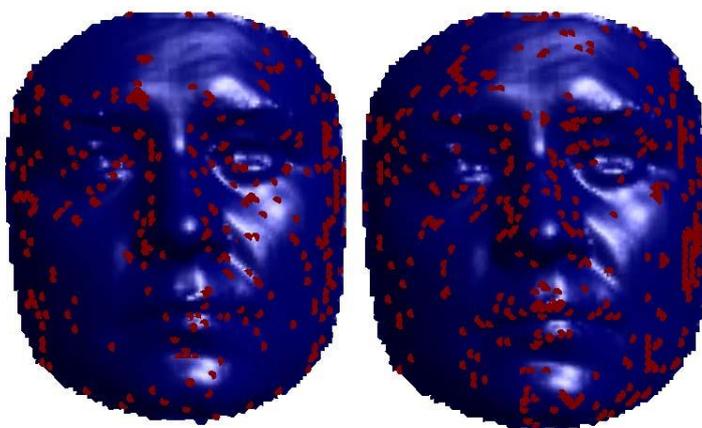


Figure 2.14 Points saillants détectés par K_{\max} (à gauche) et K_{\min} (à droite) [124].

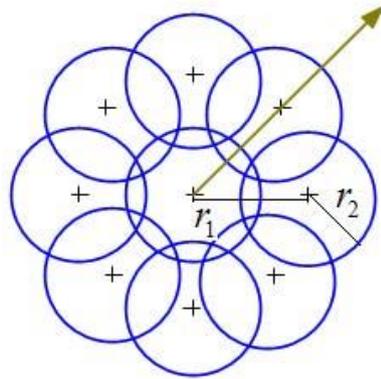


Figure 2.15 Orientations canoniques (flèche), point saillant et ses 8 sommets de voisinage (+) affectés de 9 cercles [124].

2.3.2.2 Méthodes basées sur le calcul de courbure

Ces méthodes utilisent un ensemble de courbes de surfaces de visages comme caractéristiques. Ces dernières comprennent des informations géométriques riches qui capturent les informations de forme de différentes régions faciales pour représenter un visage 3D. Par rapport aux méthodes basées sur les points clés, elles sont moins clairsemées et plus robustes contre les expressions faciales. De plus, le poids du point référence (souvent le bout du nez) est plus élevé que les autres points, car il contient des informations de forme descriptives. Les méthodes basées sur les courbes peuvent être regroupées en deux catégories :

a. Méthodes basées sur les contours

Les contours sont des courbes fermées de différentes longueurs et sans intersections. Ils sont définis comme des courbes de niveaux (level curves) classées en courbes iso-profondeur et iso-géodésiques. Les courbes iso-profondeurs sont introduites pour la première fois par *Samir et al* [125]. Elles sont obtenues en translatant un plan à travers la surface faciale dans une direction. Ces courbes sont décrites à l'aide des intersections entre la surface faciale et un plan. Pour une surface faciale S , on obtient un ensemble de courbes de niveaux c_λ , où chaque c_λ est constitué de tous les points p tels que $F(p) = t$, avec F étant une fonction de valeur de profondeur pour la composante z de point p . Une extension de ce cadre est proposée dans [126], où les courbes de niveaux d'une fonction de distance de surface, avec l'origine étant le bout du nez, sont décrites comme des courbes iso-géodésiques. Une courbe iso-géodésique c_λ est constituée de l'ensemble de tous les points, dont la distance géodésique $dist$, d'un point de référence est dans l'intervalle $[\lambda - \delta, \lambda + \delta]$, pour une petite valeur positive δ . Un cadre d'analyse Riemannien (*Bernhard Riemann*) est utilisé pour comparer les courbes faciales. Ces

derniers ont l'avantage d'être invariants aux des rotations ou aux des transformations isométriques. Cependant, à la fois les courbes iso-profondeur et les courbes iso-géodésiques (illustrées à la **Figure 2.16**), sont sensibles aux grandes expressions faciales, occlusions et les parties manquantes. Des rayures iso-géodésiques sont également été appliquées par *Berretti et al* [127]. Pour extraire les rayures, la distance géodésique normalisée $\bar{\gamma}$ est calculée entre chaque point du visage et le bout du nez, et quantifiée en N intervalles c_1, \dots, c_N . De cette façon la $n^{\text{ième}}$ bande se compose de tous les points dont les distances $\bar{\gamma}$ sont comprises dans l'intervalle c_n . Les rayures sont décrites par le descripteur 3DWW (3D Weighted Walkthroughs) et utilisées comme des nœuds dans un schéma de correspondance basé sur un graphique. Des courbes de niveaux ont également utilisées dans [128-132]. La principale limitation de la plupart de ces approches, à part occlusion, est leur manque de robustesse aux expressions faciales extrêmement grandes.

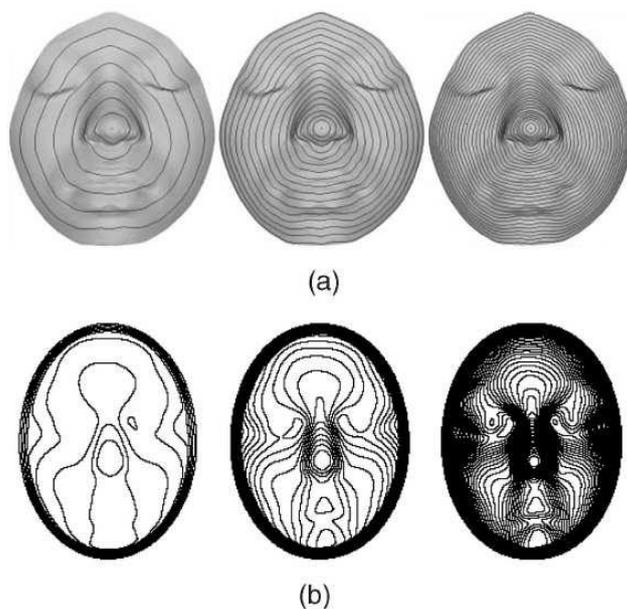


Figure 2.16 Courbes de niveaux a) de la fonction géodésique [126], b) de la fonction de profondeur [125] pour plusieurs niveaux.

b. Méthodes basées sur le profil

Les profils sont des courbes ouvertes, avec des points de départ et d'arrivée. Typiquement, les points de départ et d'arrivée sont respectivement au milieu et sur le bord du visage [133]. Des courbes radiales ont été introduites par *Drira et al* [134] et étendu dans [135]. Ces courbes sont plus efficaces que les courbes de niveaux [125, 126] car ils couvrent différentes régions du visage liées à différentes expressions faciales. Au moins certaines parties des courbes radiales sont disponibles pour gérer les occlusions et les parties manquantes. Dans l'étape d'extraction des courbes radiales, chaque courbe radiale β_α est obtenue par une coupe de la surface du visage par un plan p_α qui a le bout du nez comme origine et fait l'angle α avec le plan de référence. L'intersection du plan p_α et la surface S donne la courbe radiale β_α , comme le montre la **Figure 2.17**. Les courbes radiales sont également utilisées ensembles avec des niveaux définis dans [136], pour se rapprocher la surface de visage. La méthode bien connue l'algorithme d'apprentissage automatique AdaBoost (machine learning) est utilisé pour sélectionner les fonctions les plus efficaces. L'apprentissage automatique basée sur la méthode de sélection des caractéristiques fournit une signature très compacte d'un visage 3D et une approche de classification rapide pour la reconnaissance faciale. Lorsqu'en utilisant toutes les courbes, le temps de calcul pour la reconnaissance est 2.64s. Cependant, avec les courbes sélectionnées le temps est réduit à 0.68s, ce qui montre que la méthode de sélection améliore les performances de calcul du système.

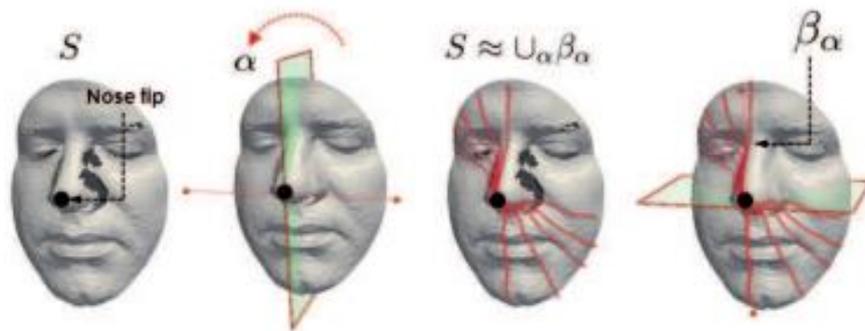


Figure 2.17 Procédure d'extraction des courbes radiales [135].

Les courbes faciales sont largement utilisées comme des méthodes basées sur l'expression faciale. Les signatures radiales angulaires (SRA) sont définies dans [137] comme un ensemble de courbes de profil pour traiter à un intervalle de θ radians ($\theta \in [0, \pi]$), émanant du bout de nez. Un masque binaire est défini sur le plan xy pour projeter les SRAs dans

différentes directions. Chaque chemin résultant se compose de 20 points, avec 3mm de distance entre deux points adjacents quelconques. La valeur de caractéristique SRA de ces points est calculée à partir de la valeur de profondeur de chaque point, en utilisant une interpolation bicubique aux coordonnées x et y . Le SRA extrait de manière significative un ensemble de vecteurs de caractéristiques d'identification discriminantes à partir de la surface faciale 3D complexe, qui atteint une efficacité de calcul dans la tâche de reconnaissance. Avec un processeur INTEL Core 2 Quad-CPU et 8 GB de RAM, l'identification des visages ne nécessite que 6.07s. En particulier, les traits extraits de régions semi-rigides sont robustes sous les expressions faciales. La **Figure 2.18** montre un masque binaire utilisé pour extraire des SRAs sur un visage. Une autre courbe faciale est obtenue en connectant des points clés SIFT, qui est introduite par *Berretti et al* [138] pour traiter les données manquantes, parce que les descripteurs SIFT ne sont pas suffisamment discriminants pour reconnaître une identité. Les courbes faciales de paires de points clés sont définies pour créer des caractéristiques efficaces. Un graphique des courbes faciales est construit entre les points clés correspondants. La performance de cette méthode en termes de précision peut être améliorée au niveau de l'appariement de courbes, lorsqu'une solution robuste est utilisée. En générale, les courbes sont moins discriminantes que les régions. Cependant, elles sont plus rapides et nécessitent moins d'espace pour le stockage.

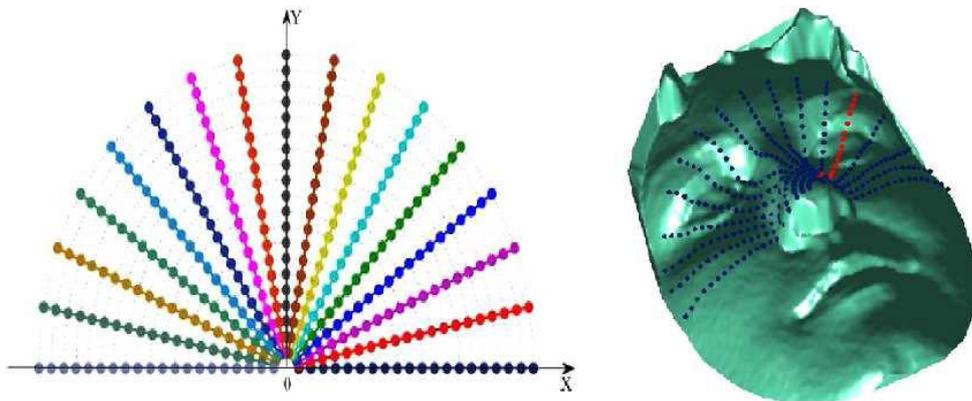


Figure 2.18 Le masque binaire et 17 SRAs [137].

2.3.2.3 Méthodes locale basée sur la surface

La plupart des méthodes basées sur la surface extraient les informations géométriques locales, à partir de plusieurs patches de la surface faciale ou de certaines régions de la surface, qui sont invariantes sous variations d'expressions faciales. Ces méthodes peuvent être divisées

en méthodes basées sur LBP, méthodes basées sur les caractéristiques géométriques et autres méthodes.

a. Méthodes basées sur LBP

LBP est généralement connu comme un descripteur efficace pour la reconnaissance de visage 2D. Récemment, LBP a été développé pour analyser les images 3D à l'expression faciale. Le descripteur local LBP a été initialement introduit par *Ojala et al* [139] pour les images 2D. LBP a été utilisé pour la première fois par *Li et al* [140] sur l'image d'intensité et la surface dans un schéma de fusion pour la reconnaissance faciale 3D. Plus tard, *Huang et al* [141] ont proposé une méthode 3DLBP basée à la fois sur des statistiques globales des caractéristiques géométriques et des statistiques locales des caractéristiques corrélatives des surfaces faciales. Dans [142], une LBP étendue multi-échelles avec une stratégie basée sur SIFT est proposée. La méthode proposée a été expérimentée sur trois ensembles de bases de données publics, à savoir FRGC v2.0, Gavab et Bosphorus. Elle affiche un taux de reconnaissance de premier rang de 97,6% et un taux de vérification de 98,4% à 0,001 FAR sur la base de données FRGC v2.0 sans aucun alignement de visage. La représentation LBP est également appliquée dans [143], où le modèle de division de visage est utilisé pour extraire la profondeur et les informations normales codées par LBP. L'opérateur LNP (Local Normal Patterns) est proposé dans [144] pour coder la composante normale de visage de la même manière que l'opérateur LBP. Le LNP est défini par le nombre décimaux du processus de codage. Les statistiques basées sur l'histogramme des valeurs LNP sont utilisées comme descripteur faciale (voir la **Figure 2.19**). Pour surmonter les variations d'expression faciale, le poids de chaque patch d'image normale du visage est appris et appliqué dans un classificateur basé sur une représentation clairsemée pondérée. La complexité de calcul de la méthode dépend de la taille de la galerie et de la dimension de l'élément. Le temps d'identification (taille de la galerie :466) sur un PC avec processeur Intel Core 2 et 2.66 GHz a été rapporté à 3.55 s pour cette méthode. De plus, l'algorithme ICP de base a été utilisé pour l'enregistrement du visage, montrant que l'application d'une méthode d'enregistrement efficace peut améliorer les résultats de la reconnaissance. *Werghi et coll* [145] ont proposé la méthode Mesh-LBP, où ils ont appliqué le descripteur LBP sur le maillage. Cette méthode a été étendue à la reconnaissance faciale dans [146].

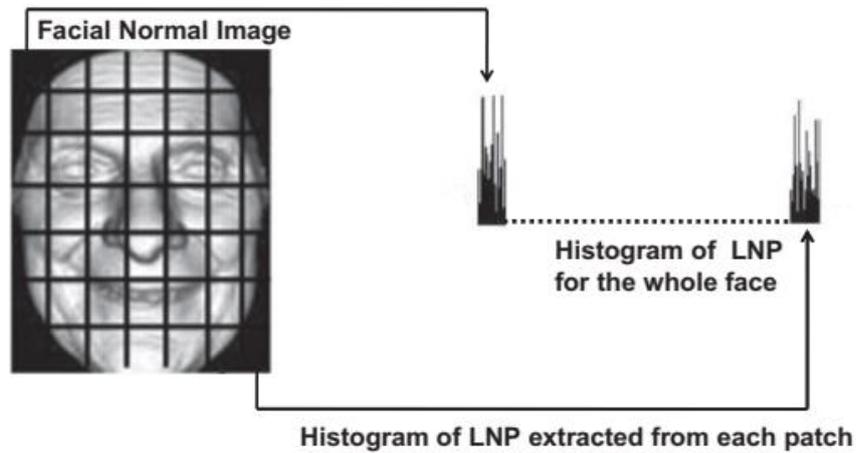


Figure 2.19 Illustration de la représentation normale du visage : histogramme de la LNP [144].

b. Méthodes basées sur les caractéristiques géométriques

Certaines méthodes sont développées sur la base de caractéristiques géométriques, *Xu et al* [147] ont proposé une méthode de reconnaissance faciale 3D utilisant de caractéristiques géométriques et informations de variation de forme. Tout d'abord, le nuage de points 3D est converti en un maillage, puis un vecteur de caractéristiques géométriques est construit, en utilisant la coordonnée de Z , $Z(v_i)$ de chaque sommet v du maillage. Les caractéristiques de forme sont extraites sur certaines régions du visage, y compris la bouche, le nez et les yeux. Deux vecteurs y compris les caractéristiques géométriques et de forme, sont concaténés pour créer le vecteur de caractéristiques. PCA est ensuite appliquée pour réduire la dimension de ce vecteur. *Li et Zhang* [148] ont utilisé plusieurs attributs géométriques intrinsèques, y compris les angles, les distances géodésiques et les courbures, pour la reconnaissance faciale 3D, où chaque visage est représentée par un maillage triangulaire, prétraité pour posséder une connectivité uniforme. Pour avoir un vecteur de caractéristique stable sous l'expression faciale, les signatures insensibles à l'expression sont construites à l'aide des attributs pondérés. *Li et al* [149] ont proposé une méthode de reconnaissance faciale 3D qui utilise les caractéristiques géométriques faciales 3D. D'abord, seule la pointe du nez doit être détectée. Puis, la distance géométrique entre les points autour du bout du nez est calculée pour représenter la géométrie du visage. La méthode proposée dépend sur deux étapes principales. Dans un premier temps, un certain nombre de candidats sont sélectionnés dans la galerie en fonction des caractéristiques géométriques. Dans la deuxième étape, l'algorithme ICP est exécuté pour faire correspondre l'image d'entrée (probe) avec les candidats afin de prendre la décision finale. Dans [150], un

vecteur insensible à l'expression EID (Expression Insensitive Descriptor) basé sur la représentation clairsemée des informations géométriques de bas niveau est proposé, où un schéma de regroupement et de classement est employé afin de sélectionner les EIDs de rang supérieur. La méthode proposée atteint des taux de reconnaissance satisfaisants.

c. Autres méthodes

Chua et al [151] ont proposé un algorithme de reconnaissance faciale basé sur la signature de points pour représenter les surfaces de forme libre. Pour gérer les variations d'expression faciale, seule les parties rigides du visage sont utilisées dans le processus de correspondance. La signature de points est également utilisée pour décrire les points caractéristiques dans les images 3D [152]. Plusieurs régions de chevauchement autour du nez sont extraites en utilisant des courbes de surface y compris la courbe moyenne H et la courbe Gaussienne K dans [153]. Une extension de ce travail est proposée par *Flatemier et al* [154]. *Mian et al* [155] ont introduit des tenseurs, où les tenseurs du troisième ordre sont indexés à l'aide d'une table de hachage 4D. Les champs tensoriels de rang 0 sont également appliqués par *Al-Osaimi et al* [156], où plusieurs champs de tenseurs locaux sont calculés sur un maillage triangulaire et utilisés comme des repères géométriques. La plupart de ces méthodes fonctionnent sur la base de l'enregistrement de surface et des descripteurs qui ne conviennent pas aux applications réelles et sont coûteuses en calcul. Récemment, *Ming* [157] a proposé un descripteur sphérique de délimitation régionale qui est efficace en termes de calcul et gère les émotions faciales avec un taux de reconnaissance élevé. Cette méthode prend 5.96 s pour l'ensemble du traitement des données, ce qui est considéré comme efficace en termes de temps. De plus, les caractéristiques 2D peuvent être calculées sur des cartes 2D extraites de maillages 3D pour décrire les caractéristiques locales telles que les coefficients de filtre de Gabor dans [108]. Une extension de ce travail utilisant des coefficients d'ondelettes a été présentée dans [158]. Les auteurs appliquent la notion des fonctionnalités pour définir des signatures compactes qui rendent la correspondance plus efficace, en particulier dans les bases de données à grande échelle. L'algorithme peut effectuer 1,800,000.00 comparaisons par seconde lorsqu'en utilise un processeur AMD 2.1GHz.

2.3.3 Discussion

Avec le développement rapide des systèmes de reconnaissance d'imagerie 3D, les scans faciaux 3D sont apparus comme une solution majeure pour résoudre les problèmes non résolus dans la reconnaissance faciale basée sur des images d'intensité 2D. Pendant ce temps, même si

les scans faciaux 3D capturent des informations de forme exactes des surfaces faciales et que les méthodes de reconnaissance faciale 3D sont théoriquement réputées pour être robustes aux variations d'éclairage, elles sont probablement plus sensibles aux variations d'expression faciale. En outre, elles nécessitent généralement une étape d'enregistrement précise avant la correspondance 3D basée sur la forme.

Généralement, comme la reconnaissance faciale basée sur une image d'intensité 2D, la façon de décrire les surfaces faciales est également un sujet central dans le domaine 3D. Les « bonnes » caractéristiques géométriques des formes faciales 3D devraient avoir les propriétés suivantes [159] : 1) ils peuvent tolérer des variations intra-classe tout en discriminant bien les différentes classes ; 2) ils doivent se trouver dans un espace avec une dimensionnalité modérée afin d'éviter des dépenses de calcul inabornables dans l'étape d'appariement. En conséquence, les techniques de reconnaissance faciale 3D existantes peuvent être grossièrement classifiées en fonction des fonctionnalités qu'elles utilisent :

(1) les techniques originales basées sur les caractéristiques utilisent toute la région du visage comme entrée pour calculer la similitude. Plusieurs travaux ont exploré le PCA directement sur les images faciales [98, 99, 160] ; tandis que certains ont appliqué l'algorithme ICP (Iterative Closest Point) [102] ou sa version modifiée sur les nuages de points faciaux pour correspondre aux surfaces [112, 161, 162]. La distance de Hausdorff a également été explorée pour faire correspondre les surfaces faciales [163, 164].

(2) les techniques basées sur des caractéristiques régionales ou des points clés détectent des points ou des zones faciales représentatives pour construire des espaces caractéristiques correspondants. Les zones des yeux et du nez sont utilisées dans [165], des régions et des lignes faciales segmentées sont utilisées dans [166], les points clés de référence anthropométriques faciaux sont étudiés dans [167].

(3) les méthodes basées sur les caractéristiques de courbe extraient des courbes discriminantes pour la représentation de la surface du visage. Dans [168], trois courbes faciales ont été trouvées pour couper la surface faciale en utilisant des plans horizontaux, verticaux et un cylindre ; le profil central avec saillie maximale ainsi que deux profils parallèles ont été recherchés dans [169] ; une union des courbes de niveaux d'une fonction de profondeur a été proposée pour décrire les surfaces faciales 3D [125].

(4) les algorithmes basés sur les caractéristiques de forme concentrent sur les attributs géométriques des surfaces locales, telles que les courbures [165], signature de point [151], image gaussienne étendue [170], carte de différence de forme signée [170], etc... .

Les techniques de la reconnaissance faciale 3D peuvent également être classées selon la nature de leurs stratégies d'appariement, c'est-à-dire d'appariement de visage holistique ou local, même si cela dépend fortement des traits du visage utilisés. Le flux holistique contient l'appariement basé sur l'algorithme ICP [162], description invariante d'isométrie [171], etc. ce schéma d'appariement est basé sur des caractéristiques faciales holistiques et nécessite donc généralement une étape de normalisation précise en ce qui concerne les changements de pose et d'échelle. De plus, les méthodes holistiques sont généralement sensibles aux expressions faciales et aux occlusions partielles. Tandis que, l'appariement basé sur les caractéristiques locales a l'avantage potentiel d'être robuste aux changements d'expression faciale, de pose, d'éclairage et même aux occlusions partielles. L'inconvénient de ce schéma d'appariement est la difficulté à extraire suffisamment les points caractéristiques informatifs à partir de surfaces faciales 3D similaires ou lisses.

2.4 La reconnaissance de visage 2D+3D

Il s'agit de techniques qui combinent des données 2D et 3D sur le visage pour améliorer les performances et la robustesse de la reconnaissance. Récemment, plusieurs approches basées sur ce principe ont été développées.

Beumier et Acheroy [172] ont proposé une méthode multimodale pour la reconnaissance faciale en utilisant une somme pondérée de mesure de similarité 2D et 3D. Ils ont utilisé un profil central et un profil latéral, chacun en 2D et 3D. Par conséquent, ils ont un total de quatre classificateurs, et une décision globale est prise en utilisant une somme pondérée des métriques de similarité. Un ensemble de données représentant plus de 100 personnes imagées sur plusieurs sessions, avec plusieurs poses par session, est acquis. Des parties de cet ensemble de données ont été utilisées par plusieurs autres chercheurs [173, 174]. Dans l'article [172], les résultats sont rapportés pour des expériences sur un sous-ensemble de données, en utilisant une galerie de 27 personnes et un ensemble d'images test (probe) de 29 personnes. Un taux d'erreur égal aussi bas que 1,4% est signalé pour la reconnaissance multimodale 3D + 2D qui fusionne plusieurs images de test (probe) par sujet. En général, la 3D + 2D multimodal s'avère plus performant que la 3D ou la 2D seule.

Xu et al [175] ont étudié les contributions apportées par les informations de profondeur et d'intensité pour la reconnaissance faciale et ils ont combiné les deux modalités démontrant des améliorations. Dans ce travail, les réponses aux ondelettes de Gabor sont extraites des images de profondeur et d'intensité, ainsi qu'une combinaison de LDA et Adaboost est utilisée pour sélectionner des caractéristiques robustes pour la classification.

Wang et al [176] ont proposé une approche multimodale 2D+3D basée sur la combinaison de deux méthodes, filtre de Gabor et signatures de points pour localiser les points d'intérêts 2D et 3D respectivement. Le PCA est utilisée pour projeter les points d'intérêts 2D et 3D chacun dans leur propre sous-espace. Les deux vecteurs résultants de forme 3D et de texture 2D sont regroupés dans un seul vecteur normalisé afin de représenter l'image faciale. La classification est effectuée par l'algorithme SVM (Support Vector Machines) avec un graphe acyclique dirigé par décision (DDAG). Les expériences sont réalisées en utilisant des images de 50 personnes, six images pour chaque personne, avec des variations de pose et d'expression faciale. Le taux de reconnaissance signalé pour la configuration 2D + 3D était supérieur à 90 % et il est démontré que la configuration multimodale a donné de meilleurs résultats que les configurations de modalités simples.

Bronstein et al [177] utilisent une approche de transformation isométrique pour analyser le visage en 3D afin de mieux faire face aux variations dues à l'expression du visage. La méthode qu'ils proposent est en fait la reconnaissance multimodale 2D +3D utilisant la décomposition propre de textures aplaties et d'images canoniques. Ils montrent des exemples de reconnaissance correcte et incorrecte par différents algorithmes, mais ne rapportent aucun résultat de performance quantitatif global pour n'importe quel algorithme.

Chang et al [178] rapportent des expériences de reconnaissance basées sur PCA réalisées à l'aide d'images 2D et 3D de 200 personnes. Une expérience utilise un seul ensemble d'images ultérieures pour chaque personne comme image d'entrée (probe). Une autre expérience utilise un plus grand ensemble de 676 images d'entrée prises dans plusieurs acquisitions sur un temps écoulé plus long. Les résultats des deux expériences sont environ 99% de reconnaissance 2D+3D de premier rang ((rank-one), 94% et 89% pour des données 3D seules et des données 2D seules respectivement.

Godil et al [179] ont présenté les résultats de la reconnaissance faciale 3D + 2D à l'aide de 200 personnes de données tirées de la base de données anthropométrique CAESAR. Ils ont utilisé PCA pour faire correspondre à la fois 2D et 3D, avec la 3D représentée comme une

image de gamme. Les données de visage 3D de cette base de données peuvent être assez grossière, avec environ 4000 points rapportés sur le visage. Plusieurs approches de fusion au niveau des scores les deux résultats sont explorées. Une performance aussi élevée que 82% de reconnaissance de premier rang est signalée.

Hüsken et al [180] décrivent l'approche appelée « Visage » pour la reconnaissance multimodale. La correspondance 3D suit le style de correspondance graphique hiérarchique déjà utilisé dans la technologie de reconnaissance faciale 2D de l'approche « Visage ». On estime que cela permet une plus grande vitesse d'appariement par rapport aux techniques basées sur ICP ou des techniques itératives similaires. La fusion des résultats des deux modalités se fait au niveau du score. Les performances multimodales sur la base de données FRGC v2.0 sont signalées comme une vérification de 93% à 0.01 FAR. En outre, il est rapporté que les performances de 2D seule ne sont que légèrement inférieures aux performances multimodales, et que les performances de 3D seule sont nettement inférieures à celles de 2D seule. Dans ce contexte, il peut être intéressant de noter que les résultats d'un groupe (Geometrix) qui se concentrait à l'origine sur la reconnaissance de visage 3D montre que la 3D seule surpasse la 2D seule, alors que les résultats d'un groupe « Visage » qui concentrait à l'origine sur la reconnaissance de visage 2D montre que la 3D seule surpasse la 2D seule.

Lu et al [181] s'appuient sur des travaux antérieurs avec la correspondance de style ICP de la forme 3D [182] pour créer un système multimodale 2D+3D. ils utilisent une approche d'analyse discriminante linéaire pour l'appariement de composant 2D. L'ensemble de données expérimentales consiste en plusieurs scans de chacune des 100 personnes. Cinq scans avec un système Minolta Vivid 910 sont effectués afin de créer un modèle de visage 3D pour l'inscription d'une personne. L'inscription se fait avec une expression neutre. Six scans sont effectués de chaque personne, trois avec une expression neutre et trois avec une expression souriante, à utiliser comme des images d'entrée individuelles pour les tests. Ils rapportent de meilleures performances avec la correspondance 3D seule qu'avec correspondance 2D seule. Ils rapportent également 98% de reconnaissance de premier rang pour la reconnaissance 2D + 3D sur les seules expressions neutres, et 91% sur le plus grand ensemble d'expressions neutres et souriantes.

Papatheodorou and Rueckert [183] effectuent une reconnaissance faciale multimodale 2D+3D en utilisant une généralisation de l'algorithme ICP basée sur des distances de points dans un espace 4D ($x, y, z, intensité$). Cette approche intègre les informations de forme et de

texture à un stade précoce, plutôt que de prendre une décision en utilisant chaque mode indépendamment et en combinant les décisions. Les résultats d'expériences ont été évalués sur 62 sujets de différentes poses et expressions faciales. Ils rapportent une reconnaissance correcte de 98 à 100%.

Al-Osaimi et al [184] proposent une fusion spatialement optimisée de données 3D et de texture pour modéliser des variations d'expressions faciales et d'illuminations discriminants des informations interpersonnelles. Dans [185], des données de visage 4D sont créées pour la reconnaissance de visage, qui est un maillage facial de sommets avec des attributs de texture et de géométrie. Ils développent un algorithme ICP 4D pour aligner les visages pour mesurer les similitudes pour la reconnaissance faciale. Une précision de 100% est obtenue en vue frontale sur un ensemble de données de 62 sujets avec cinq postures de la tête et trois expressions.

Zhang et Wang [186] ont proposé des caractéristiques invariantes de résolution locale extraites comme des extrema d'espace d'échelle à partir d'images de visage 2D et 3D. L'appariement des caractéristiques est obtenu à l'aide de six mesures de similarité différentes qui sont fusionnées au niveau du score.

Shih and Liu [187] extrait des caractéristiques de convolution basées sur quatre fonctions d'ondelettes à partir des images 2D et 3D. Un schéma de fusion linéaire est utilisé pour gérer la mesure de similitude finale.

Elaiwat et al [188] ont identifié des points clés répétables sur des surfaces de visage 3D texturées à l'aide d'un détecteur multimodale. Des descripteurs de surface locaux basés sur « Curvelet » sont ensuite construits autour de points clés détectés par cache, à la fois sur la texture et les caractéristiques locales 3D, en considérant différentes orientations.

En conclusion, la plupart des approches multimodales proposées dans la reconnaissance faciale confirment que la fusion de deux modalités 2D et 3D surpasse effectivement la performance de données 2D seule ou 3D seule.

2.5 Conclusion

Dans ce chapitre différentes approches de la reconnaissance faciale sont discutées. Généralement, les méthodes de la reconnaissance faciale proposées sont basées à la phase d'acquisition sur des données unimodaux 2D, 3D ou des données multimodaux 2D+3D.

La majorité des approches proposées sont basées sur l'image 2D d'intensité ou couleur. Ces approches peuvent atteindre un taux de reconnaissance plus de 90% dans les conditions bien contrôlées. En fait, certains des défis fondamentaux, à savoir les variations d'éclairage, de pose et d'expression faciale peuvent dégrader la performance des systèmes basés sur l'image faciale 2D.

Les approches basées sur les scans faciaux 3D sont généralement conçues pour résoudre les problèmes non résolus dans la reconnaissance faciale 2D. L'utilisation des scans faciaux 3D améliore effectivement la performance de reconnaissance grâce à sa robustesse aux variations d'éclairage et pose. Cependant, les méthodes de reconnaissance faciale 3D sont plus sensibles aux variations d'expression faciale. En outre, il y a d'autres critères qui doivent être pris en compte tels que la précision d'enregistrement de la surface, la précision de la détection des points clés, et la vitesse d'acquisition des données.

En conclusion, l'intérêt de la fusion des deux modalités (2D et 3D) en vue d'obtenir des taux de reconnaissances plus élevés est une solution proposée par plusieurs méthodes. Les méthodes basées sur les données 2D+3D ont prouvé ses efficacités par rapport aux méthodes qui ont exploité les données 2D seules ou 3D seules.

Chapitre 3

Apprentissage sous- espace multilinéaire

3.1 Introduction.....	54
3.2 Propriétés des tenseurs pour la modélisation de données multidimensionnelles	54
3.2.1 Notations et concepts.....	54
3.2.1.1 Définition d'un tenseur.....	54
3.2.2 Opérations de bases d'algèbre multilinéaire	57
3.2.2.1 Vectorisation : transformation tenseur en vecteur.....	57
3.2.2.2 Déploiement (Matricisation) : Transformation tenseur en matrice	57
3.2.2.3 Multiplication des tenseurs.....	58
3.2.2.4 Produit scalaire et la norme Frobenius	60
3.2.2.5 Tenseur de rang un (Rank-one tensor)	60
3.2.2.6 La distance euclidienne	61
3.2.2.7 Produit Kronecker	61
3.2.3 Projections multilinéaires	61
3.2.3.1 Projection : vecteur à vecteur (PVV)	62
3.2.3.2 Projection : tenseur à tenseur (PTT).....	63
3.2.3.3 Projection : tenseur à vecteur (PTV)	64
3.2.4 Relation entre les projections multilinéaires.....	66
3.3 Multilinear Whitening Principal Component Analysis (MWPCA)	69
3.4 Multilinear Discriminant Analysis (MDA).....	69
3.5 Within-Class Covariance Normalization (WCCN).....	70
3.6 L'algorithme MDA+WCCN.....	72
3.7 Conclusion :	73

Chapitre 3 : Apprentissage sous-espace multilinéaire

3.1 Introduction

Dans les dernières années, il y a un intérêt croissant pour les techniques d'apprentissage de sous-espace pour la reconnaissance faciale. Cependant, les images faciales capturées par les nouveaux dispositifs généralement ont des structures multi-dimensionnelles. Ceci conduit à une forte demande d'utilisation des techniques multilinéaires (tenseurs). L'intérêt de ce chapitre est de présenter les propriétés des tenseurs pour la modélisation de données multidimensionnelles, ainsi que les différentes techniques multilinéaires utilisées dans la littérature.

Dans un premier temps, nous présentons les notations et les concepts nécessaires à la définition du tenseur. Ensuite, nous étudions les bases d'algèbre multilinéaire et les différentes projections multilinéaires, notamment la projection vecteur à vecteur, la projection tenseur à vecteur, et la projection tenseur à tenseur, et leurs relations sous-jacentes. Enfin, nous présentons les techniques multilinéaires telles que MWPCA, MDA, MDA+WCCN qui sont élaborées à expérimenter les résultats de l'approche proposée.

3.2 Propriétés des tenseurs

3.2.1 Notations et concepts

Dans cette thèse, nous suivons les conventions de notations en algèbre multilinéaire, la reconnaissance des formes et l'apprentissage adaptatif [189-191]. Alors, les tenseurs sont notés par des scripts ($\mathcal{A}, \mathcal{B}, \mathcal{X}, \mathcal{Y}$), les matrices par des lettres majuscules en gras (\mathbf{U}, \mathbf{S}), les vecteurs par des lettres minuscules en gras (\mathbf{x}, \mathbf{y}) et les scalaires sont désignés par des lettres minuscules et majuscules (i, j, L, N).

3.2.1.1 Définition d'un tenseur

Le terme tenseur a différentes significations en mathématiques et en physique. L'usage de ce mot dans cette thèse fait référence à sa signification mathématique, en particulier en algèbre multilinéaire. Le tenseur est considéré comme un tableau multidimensionnel [5]. La **Figure 3.1** illustre différentes représentations tensorielles des données multidimensionnelles.

a) Ordre du tenseur :

L'ordre d'un tenseur est le nombre de ses dimensions. Les scalaires peuvent donc être interprétés comme des tenseurs d'ordre zéro ($\mathbf{x} \in \mathbb{R}$), les vecteurs comme des tenseurs du premier ordre ($\mathbf{x} \in \mathbb{R}^1$) et les matrices comme des tenseurs du second ordre ($\mathbf{X} \in$

$\mathbb{R}^{I_1 \times I_2}$). Les tenseurs d'ordre trois ou plus ($\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$) sont appelés des tenseurs d'ordre supérieur. La **Figure 3.2** montre comment on peut passer des scalaires aux tenseurs.

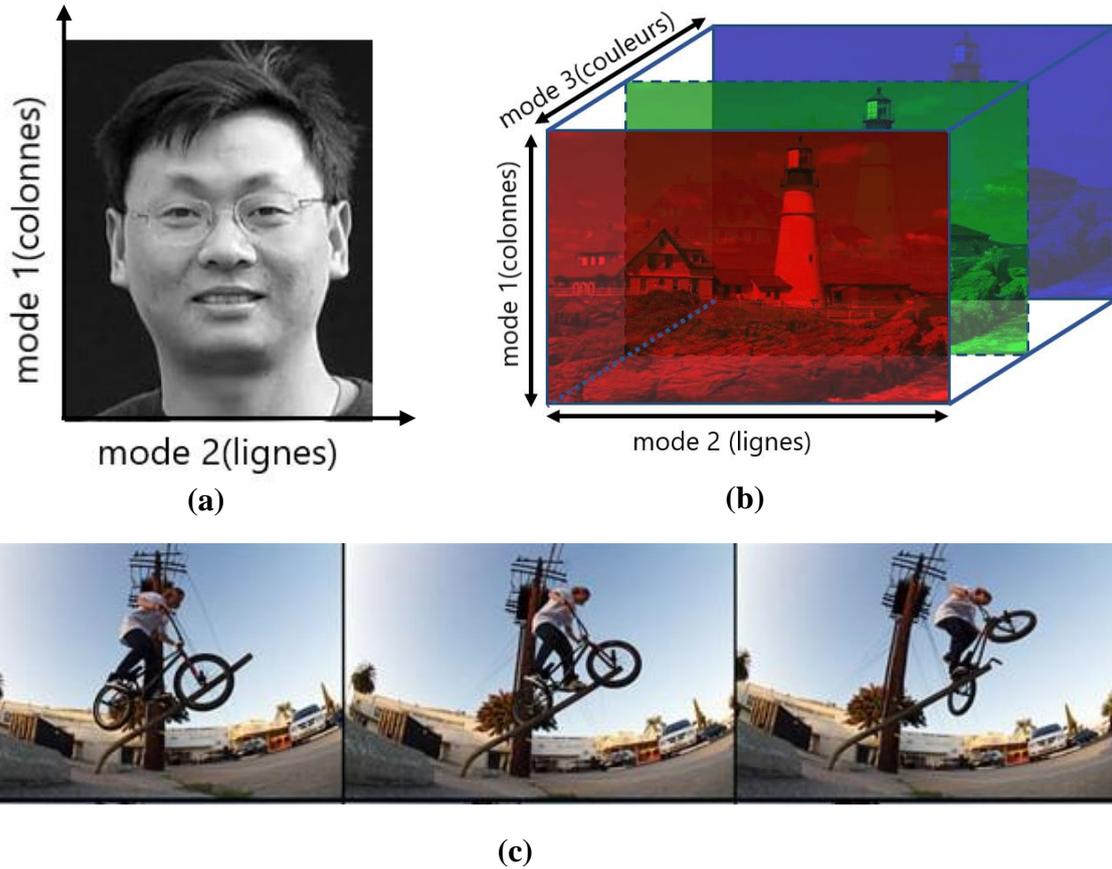


Figure 3.1 Exemples de la représentation tensorielle : (a) image en niveau de gris (mode 2) ; (b) image couleur (mode 3) ; (c) séquence vidéo en couleur (mode 4).

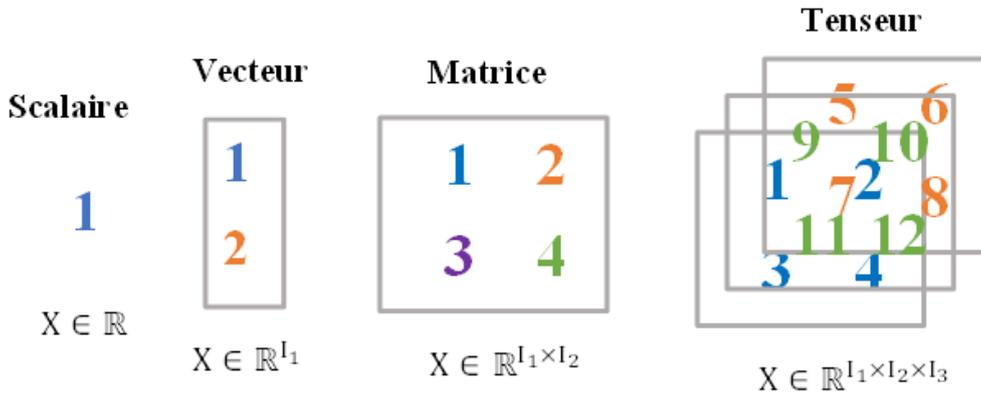


Figure 3.2 Passage des scalaires aux tenseurs.

b) Indexation des tenseurs

Nous pouvons créer des sous-tableaux (ou sous-champs) en fixant certains indices de tenseurs. Des fibres sont créées lors de la fixation de tous les indices sauf un, des tranches (ou dalles) sont créées lors de la fixation de tous les indices sauf deux. Pour un tenseur du troisième ordre, les fibres sont données sous forme $x_{:,jk}$ (colonne), $x_{j,:k}$ (ligne) et $x_{jk,:}$ (tube) ; les tranches sont données sous forme $X_{::k} = X_k$ (frontale), $X_{:,j}$ (latérale), et $X_{i,:}$ (horizontale). La **Figure 3.3** représente des exemples graphiques de fibres et de tranches pour un tenseur du troisième ordre.

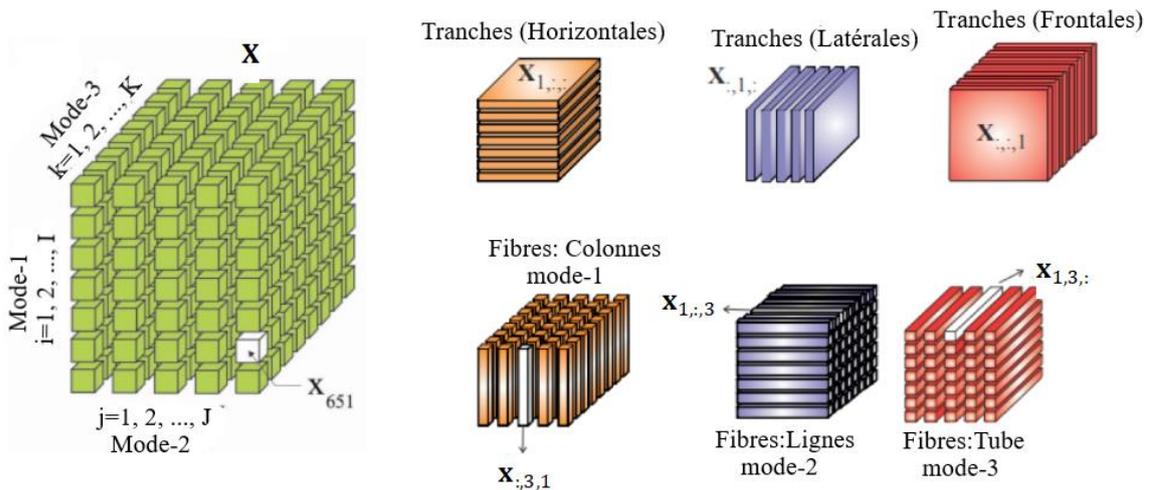


Figure 3.3 Exemples d'un tenseur du troisième ordre sous forme des fibres et des tranches.

3.2.2 Opérations de bases d'algèbre multilinéaire

3.2.2.1 Vectorisation : transformation tenseur en vecteur

La vectorisation de tenseur est une transformation linéaire qui convertit le tenseur $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ en un vecteur colonne $\mathbf{a} \in \mathbb{R}^{\prod_{n=1}^N I_n}$, dénoté par $\mathbf{a} = \text{vec}(\mathcal{A})$.

3.2.2.2 Déploiement (Matricisation) : Transformation tenseur en matrice

La transformation tenseur en matrice, également connue sous le nom de déploiement ou d'aplatissement, est le processus de réorganisation des éléments d'un tableau à N voies dans une matrice. Par exemple, un tenseur $2 \times 3 \times 4$ peut être arrangé comme une matrice 6×4 ou une matrice 3×8 , et ainsi de suite.

Le déploiement du tenseur $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ dans le mode n est dénoté par $\mathbf{A}_{(n)} \in \mathbb{R}^{I_n \times (I_1 \times \dots \times I_{n-1} \times I_{n+1} \times \dots \times I_N)}$, où les vecteurs colonnes de $\mathbf{A}_{(n)}$ sont les vecteurs mode n du tenseur \mathcal{A} .

L'élément tensoriel (i_1, i_2, \dots, i_N) correspond à l'élément de matrice (i_n, j) , où

$$j = 1 + \sum_{\substack{k=1 \\ k \neq n}}^N (i_k - 1) J_k, \quad \text{with} \quad J_k = \prod_{\substack{m=1 \\ m \neq n}}^{k-1} I_m \quad 3.1)$$

Un exemple graphique sur le déploiement du tenseur est illustré par la **Figure 3.4** où le tenseur d'ordre 3 est transformé en trois matrices de déploiement via la structure vectorielle (fibres).

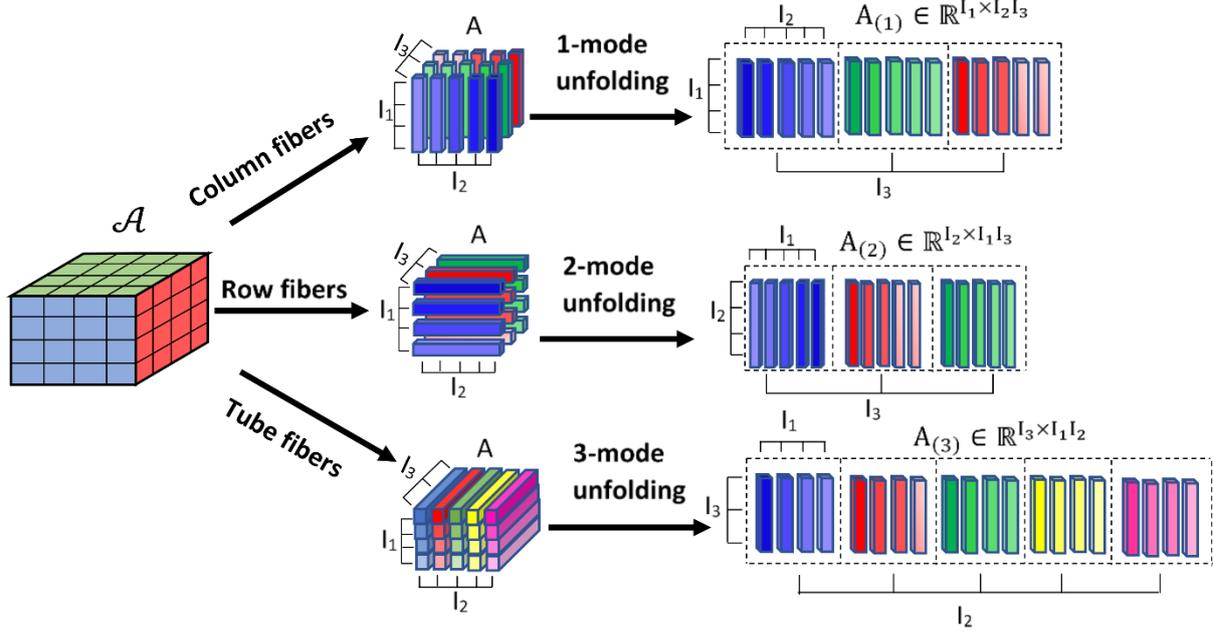


Figure 3.4 Exemple de déploiement d'un tenseur d'ordre 3 dans mode 1, 2 et 3 [192].

3.2.2.3 Multiplication des tenseurs

a) Produit mode n : Multiplication tenseur matrice

Le produit en mode n du tenseur $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ avec la matrice $\mathbf{U} \in \mathbb{R}^{J_n \times I_n}$ est un tenseur $\mathcal{B} \in \mathbb{R}^{I_1 \times \dots \times I_{n-1} \times J_n \times I_{n+1} \times \dots \times I_N}$, dénoté par

$$\mathcal{B} = \mathcal{A} \times_n \mathbf{U} \quad (3.2)$$

où chaque entrée de tenseur \mathcal{B} est définie comme la somme des produits des entrées correspondantes en \mathcal{A} et \mathbf{U} :

$$(\mathcal{A} \times_n \mathbf{U})(i_1, \dots, i_{n-1}, j_n, i_{n+1}, \dots, i_N) = \sum_{i_n} \mathcal{A}(i_1, \dots, i_n) \cdot \mathbf{U}(j_n, i_n) \quad (3.3)$$

Cela est équivalent à prémultiplier chaque mode n vecteur de tenseur \mathcal{A} par la matrice \mathbf{U} . Ainsi, le produit mode n ci-dessus peut être écrit en utilisant les matrices dépliantes suivantes :

$$\mathbf{B}_{(n)} = \mathbf{U} \mathbf{A}_{(n)} \quad (3.4)$$

Pour le cas d'un tenseur (matrice) de deuxième ordre \mathbf{A} et \mathbf{U} de tailles appropriées, en écrivant :

$$\mathbf{A} \times_1 \mathbf{U} = \mathbf{U} \mathbf{A}, \quad \mathbf{A} \times_2 \mathbf{U} = \mathbf{A} \mathbf{U}^T \quad (3.5)$$

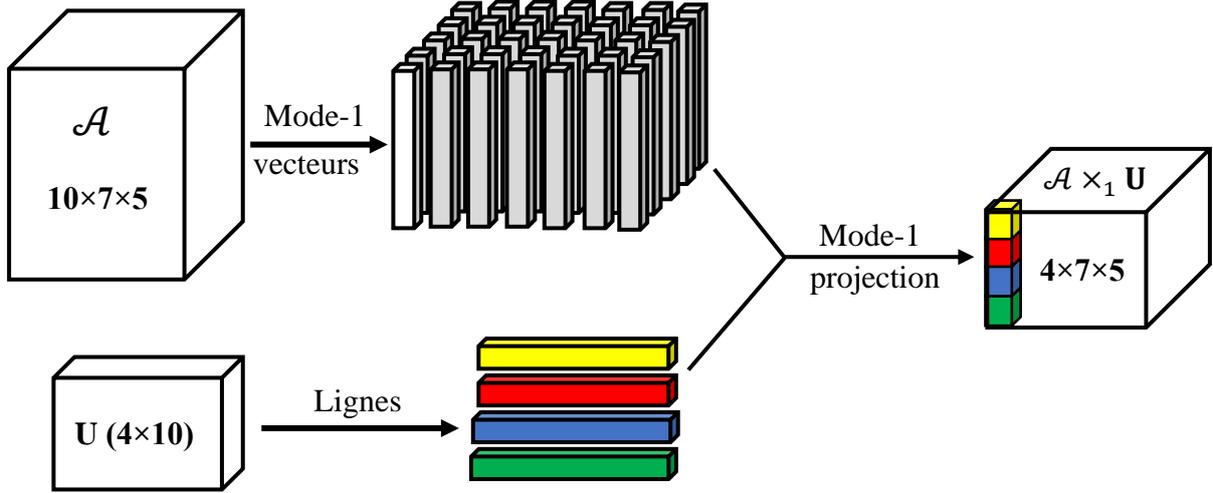


Figure 3.5 Illustration visuelle de la multiplication tenseur-matrice dans le mode 1.

Le produit $\mathcal{A} \times_1 \mathbf{U}$ est calculé comme les produits internes entre le mode-1 vecteurs de tenseur \mathcal{A} et les lignes de la matrice \mathbf{U} . Dans la **Figure 3.5**, chaque mode-1 vecteur de tenseur $\mathcal{A} \in \mathbb{R}^{10 \times 7 \times 5}$ est projeté par la matrice $\mathbf{U} \in \mathbb{R}^{4 \times 10}$ pour obtenir un vecteur ($\mathbf{v} \in \mathbb{R}^4$).

La multiplication tenseur matrice a deux propriétés [189] :

Propriété 3.1 : En donnant un tenseur $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$, et deux matrices $\mathbf{U} \in \mathbb{R}^{J_n \times I_n}$ et $\mathbf{V} \in \mathbb{R}^{I_m \times I_m}$, où $m \neq n$, nous avons :

$$(\mathcal{A} \times_m \mathbf{U}) \times_n \mathbf{V} = (\mathcal{A} \times_n \mathbf{V}) \times_m \mathbf{U} \quad (3.6)$$

Propriété 3.2 : En donnant un tenseur $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$, et deux matrices $\mathbf{U} \in \mathbb{R}^{J_n \times I_n}$ et $\mathbf{V} \in \mathbb{R}^{K_n \times I_n}$, où $m \neq n$, nous avons :

$$(\mathcal{A} \times_n \mathbf{U}) \times_n \mathbf{V} = \mathcal{A} \times_n (\mathbf{V} \cdot \mathbf{U}) \quad (3.7)$$

b) Produit mode n : Multiplication tenseur vecteur

Le produit en mode n du tenseur $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ avec le vecteur $\mathbf{u} \in \mathbb{R}^{I_n \times 1}$ est un tenseur $\mathcal{Z} \in \mathbb{R}^{I_1 \times \dots \times I_{n-1} \times 1 \times I_{n+1} \times \dots \times I_N}$, dénoté par :

$$\mathcal{Z} = \mathcal{A} \times_n \mathbf{u}^T \quad (3.8)$$

Où chaque entrée de tenseur \mathcal{Z} est définie par :

$$\mathcal{Z}(I_1 \times \dots \times I_{n-1} \times 1 \times I_{n+1}, \dots, I_N) = \sum_{i_n} \mathcal{A}(i_1, \dots, i_N) \cdot \mathbf{u}(i_n) \quad (3.9)$$

La multiplication d'un tenseur par un vecteur peut être considéré comme un cas particulier de la multiplication d'un tenseur par une matrice où $J_n = 1$ (donc, $\mathbf{U} \in \mathbb{R}^{J_n \times I_n} = \mathbf{u}^T$). Ce produit $\mathcal{A} \times_n \mathbf{u}^T$ peut-être calculer comme des produits internes entre le mode-1 vecteurs de tenseur \mathcal{A} et le vecteur \mathbf{u} .

3.2.2.4 Produit scalaire et la norme Frobenius

- **Définition 3.1** : le produit scalaire (ou produit interne) entre deux tenseurs $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ ont la même dimension est donné par :

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i_1} \sum_{i_2} \dots \sum_{i_N} \mathcal{A}(i_1, i_2, \dots, i_N) \cdot \mathcal{B}(i_1, i_2, \dots, i_N) \quad (3.10)$$

- **Définition 3.2** : La norme Frobenius du tenseur \mathcal{A} est définie par :

$$\|\mathcal{A}\|_F = \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle} \quad (3.11)$$

3.2.2.5 Tenseur de rang un (Rank-one tensor)

Le rang d'un tenseur \mathcal{X} , noté $\text{rang}(\mathcal{X})$, est défini comme le plus petit nombre de tenseur de rang un.

- **Définition 3.3** : un tenseur à N voies $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ est de rang un s'il peut être écrit comme le produit extérieur de N vecteurs c'est-à-dire,

$$\mathcal{X} = \mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \dots \circ \mathbf{a}^{(N)} \quad (3.12)$$

Ce qui signifie que chaque élément du tenseur est le produit des éléments vectoriels correspondants :

$$\mathbf{x}_{i_1, i_2, \dots, i_N} = \mathbf{a}_{i_1}^{(1)} \cdot \mathbf{a}_{i_2}^{(2)} \cdot \dots \cdot \mathbf{a}_{i_N}^{(N)}, \quad 1 \leq i_n \leq I_n \quad (3.13)$$

La **Figure 3.6** illustre un exemple d'un tenseur de rang du troisième ordre.

Tenseur de rang 1

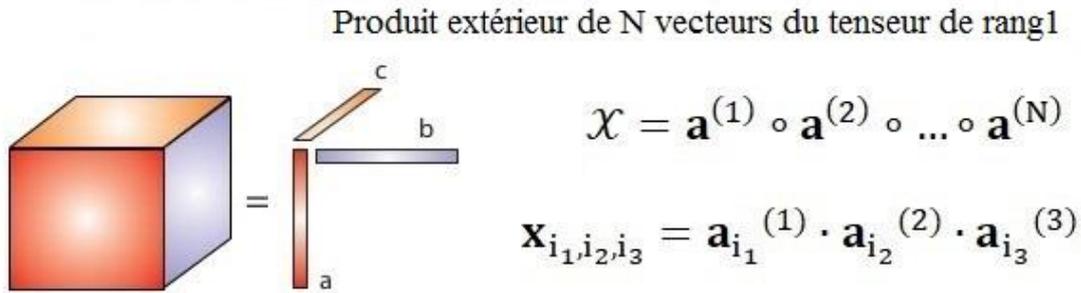


Figure 3.6 Exemple d'un tenseur de rang 1 du troisième ordre 3.

3.2.2.6 La distance euclidienne

La distance entre les tenseurs \mathcal{A} et \mathcal{B} peut-être mesurer par la Frobenius norme[10] :

$$dist(\mathcal{A}, \mathcal{B}) = \|\mathcal{A} - \mathcal{B}\|_F \quad (3.14)$$

Bien qu'il s'agisse d'une mesure basée sur un tenseur, elle équivalente à une mesure de distance des représentations vectorielles correspondantes comme prouvée dans [193]. Soit $vec(\mathcal{A})$ la représentation vectorielle de \mathcal{A} , alors

$$dist(\mathcal{A}, \mathcal{B}) = \|vec(\mathcal{A}) - vec(\mathcal{B})\|_F \quad (3.15)$$

Cela implique que la distance entre deux tenseurs définis dans (3.14) est égale à la distance euclidienne entre leurs représentations vectorisées.

3.2.2.7 Produit Kronecker

Le produit Kronecker de matrices $\mathbf{A} \in \mathbb{R}^{I \times J}$ et $\mathbf{B} \in \mathbb{R}^{K \times L}$ est dénoté par $\mathbf{A} \otimes \mathbf{B}$. Le résultat est une matrice de taille $(IK) \times (JL)$ donnée par :

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11} \mathbf{B} & a_{12} \mathbf{B} \cdots & a_{1J} \mathbf{B} \\ a_{21} \mathbf{B} & a_{22} \mathbf{B} \cdots & a_{2J} \mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{I1} \mathbf{B} & a_{I2} \mathbf{B} \cdots & a_{IJ} \mathbf{B} \end{bmatrix} \quad (3.16)$$

$$= [a_1 \otimes b_1 a_1 \otimes b_2 a_1 \otimes b_3 \dots a_j \otimes b_{L-1} a_j \otimes b_L]$$

3.2.3 Projections multilinéaires

L'apprentissage sous-espace multilinéaire est une méthode de la réduction de la dimensionnalité. Cette réduction peut être effectuée sur un tenseur de données dont les observations des vecteurs ou matrices concaténées en un tenseur de données. Par conséquent, la projection multilinéaire c'est le mappage d'un espace vectoriel de haute dimension à un ensemble d'espaces vectoriels de dimension inférieur en effectuant N projections linéaire multiples. Alors, il y a trois projections multilinéaires de base nommées en fonction de l'entrée et la sortie de la projection [194]: la projection traditionnelle vecteur à vecteur, la projection tenseur à tenseur et la projection tenseur à vecteur. De plus, nous étudions les relations entre ces projections.

3.2.3.1 Projection : vecteur à vecteur (PVV)

C'est une transformation linéaire qui projette un vecteur sur un autre vecteur. Soit $\mathbf{a} \in \mathbb{R}^I$ et $\mathbf{b} \in \mathbb{R}^P$ deux vecteurs, pour projeter le vecteur \mathbf{a} sur le vecteur \mathbf{b} en utilisant la matrice de projection $\mathbf{U} \in \mathbb{R}^{I \times P}$:

$$\mathbf{b} = \mathbf{U}^T \mathbf{a} = \mathbf{a} \times_1 \mathbf{U}^T \quad (3.17)$$

Si nous nommons cette projection en fonction de ses représentations d'entrée et de sortie, la projection linéaire est une projection vecteur à vecteur car elle mappe un vecteur à un autre vecteur. Lorsque les données d'entrée sont une matrice ou un tenseur d'ordre supérieur, il doit être vectorisé (remodelé en vecteur) avant la projection. La **Figure 3.7** illustre la projection de vecteur à vecteur d'un objet tensoriel \mathcal{A} .

Dénotons chaque colonne de \mathbf{U} par \mathbf{u}_p , donc $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2, \dots \ \mathbf{u}_p]$. Puis, chaque colonne \mathbf{u}_p projette \mathbf{a} sur un scalaire $\mathbf{b}(p)$ ou \mathbf{b}_p :

$$\mathbf{b}_p = \mathbf{b}(p) = \mathbf{u}_p^T \mathbf{a} \quad (3.18)$$

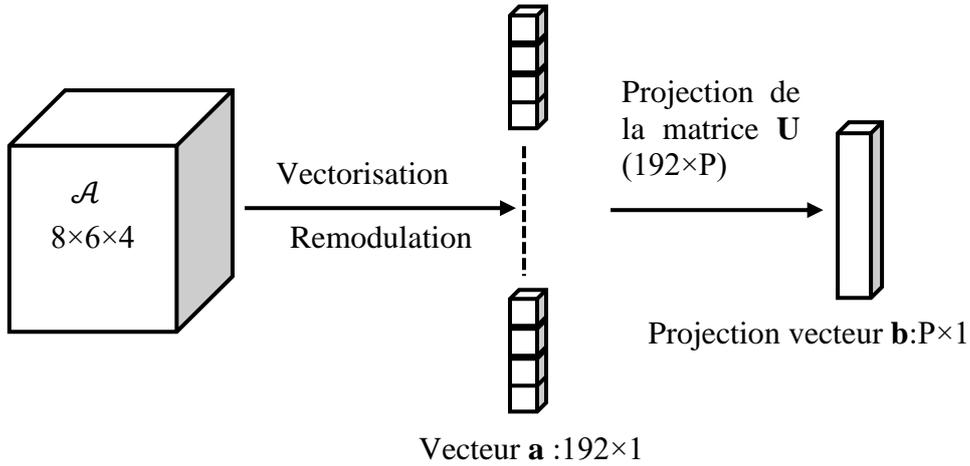


Figure 3.7 Illustration de la projection : vecteur à vecteur (PVV) [195].

3.2.3.2 Projection : tenseur à tenseur (PTT)

Un tenseur peut être projeté sur un autre tenseur du même ordre, comme un tenseur \mathcal{X} d'ordre N réside dans l'espace tensoriel $\mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \dots \otimes \mathbb{R}^{I_N}$ [189, 194], l'espace tensoriel peut être considéré comme le produit de Kronecker de N espaces vectoriels (linéaires) $\mathbb{R}^{I_1}, \mathbb{R}^{I_2}, \dots, \mathbb{R}^{I_N}$. Pour projeter un tenseur \mathcal{X} dans un espace tensoriel $\mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \dots \otimes \mathbb{R}^{I_N}$ à un autre tenseur \mathcal{Y} dans un espace tensoriel inférieur $\mathbb{R}^{P_1} \otimes \mathbb{R}^{P_2} \dots \otimes \mathbb{R}^{P_N}$, où $P_n \leq I_n$ pour toutes les n , les N matrices de projection $\{\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times P_n}, n = 1, \dots, N\}$ (avec colonnes orthonormées) sont utilisées de telle sorte que [196]:

$$\mathcal{Y} = \mathcal{X} \times_1 \mathbf{U}^{(1)T} \times_2 \mathbf{U}^{(2)T} \dots \times_N \mathbf{U}^{(N)T} \quad (3.19)$$

Ces N matrices de projection utilisées peuvent être écrites de manière concise sous la forme $\{\mathbf{U}^{(N)}\}$. La Figure 3.8 illustre la projection d'un objet tenseur à un tenseur de dimension plus faible $P_1 \times P_2 \times P_3$. La manière dont cette projection multilinéaire est réalisée peut-être mieux comprise en se référant à l'illustration de la multiplication en mode n sur la Figure 3.5.

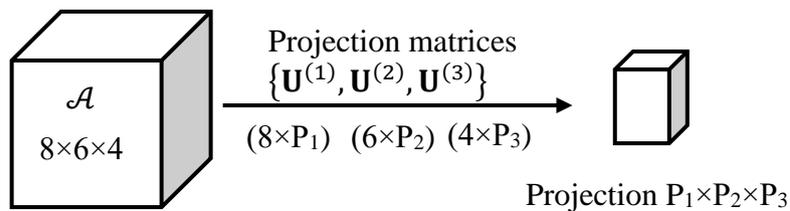


Figure 3.8 Projection Tenseur à tenseur (PTT) [195].

3.2.3.3 Projection : tenseur à vecteur (PTV)

La troisième projection multilinéaire est d'un espace tensoriel à un espace vectoriel, et on l'appelle la projection tenseur à vecteur. Il est formulé sur la base du modèle CANDECOMP / PARAFAC.

Un vecteur peut être considéré comme plusieurs scalaires. Aussi, la projection d'un tenseur à un vecteur peut être considérée comme des multiples projections, chacune projette un tenseur à un scalaire, comme indiqué la **Figure 3.9**. La projection d'un tenseur $\mathcal{A} \in \mathbb{R}^{8 \times 6 \times 4}$ à un vecteur $P \times 1$ consiste de P projections, chacune projette \mathcal{A} à un scalaire. Ainsi, la projection d'un tenseur vers un scalaire est visualisée comme des blocs de construction et elle est considérée en premier.

Un tenseur du second ordre (matrice) $\mathbf{X} \in \mathbb{R}^{I_1 \times I_2}$ peut être projeté à un scalaire y à travers deux vecteurs de projection unitaire $\mathbf{u}^{(1)}$ et $\mathbf{u}^{(2)}$ comme,

$$y = \mathbf{X} \times_1 \mathbf{u}^{(1)T} \times_2 \mathbf{u}^{(2)T} = \mathbf{u}^{(1)T} \mathbf{X} \mathbf{u}^{(2)}, \|\mathbf{u}^{(1)}\| = \|\mathbf{u}^{(2)}\| = 1 \quad (3.20)$$

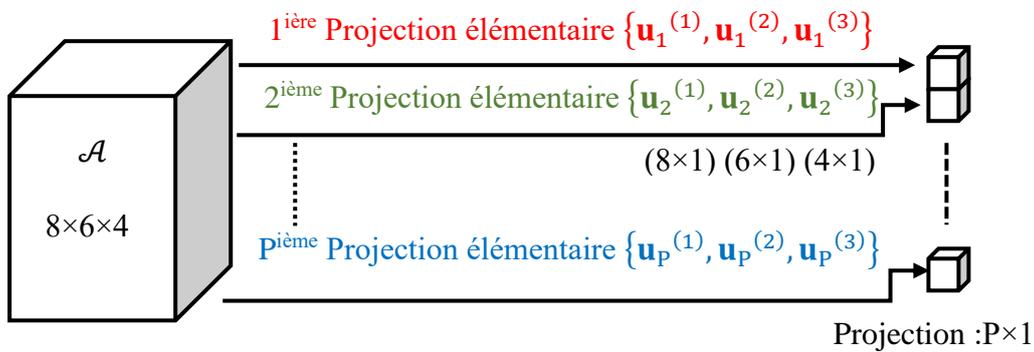


Figure 3.9 Projection tenseur à vecteur (PTV) [195].

où $\|\cdot\|$ est la norme Euclidienne pour les vecteurs qui peut être écrit comme le produit intérieur entre \mathbf{X} et les produits extérieurs du $\mathbf{u}^{(1)}$ et $\mathbf{u}^{(2)}$:

$$y = \langle \mathbf{X}, \mathbf{u}^{(1)} \mathbf{u}^{(2)T} \rangle \quad (3.21)$$

Un tenseur général $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ peut être projeté en un point y à travers N vecteurs de projection unitaires $\{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \dots, \mathbf{u}^{(N)}\}$, qui peuvent aussi s'écrire $\{\mathbf{u}^{(n)}, n = 1, \dots, N\}$ ou simplement comme $\{\mathbf{u}^{(n)}\}$:

$$y = \mathcal{X} \times_1 \mathbf{u}^{(1)T} \times_2 \mathbf{u}^{(2)T} \dots \times_N \mathbf{u}^{(N)T}, \|\mathbf{u}^{(n)}\| = 1 \text{ pour } n = 1, \dots, N \quad (3.22)$$

Il peut être écrit comme le produit scalaire ((3.10) du \mathcal{X} et le produit externe du $\{\mathbf{u}^{(n)}\}$:

$$y = \langle \mathcal{X}, \mathbf{u}^1 \circ \mathbf{u}^2 \circ \dots \circ \mathbf{u}^N \rangle \quad (3.23)$$

Notons $\mathcal{U} = \mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ \dots \circ \mathbf{u}^{(N)}$, alors $y = \langle \mathcal{X}, \mathcal{U} \rangle$. Cette projection multilinéaire à travers $\{\mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ \dots \circ \mathbf{u}^{(N)}\}$ est appelée une projection multilinéaire élémentaire (PME) [197], qui est une projection d'un tenseur à un scalaire. Il s'agit d'une projection de rang un et il se compose d'une projection de vecteur dans chaque mode. La **Figure 3.10** illustre une projection élémentaire multilinéaire du tenseur $\mathcal{A} \in \mathbb{R}^{8 \times 6 \times 4}$.

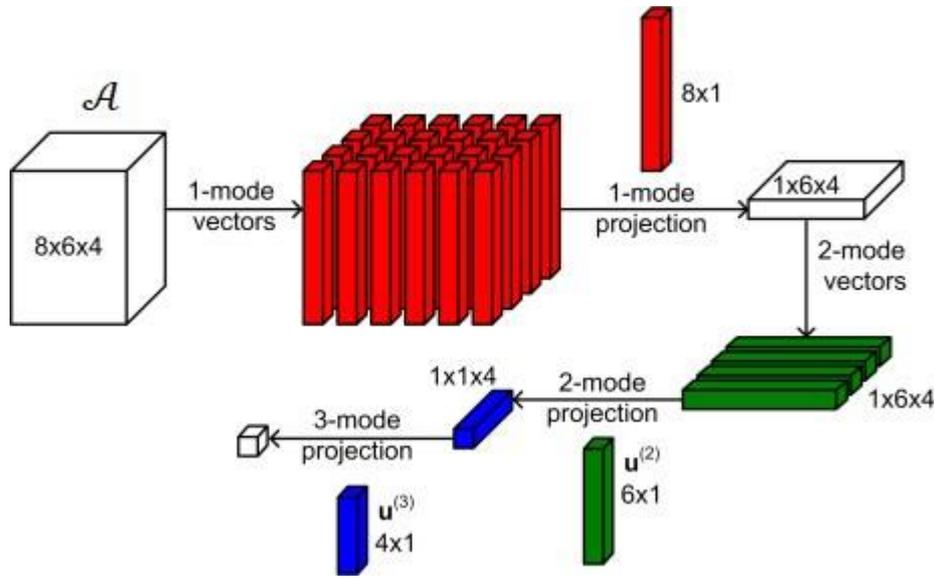


Figure 3.10 Illustration de la projection d'un tenseur à vecteur à travers la projection multilinéaire élémentaire [195].

Ainsi, la projection du tenseur \mathcal{X} à vecteur $\mathbf{y} \in \mathbb{R}^P$ d'un objet tensoriel dans un espace vectoriel de P-dimension constitué de P projections multilinéaires élémentaires $\{\mathbf{u}_p^{(1)T}, \mathbf{u}_p^{(2)T}, \dots, \mathbf{u}_p^{(N)T}\}, p = 1, \dots, P$, qui peut être écrit de manière concise comme $\{\mathbf{u}_p^{(n)T}, n = 1, \dots, N\}_{p=1}^P$. Alors, la projection tenseur \mathcal{X} à vecteur \mathbf{y} est donnée par :

$$\mathbf{y} = \mathcal{X} \times_{n=1}^N \left\{ \mathbf{u}_p^{(n)T}, n = 1, \dots, N \right\}_{p=1}^P \quad (3.24)$$

Où le $P^{\text{ième}}$ composant de \mathbf{y} est obtenu à partir du $P^{\text{ième}}$ de la projection multilinéaire élémentaire :

$$\mathbf{y}(p) = \mathcal{X} \times_1 \mathbf{u}^{(1)T} \times_2 \mathbf{u}^{(2)T} \dots \times_N \mathbf{u}^{(N)T} \quad (3.25)$$

3.2.4 Relation entre les projections multilinéaires

Après on a vu les trois projections multilinéaires de base possibles. Il est nécessaire d'étudier leurs relations.

- **Conditions dégénérées :** il est facile de vérifier que la projection vecteur à vecteur est un cas particulier de la projection tenseur à tenseur et aussi la projection tenseur à vecteur pour $N=1$. La projection multilinéaire élémentaire est la version dégénérée de la projection tenseur à tenseur avec $P_n = 1$ pour tout n .

- **Relation entre la projection multilinéaire et PTT :**

Chaque élément projeté dans la projection tenseur-à-tenseur peut être vu comme une projection multilinéaire élémentaire formée en prenant une colonne de chaque matrice de projection. Ainsi, un tenseur projeté dans la projection tenseur à tenseur est obtenu par $\prod_{n=1}^N P_n$ projections multilinéaires élémentaires avec vecteurs de projection partagés (à partir des matrices de projection) en effet, alors que dans la projection tenseur à vecteur, les P projections multilinéaires élémentaires n'ont pas de vecteurs de projection partagés.

- **Equivalence entre PME et PVV :**

Rappelons que la projection utilisant une projection multilinéaire élémentaire $\{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \dots, \mathbf{u}^{(N)}\}$ peut-être écrite par :

$$\mathbf{y} = \langle \mathcal{X}, \mathcal{Y} \rangle = \langle \text{vec}(\mathcal{X}), \text{vec}(\mathcal{Y}) \rangle = [\text{vec}(\mathcal{Y})]^T \text{vec}(\mathcal{X}) \quad (3.26)$$

Ainsi, une projection multilinéaire élémentaire est équivalente à une projection linéaire de $\text{vec}(\mathcal{X})$, la représentation vectorielle du \mathcal{X} par un vecteur $\text{vec}(\mathcal{Y})$ est donnée par l'équation (3.18). Puisque $\mathcal{U} = \mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ \dots \circ \mathbf{u}^{(N)}$, l'équation (3.26) indique que la projection multilinéaire élémentaire est équivalente à une projection linéaire avec une contrainte sur le vecteur de projection telle que s'agit de la représentation vectorisée d'un tenseur de rang un.

- **Equivalence entre PTT et PTV :**

Donnant une projection tenseur à vecteur $\{\mathbf{u}_p^{(n)}\}_N^P$, on peut former N matrices $\{\mathbf{V}^n\}$, où

$$\mathbf{V}^n = [\mathbf{u}_1^{(n)}, \dots, \mathbf{u}_p^{(n)}, \dots, \mathbf{u}_p^{(n)}] \in \mathbb{R}^{I_n \times P} \quad (3.27)$$

Ces matrices peuvent être vues comme une projection tenseur à tenseur où $P_n = P$ pour $n = 1, \dots, N$. Ainsi, la projection d'un tenseur par un vecteur $\{\mathbf{u}_p^{(n)}\}_N^P$ est équivalente à la diagonale

de la projection tenseur à tenseur correspondante du même tenseur par $\{\mathbf{V}^n\}$ tel que défini dans l'équation (3.27):

$$\mathbf{y} = \mathcal{X} \times_{n=1}^N \left\{ \mathbf{u}_p^{(n)} \right\}_N^P = \text{diag}(\mathcal{X} \times_1 \mathbf{V}^{(1)T} \mathcal{X} \times_2 \mathbf{V}^{(2)T} \dots \mathcal{X} \times_N \mathbf{V}^{(N)T}) \quad (3.28)$$

Dans le cas du second ordre, cette équivalence est

$$\mathbf{y} = \text{diag}(\mathbf{V}^{(1)T} \mathbf{X} \mathbf{V}^{(2)}) \quad (3.29)$$

▪ **Nombre de paramètres à estimer :**

Le nombre de paramètres à estimer dans une projection particulière est une préoccupation importante dans la pratique. Comparant la projection vecteur de taille $I \times 1$ dans la projection vecteur à vecteur spécifié par I les paramètres ($I = \prod_{n=1}^N I_n$) pour un tenseur d'ordre n . la projection multilinéaire élémentaire dans la projection tenseur à vecteur peut être spécifiée par $\sum_{n=1}^N I_n$ paramètres. Ainsi, $P \cdot \sum_{n=1}^N I_n$ est le nombre de paramètres doit être estimé pour projeter un tenseur de taille $\prod_{n=1}^N I_n$ sur un vecteur de taille $P \times 1$. Tandis que, la projection vecteur à vecteur besoin seulement $P \cdot \prod_{n=1}^N I_n$ paramètres estimés. L'implication est que la projection tenseur à vecteur a moins de paramètres à estimer alors qu'étant plus contraint sur les solutions, et la projection tenseur à tenseur a moins de contrainte sur les solutions recherchées tout en ayant plus de paramètres à estimer. En d'autres termes, la projection tenseur à vecteur a un modèle plus simple que la projection vecteur à vecteur. Pour la projection tenseur à tenseur avec la même quantité de réduction de dimensionnalité, il y a $\sum_{n=1}^N P_n \times I_n$ paramètres doivent être estimés. Ainsi, en raison des vecteurs de projection partagés, la projection tenseur à tenseur peut avoir besoin d'estimer encore moins de paramètres et son modèle peut être encore plus simple.

Le **Tableau 3.1** compare le nombre de paramètres à estimer par les trois projections pour la même réduction de dimensionnalité pour divers cas. La **Figure 3.11** illustre en outre les trois premiers cas, où le nombre de paramètres est normalisé par rapport à celui de la projection vecteur à vecteur pour une meilleure visualisation. À partir du tableau et de la figure, nous pouvons voir que pour les tenseurs d'ordre supérieur, le modèle de la projection vecteur à vecteur conventionnel devient extrêmement complexe et l'estimation des paramètres devient extrêmement difficile. Pratiquement, cela conduit souvent au problème de petite taille de l'échantillon (SSS, smallsample size) lorsque le nombre d'échantillons d'apprentissage disponibles est limité.

Tableau 3.1 Nombre de paramètre à estimer par les trois projections multilinéaires [195].

Entrée	Sortie	PVV	PTV	PTT
$\prod_{n=1}^N I_n$	P	$P \cdot \prod_{n=1}^N P_n$	$P \cdot \sum_{n=1}^N I_n$	$\sum_{n=1}^N P_n \times I_n$
10×10	4	400	80	40($P_n = 2$)
100×100	4	40,000	800	400($P_n = 2$)
100×100×100	8	8,000,000	2400	600($P_n = 2$)
$\prod_{n=1}^4 100$	16	1,600,000,000	6400	800($P_n = 2$)

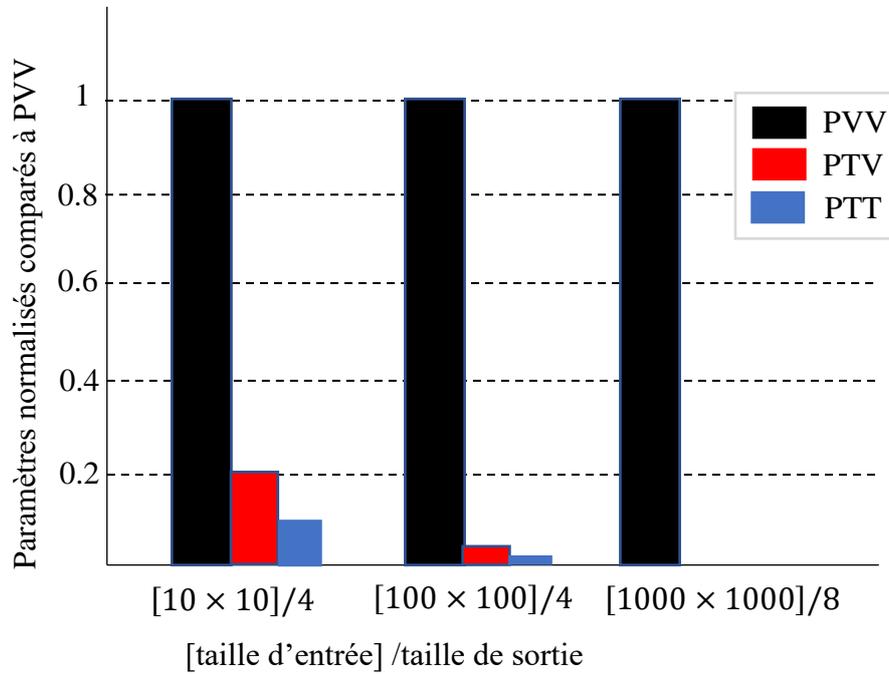


Figure 3.11 Comparaison du nombre de paramètres à estimer par PVV, PTV et PTT normalisés par rapport au nombre VVP pour visualisation [195].

3.3 Analyse en Composante Principale Multilinéaire (MPCA)

MPCA (Multilinear Principal Component Analysis) est l'extension de PCA proposée par Lu et al[9] pour résoudre le problème de la représentation des objets multidimensionnels. Pour comprendre l'idée de MPCA, soit un ensemble de M d'objets tensoriels $\{\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_M\}$. Chaque tenseur d'objet $\mathcal{A}_m \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ prend des valeurs dans un espace tensoriel $\mathbb{R}^{I_1} \otimes \mathbb{R}^{I_2} \dots \otimes \mathbb{R}^{I_N}$, où I_n est la dimension de mode- n du tenseur. L'objectif de MPCA est de définir une projection tenseur à tenseur $\{\mathbf{U}^{(n)} \in \mathbb{R}^{I_n \times P_n}, n = 1, \dots, N\}$ qui mappe l'espace tensoriel original $\mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ dans un sous-espace tensoriel $\mathbb{R}^{P_1 \times P_2 \times \dots \times P_N}$ (avec $P_n \leq I_n$, pour $n = 1, \dots, N$) :

$$\mathcal{B}_m = \mathcal{A}_m \times_1 \mathbf{U}^{(1)T} \times_2 \mathbf{U}^{(2)T} \dots \times_N \mathbf{U}^{(N)T}, m = 1, \dots, M \quad (3.30)$$

Tel que $\{\mathcal{B}_m \in \mathbb{R}^{P_1} \otimes \mathbb{R}^{P_2} \dots \otimes \mathbb{R}^{P_N}, m = 1, \dots, M\}$ capture la plupart de la variation observée dans les objets tensoriels d'origine, en supposant que ces variations sont mesurées par la dispersion totale définie pour les tenseurs. En d'autres termes, l'objectif de MPCA est de déterminer des N projections de la matrices $\mathbf{U}^{(n)}$ qui maximisent la dispersion totale $\psi_{\mathcal{B}}$ du tenseur :

$$\{\mathbf{U}^{(n)}, n = 1, \dots, N\} = \arg \max_{\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \dots, \mathbf{U}^{(N)}} \psi_{\mathcal{B}} \quad (3.31)$$

Où $\psi_{\mathcal{B}} = \sum_{m=1}^M \|\mathcal{B}_m - \bar{\mathcal{B}}\|_F^2$ et $\bar{\mathcal{B}}$ est le moyen du tenseur calculé par

$$\bar{\mathcal{B}} = \frac{1}{M} \sum_{n=1}^N \mathcal{B}_n \quad (3.32)$$

La matrice de dispersion totale en mode k de N échantillons est définie comme suit :

$$\mathbf{S}_{T_{\mathcal{B}}}^{(k)} = \sum_{n=1}^N (\mathbf{B}_{n(k)} - \bar{\mathbf{B}}_{(k)}) (\mathbf{B}_{n(k)} - \bar{\mathbf{B}}_{(k)})^T \quad (3.33)$$

où $\mathbf{B}_{n(k)}$ est la matrice dépliée en mode k du tenseur \mathcal{B}_n .

Une autre extension appelée Whitened MPCA (MWPCA) est proposée dans [198] pour résoudre le problème de la petite taille d'échantillon (Small Sample Size) dans l'espace de grand dimension et pour améliorer le pouvoir de la discrimination obtenu par l'analyse en composantes principales multilinéaire classique (MPCA). Alors, Le blanchiment (whitening) transforme linéairement les données pour que sa covariance devienne la matrice d'identité.

Après le blanchiment, les données deviennent moins corrélées et avec une variance uniforme dans toutes les directions. Les matrices blanchies de projection sont obtenues en multipliant les vecteurs propres sélectionnés par la racine carrée des valeurs propres correspondantes comme suit :

$$\mathbf{W}_k = \Gamma_{(k)}^{-1/2} \mathbf{U}_{(k)} \quad (3.34)$$

Où $\Gamma_{(k)} = [\gamma_1 \gamma_2 \cdots \gamma_{n(k)}]$ et $\mathbf{U}_{(k)}$ sont les vecteurs propres sélectionnés correspondants au $n(k)$ plus grandes valeurs propres.

3.4 Analyse Discriminante Multilinéaire (MDA)

MDA (Multilinear Discriminant Analysis) est une approche de réduction de dimensionnalité supervisée proposée dans [10] pour résoudre des problèmes de sous-espace en dépliant le tenseur dans différentes directions. Dans notre algorithme, nous considérons les échantillons d'apprentissage obtenus par MWPCA comme des échantillons d'apprentissage d'entrée pour l'algorithme MDA afin d'obtenir des échantillons plus réduits à la sortie.

Les échantillons d'apprentissage sont représentés par un tenseur $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_m}$ d'ordre m appartenant aux différentes classes C , où l'ensemble de données d'apprentissage peut être montré comme un tenseur $\check{\mathcal{A}} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_m \times M}$ d'ordre $(m+1)$, où M est le nombre d'échantillons d'apprentissage.

L'algorithme MDA est caractérisé par plusieurs sous-espaces interdépendants, qui peuvent coopérer pour discriminer différentes classes en maximisant la dispersion inter-classe et en minimisant au même temps la dispersion intra-classe dans chaque mode du tenseur.

$$\mathbf{U}_k^* \Big|_{k=1}^m = \arg \max_{\mathbf{U}_k \Big|_{k=1}^m} \frac{\sum_{j=1}^C n_j \|\bar{\mathcal{A}}_j \times_1 \mathbf{U}_1 \cdots \times_m \mathbf{U}_m - \bar{\mathcal{A}} \times_1 \mathbf{U}_1 \cdots \times_m \mathbf{U}_m\|^2}{\sum_{i=1}^M \|\mathcal{A}_i \times_1 \mathbf{U}_1 \cdots \times_m \mathbf{U}_m - \bar{\mathcal{A}}_{n_i} \times_1 \mathbf{U}_1 \cdots \times_m \mathbf{U}_m\|^2} \quad (3.35)$$

Où n_j est le nombre d'échantillons d'apprentissage de la classe C , $\bar{\mathcal{A}}_j$ est le moyen du tenseur des échantillons d'apprentissage appartenant à la classe C et $\bar{\mathcal{A}}$ est le moyen total du tenseur de tous les échantillons d'apprentissage.

Dans l'équation (3.35), il y a un problème d'optimisation non linéaire d'ordre supérieur avec une contrainte non linéaire d'ordre supérieur. La découverte d'une solution fermée simple n'est pas évidente. Par conséquent, une approche d'optimisation itérative pour estimer des sous-espaces discriminants interdépendants est proposée.

$$\mathbf{U}_k^* = \arg \max_{\mathbf{U}_k} \frac{\sum_{j=1}^C n_j \|\bar{\mathcal{A}}_j \times_k \mathbf{U}_k - \bar{\mathcal{A}} \times_k \mathbf{U}_k\|^2}{\sum_{i=1}^M \|\mathcal{A}_i \times_k \mathbf{U}_k - \bar{\mathcal{A}}_{n_i} \times_k \mathbf{U}_k\|^2} \quad (3.36)$$

Ce problème d'optimisation se réduit à une analyse discriminante spéciale où le tenseur se déplie dans le mode-k et les vecteurs colonnes des matrices dépliées sont étiquetés avec l'étiquette originale du tenseur. Le problème d'optimisation dans l'équation (3.36) est reformulé comme une analyse discriminante spéciale :

$$\mathbf{U}_k^* = \arg \max_{\mathbf{U}_k} \frac{\text{Tr}(\mathbf{U}_k^T \mathbf{S}_b \mathbf{U}_k)}{\text{Tr}(\mathbf{U}_k^T \mathbf{S}_w \mathbf{U}_k)} \quad (3.37)$$

où \mathbf{S}_b et \mathbf{S}_w sont les matrices de dispersion inter-classe et intra-classe respectivement données par :

$$\mathbf{S}_b^k = \sum_{j=1}^L n_j (\bar{\mathbf{A}}_j^k - \bar{\mathbf{A}}^k)(\bar{\mathbf{A}}_j^k - \bar{\mathbf{A}}^k)^T \quad (3.38)$$

$$\mathbf{S}_w^k = \sum_{j=1}^L \sum_{i=1}^{n_j} (\mathbf{A}_{j,i}^k - \bar{\mathbf{A}}_j^k)(\mathbf{A}_{j,i}^k - \bar{\mathbf{A}}_j^k)^T \quad (3.39)$$

La matrice $\mathbf{A}_{j,i}^k$ est la matrice dépliée en mode k du tenseur \mathcal{A}_i , $\bar{\mathbf{A}}_j^k$ est la matrice moyenne de la classe j et $\bar{\mathbf{A}}^k$ est la matrice moyenne de données d'apprentissages globales.

A chaque itération de la solution d'équation (3.35), les matrices $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_m$ sont censées être connues, conduisons à :

$$\begin{aligned} & \mathbf{U}_k^* \Big|_{k=1}^m \\ &= \arg \max_{\mathbf{U}_k \Big|_{k=1}^m} \frac{\sum_{j=1}^C n_j \|\bar{\mathcal{A}}_j \times_1 \mathbf{U}_1 \cdots \times_m \mathbf{U}_m - \bar{\mathcal{A}} \times_1 \mathbf{U}_1 \cdots \times_m \mathbf{U}_m\|^2}{\sum_{i=1}^M \|\mathcal{A}_i \times_1 \mathbf{U}_1 \cdots \times_m \mathbf{U}_m - \bar{\mathcal{A}}_{n_i} \times_1 \mathbf{U}_1 \cdots \times_m \mathbf{U}_m\|^2} \end{aligned} \quad (3.40)$$

Notons $\mathcal{B}_i = \mathcal{A}_i \times_1 \mathbf{U}_1 \cdots \times_{k-1} \mathbf{U}_{k-1} \times_{k+1} \mathbf{U}_{k+1} \cdots \times_m \mathbf{U}_m$, nous obtenons :

$$\mathbf{U}_k^* = \arg \max_{\mathbf{U}_k} \frac{\sum_{j=1}^C n_j \|\bar{\mathcal{B}}_j \times_k \mathbf{U}_k - \bar{\mathcal{B}} \times_k \mathbf{U}_k\|^2}{\sum_{i=1}^M \|\mathcal{B}_i \times_k \mathbf{U}_k - \bar{\mathcal{B}}_{n_i} \times_k \mathbf{U}_k\|^2} \quad (3.41)$$

Cette équation est équivalente à l'équation (3.36) en changeant \mathcal{B}_i par \mathcal{A}_i . Par conséquent, il peut être résolu à l'aide de l'équation (3.37).

Les matrices de projection peuvent être optimisées de manière itérative pour chaque mode. Le processus itératif de MDA s'arrête lorsqu'un nombre maximal d'itérations est atteint ou la différence entre deux itérations successives est inférieure au seuil $n'_k n_k \varepsilon$.

$$\|\mathbf{U}_k^{\text{itération}} - \mathbf{U}_k^{\text{itération-1}}\| < n'_k n_k \varepsilon, \quad \mathbf{U}_k^{\text{itération}} \in \mathbb{R}^{n'_k \times n_k}.$$

3.5 Normalisation de la Covariance intra-classe (WCCN)

WCCN (Within Class Covariance Normalization) est une méthode de normalisation appliquée à l'origine dans le système de reconnaissance de locuteur basé sur SVM[11]. Le WCCN est proposé pour minimiser le taux d'erreurs de classification attendu sur les échantillons d'apprentissage en utilisant l'inverse de la matrice de covariance intra-classe \mathbf{W}_k , comme :

$$\mathbf{W}_k = \sum_j^C \sum_i^M ((\mathbf{U}_k^*)^T \mathbf{A}_{j,i}^k - (\mathbf{U}_k^*)^T \bar{\mathbf{A}}_j^k) ((\mathbf{U}_k^*)^T \mathbf{A}_{j,i}^k - (\mathbf{U}_k^*)^T \bar{\mathbf{A}}_j^k)^T \quad (3.42)$$

où \mathbf{U}_k^* est la matrice de projection MDA dans chaque mode calculée par l'équation (3.37). Quand on applique la méthode WCCN, chaque vecteur de caractéristiques de la matrice \mathbf{U}_k^* est prémultiplié par une matrice triangulaire supérieure \mathbf{Z} obtenue par décomposition de Cholesky de la matrice $\mathbf{W}_k^{-1} = \mathbf{Z}_k \mathbf{Z}_k^T$.

3.6 L'algorithme MDA+WCCN

L'approche proposée dans ce mémoire de thèse est basée sur deux méthodes multilinéaires, MWPCA et MDA+WCCN. La première méthode est utilisée pour bien organiser les données multidimensionnelles en tenseur au lieu de les représentées par les méthodes traditionnelles, telles que PCA et LDA qui se basent sur la vectorisation des données ce qui introduit la destruction de la nature de ces données. La deuxième méthode est une intégration entre MDA et WCCN. MDA est généralement utilisée pour la réduction de dimensionnalité d'une parte, et à la classification d'une autre parte. L'intégration de WCCN permet d'améliorer les résultats de la classification par MDA et de minimiser les erreurs qui se produisent durant le processus de classification. L'algorithme complet de l'approche proposée est détaillé comme suit:

Tableau 3.2 L'algorithme de l'approche proposée

Algorithme 1 : MDA+WCCN**Entrée :**

- Le tenseur $\tilde{\mathcal{A}} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_m \times M}$ de M échantillons d'apprentissage appartenant à la classe C et chaque classe \mathcal{A}_j comprend n_j échantillons.
- Le nombre maximal d'itérations : itr_{\max}
- Dimensions inférieurs finales : $I'_1 \times I'_2 \times \dots \times I'_m$

Sortie :

Matrices de projections : $U_k = U_k^t \in \mathbb{R}^{I_k \times I'_k}, k = 1, 2, \dots, m.$

Algorithme :

1. Initialisation : $U_1^0 = I_{I_1}, U_2^0 = I_{I_2}, \dots, U_m^0 = I_{I_m}.$
2. Pour : $t=1, 2, \dots, \text{itr}_{\max}$ faire
 - a) Pour $k=1, 2, \dots, m$ faire
 - $B_i = \mathcal{A}_i \times_1 U_1^t \dots \times_{k-1} U_{k-1}^t \times_{k+1} U_{k+1}^{t-1} \dots \times_m U_m^{t-1}$
 - $A_i^k \leftarrow \mathcal{A}_i$
 - $S_b^k = \sum_{j=1}^L n_j (\bar{A}_j^k - A^k) (\bar{A}_j^k - \bar{A}^k)^T$
 - $S_w^k = \sum_{j=1}^L \sum_{i=1}^{n_j} (A_{j,i}^k - \bar{A}_j^k) (A_{j,i}^k - \bar{A}_j^k)^T$
 - $W_k = \sum_j^C \sum_i^M ((U_k^*)^T A_{j,i}^k - (U_k^*)^T \bar{A}_j^k) ((U_k^*)^T A_{j,i}^k - (U_k^*)^T \bar{A}_j^k)^T$
 - Calculer les vecteurs propres et valeurs propres : $U_k^t = S_w^{k-1} \times S_b^k$
 - Calculer la matrice de projection de WCCN: $WCCN = \text{chol}(W_k^{-1})$
 - Calculer la nouvelle matrice de projection $U_k^t = WCCN^T U_k^t$
 - b) Si $t > 2$ et $\|U_k^t - U_k^{t-1}\| < I_k I'_k \varepsilon, k = 1, \dots, m$ pause et aller à l'étape 3.
3. Projection : $U_k = U_k^t \in \mathbb{R}^{I_k \times I'_k}, k = 1, \dots, m$

3.7 Conclusion :

Dans ce chapitre, nous avons introduit une formulation générale de l'apprentissage de sous espace multilinéaire. Nous avons présenté les fondamentaux des tenseurs, y compris les opérations multilinéaires de base et une mesure de distance basée sur un tenseur. Ainsi, nous avons introduit les trois types de projection multilinéaire : la projection de vecteur à vecteur, la

projection de tenseur à tenseur et la projection de tenseur à vecteur. De plus, les connexions entre ces trois projections sont révélées. Enfin, nous avons étudié deux techniques multilinéaires conçues pour la réduction de dimensionnalité. Ces deux techniques sont MWPCA et MDA+WCCN qui constituent la base de notre approche.

Généralement, l'utilisation des vecteurs pour la représentation des objets multidimensionnels conduit souvent au problème de petite taille de l'échantillon (small sample size problem). Ce problème se pose lorsque le nombre de paramètres est bien supérieurs au nombre des échantillons et la possibilité de le réduire est extrêmement difficile. Tandis que, l'utilisation des tenseurs donnent une représentation naturelle pour les objets multidimensionnels. Ils sont appliqués directement sur les données sans les transformées en vecteurs. En plus, le nombre de paramètres à estimer par les projections multilinéaires est plus réduit par rapport au nombre de paramètres à estimer par les projections vectorielles. Nous concluons que les approches multilinéaires sont plus efficaces que les approches classiques. Les deux approches multilinéaires MPCA et MDA sont des extensions des approches linéaires classique PCA et LDA respectivement. Elles sont capables de résoudre le problème de la grande dimensionnalité, ainsi qu'augmenter les résultats de reconnaissance faciale.

Chapitre 4

Résultats expérimentaux

4.1 Introduction.....	75
4.2 Aperçu de l’approche de vérification de visage proposée.....	75
4.2.1 Pré-traitement de visage 2D et 3D.....	75
4.2.2 Extraction des caractéristiques	75
4.2.2.1 Descripteur LPQ.....	76
4.2.2.2 Descripteur BSIF.....	75
4.2.3 Réduction de la dimensionalité, comparaison et décision	80
4.3 Résultats expérimentaux	80
4.3.1 Les bases de données	81
4.3.1.1 FRGC v2.0	81
4.3.1.2 Bosphorus.....	82
4.3.1.3 CASIA 3D.....	83
4.4 Configuration de paramètres	83
4.5 Discussion et comparaison.....	83
4.5.1 Descripteurs locaux LPQ et BSIF.....	84
4.5.2 L’approche proposée par rapport à la méthode MDA	86
4.5.3 Les performances de la modalité 2D et 3D.....	88
4.5.4 Les performances de la vérification multimodale de visage 2D + 3D.....	89
4.5.5 L’approche proposée VS l’état de l’art	90
4.6 Conclusion	93

Chapitre 4 : Résultats et Discussions

4.1 Introduction

Dans ce chapitre, nous détaillons le système proposé pour vérifier le visage d'une personne en utilisant la fusion multimodale des images faciales 2D et 3D. Notre système se déroule en deux phases : la phase d'apprentissage et celle de teste. Chaque phase est constituée de quatre étapes principales qui sont : le pré-traitement, l'extraction des caractéristiques, la réduction de la dimensionalité et la comparaison. Ensuite, nous présentons les bases de données utilisées pour tester notre algorithme MDA+WCCN. En plus, nous commentons les résultats obtenus à partir de l'exécution de notre algorithme. En fin, nous comparons nos résultats avec les travaux récent dans littérature afin de prouver la performance et l'efficacité de l'approche proposée.

4.2 Approche proposée pour la vérification de visage

La **Figure 4.1** représente le schéma fonctionnel de notre système de vérification multimodale de visage 2D+3D. nous expliquons en détail ci-dessus étape par étape l'approche proposée.

4.2.1 Pré-traitement de visage 2D et 3D

L'étape de pré-traitement est nécessaire pour chaque système de reconnaissance faciale pour améliorer la qualité des données 2D et 3D. Tout d'abord, la région d'image 2D et 3D est détectée et recadrée avec une sphère de rayon de 80 mm centrée au bout du nez. Pour supprimer les pics générés par les scans 3D, un filtre médian est appliqué uniquement sur les sommets prédéfinis comme les pics potentiels. Après avoir retiré toutes les pointes, les trous apparus en raison de l'auto-occlusion sont remplis à l'aide d'une interpolation cubique des pixels voisins. Un processus d'enregistrement itératif basé sur l'algorithme ICP [199] est utilisé pour ajuster l'alignement des images de visage 2D et 3D à la base de données d'origine (expression neutre). Un deuxième recadrage d'images de visage 2D et 3D avec une sphère de rayon 75 mm centrée à l'extrémité du nez est effectué pour éliminer le bruit produit par le processus d'alignement. Afin d'avoir la même résolution pour toutes les images, les images de visage 3D et leurs images 2D correspondantes sont rééchantillonnées en une grille carrée uniforme de 1 mm. Un procédé photométrique est ensuite appliqué sur les images faciales 2D afin de réduire les effets de variation de la lumière.

4.2.2 Extraction des caractéristiques

L'extraction des caractéristiques est une étape clé dans le système de reconnaissance faciale. La structure interne des images faciale 2D et 3D sont difficiles à comprendre. Les descripteurs locaux sont connus par leur succès d'analyser les textures locales des images dans plusieurs applications. Dans notre cas, nous avons utilisé deux descripteurs locaux très performants pour décrire les textures locales, sont LPQ (Local Phase Quantization) et BSIF (Binarized Statistical Image Features). De plus, nous présentons une investigation de LPQ et BSIF en les évaluant dans de nombreuses expériences sur des bases de données FRGC v 2.0, Bosphorus et CASIA 3D.

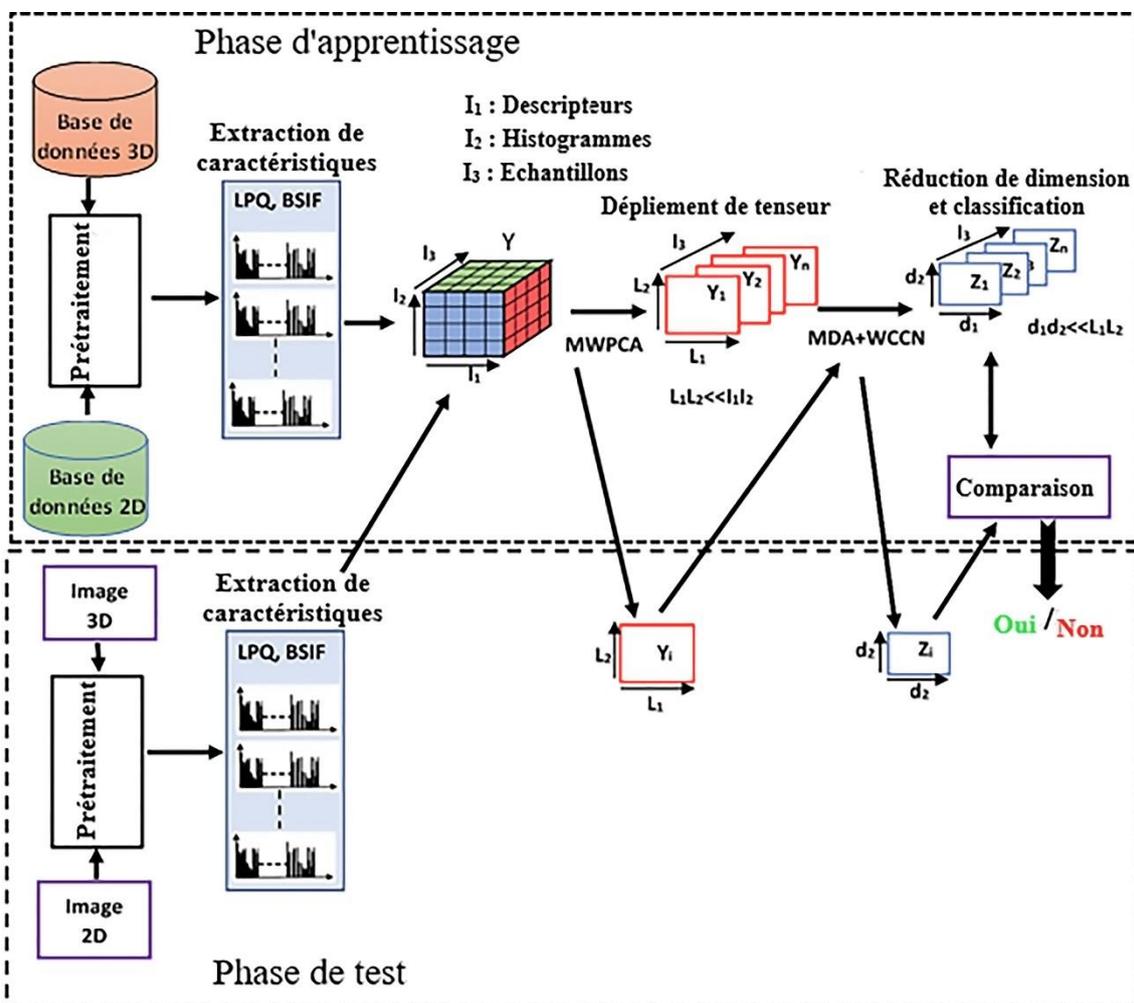


Figure 4.1 Système proposé pour la vérification multimodale de visage 2D+3D.

4.2.2.1 Descripteur LPQ

L'opérateur de quantification de phase locale LPQ a été initialement proposé par Ojansivu et Heikkila[65] pour décrire les textures. L'opérateur LPQ est avéré robuste et performant plus que l'opérateur LBP dans la classification des textures [139]. LPQ est conçu

pour conserver une image dans les informations locales invariantes aux artefacts générés par différentes formes de flou. Inspirés par cette idée, nous proposons le LPQ comme méthode efficace pour résoudre le problème de variations d'expressions.

Le descripteur LPQ est construit sur le concept de quantification de la phase autour d'un bloc $M \times M$ de voisinage N_x pour chaque pixel x de l'image $f(x)$ en utilisant la transformée de Fourier à court terme STFT. Le STFT $F(u, x)$ pour chaque position de pixel x à la fréquence u est donnée par :

$$F(u, x) = \sum_{b \in N_x} f(x - y) e^{-j2\pi u^T y} = w_u^T f_x \quad (4.1)$$

où w_u est le vecteur de base de la transformée discrète de Fourier 2D à la fréquence u , et f_x est un autre vecteur contenant tous les échantillons d'image M^2 de N_x . Seuls les quatre coefficients de basse fréquence correspondant aux fréquences 2D sont pris en compte $u_1 = [a, 0]^T$, $u_2 = [0, a]^T$, $u_3 = [a, a]^T$ et $u_4 = [a, -a]^T$, où a est un suffisamment petit scalaire pour satisfaire $H(u_i) > 0$. Pour chaque position de pixel, il en résulte un vecteur :

$$F_a^c = [F(u_1, x), F(u_2, x), F(u_3, x), F(u_4, x)] \quad (4.2)$$

Les informations de phase dans les coefficients de Fourier sont enregistrées en observant les signes des parties réelles et imaginaires de chaque composant dans F_x . Ceci est fait par

$$q_j = \begin{cases} 1 & \text{si } g_j(x) > 0 \\ 0 & \text{si non} \end{cases} \quad (4.3)$$

où $g_j(x)$ est le $j^{\text{ième}}$ composant de $G(x) = [Re\{F(x)\}, Im\{F(x)\}]$ et $q_j(x)$ sont les huit coefficients binaires représentés sous forme de valeurs entières comprises entre 0 et 255 en utilisant un codage binaire similaire à la méthode LBP comme suit :

$$f_{LPQ}(x) = \sum_{j=1}^8 q_j 2^{j-1} \quad (4.4)$$

En conséquence, nous obtenons l'image d'étiquette f_{LPQ} dont les valeurs sont les étiquettes LPQ invariantes de flou. La **Figure 4.2** représente la structure de générer le code LPQ.

Pour utiliser LPQ dans la description du visage, nous appliquons la procédure de [200]. D'abord, l'image du visage est étiquetée avec l'opérateur LPQ. Ensuite, l'image de l'étiquette est divisée en régions rectangulaires non superposées de taille égale et un histogramme

d'étiquettes est calculé indépendamment dans chaque région. Enfin, les histogrammes de différentes régions sont concaténés pour construire une description du visage.

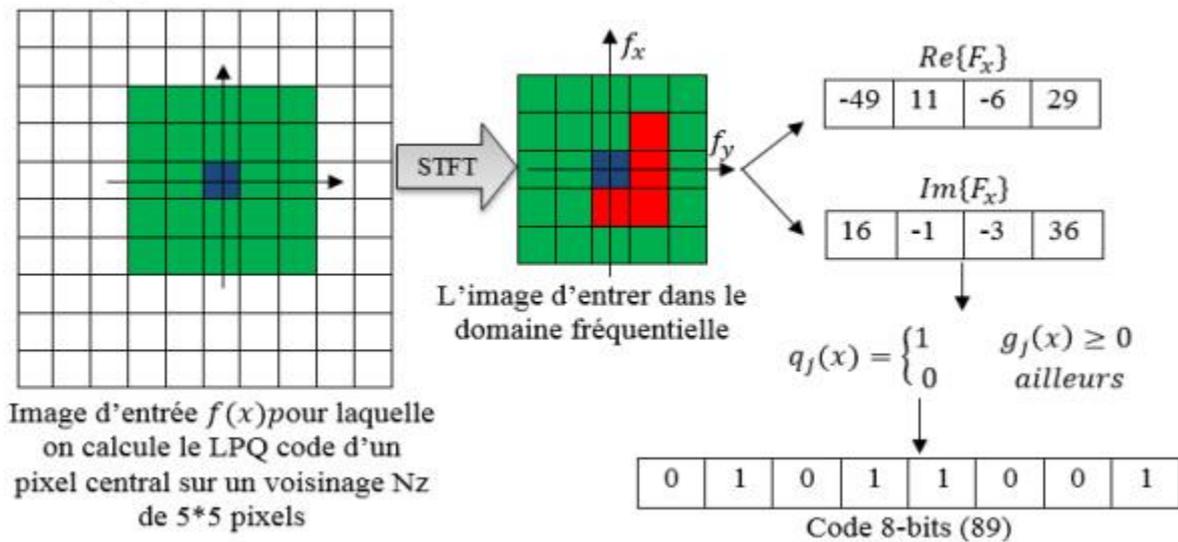


Figure 4.2 Procédures de générer le code LPQ.

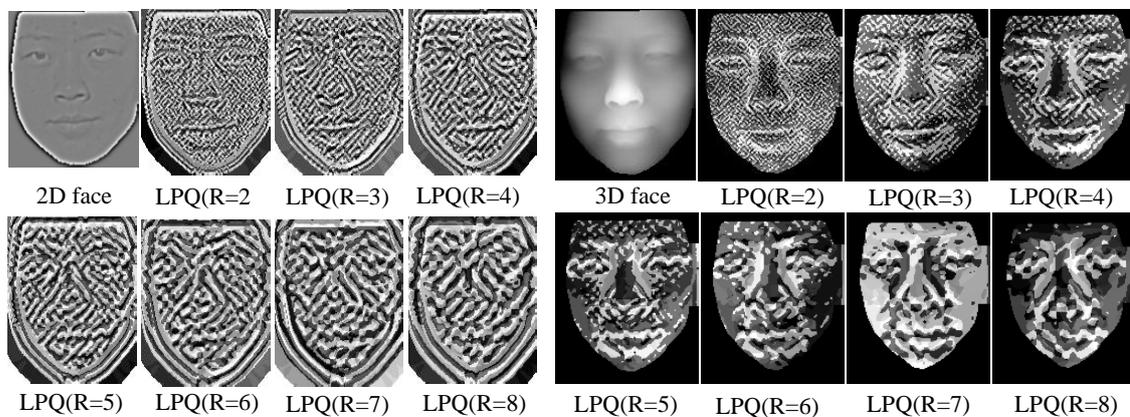


Figure 4.3 Extraction des caractéristiques pour l'image de profondeur et d'intensité en utilisant LPQ multi-échelles.

La Figure 4.3 illustre un exemple d'extraction des caractéristiques en appliquant le descripteur LPQ multi-échelles sur une image (profondeur et intensité) de la base de données FRGC v2.0.

4.2.2.2 Descripteur BSIF

BSIF (Binarized Statistical Image Features) est un nouvel descripteur local proposée par Kannala and Rahtu [67] pour l'utiliser dans des tâches de reconnaissance de texture de

manière similaire de descripteurs LBP et LPQ. Le BSIF calcule une chaîne de code binaire pour les pixels d'une image donnée. La valeur de code d'un pixel est considérée comme un descripteur local de l'image motif d'intensité dans l'environnement du pixel. Davantage, les histogrammes des valeurs de code des pixels permettent de caractériser les propriétés de textures dans les sous-régions d'image. La valeur de chaque élément (c'est-à-dire bit) est calculée en binarisant la réponse d'un filtre linéaire avec un seuil à zéro. Chaque bit est associé à un filtre différent et à la longueur souhaitée de la chaîne de bits détermine le nombre de filtres utilisés. Plus de détail est décrit ci-dessous.

Détails. Étant donné un patch d'image X de taille $l \times l$ pixels et un filtre linéaire W_i de même taille, la réponse du filtre s_i est obtenu par :

$$s_i = \sum_{u,v} W_i(u,v)X(u,v) = \mathbf{w}_i^T \mathbf{x} \quad (4.5)$$

où \mathbf{w}_i et \mathbf{x} sont deux vecteurs contenant les pixels de W_i et X respectivement.

Étant donné n filtres linéaires W_i , on peut les empiler en une matrice \mathbf{W} de taille $n \times l^2$ et calculer toutes les réponses à la fois, c'est-à-dire $\mathbf{s} = \mathbf{W}\mathbf{x}$ et on obtient la chaîne de bits b en binarisant chaque élément s_i de \mathbf{s} comme ci-dessus.

$$b_i = \begin{cases} 1 & \text{si } s_i > 0 \\ 0 & \text{autrement} \end{cases} \quad (4.6)$$

Afin d'obtenir un ensemble utile de filtres W_i , la même idée de [201] est appliquée pour estimer les filtres en maximisant l'indépendance statistique de s_i . Cependant, afin d'utiliser les algorithmes standard d'analyse en composantes indépendantes (ICA) pour estimer les composantes indépendantes, il faut décomposer le filtre matrice \mathbf{W} en deux parties par

$$\mathbf{s} = \mathbf{W}\mathbf{x} = \mathbf{U}\mathbf{V}\mathbf{x} = \mathbf{U}\mathbf{z} \quad (4.7)$$

Où $\mathbf{z} = \mathbf{V}\mathbf{x}$, et \mathbf{U} est une matrice carrée $n \times n$ qui sera estimée via l'algorithme ICA, et la matrice \mathbf{V} effectue le prétraitement canonique, c'est-à-dire blanchiment simultané et réduction de la dimension des échantillons d'apprentissage \mathbf{x} [201]. Le prétraitement canonique utilise la principale analyse des composants PCA pour réduire la dimension de \mathbf{x} et en ne conservant que les n premiers composants principaux qui sont ensuite divisés par leur écart type pour obtenir des échantillons de données blanchis \mathbf{z} . En détail, étant donné la composition propre $\mathbf{C} = \mathbf{E}\mathbf{D}\mathbf{E}^T$ de la matrice de covariance \mathbf{C} des échantillons \mathbf{x} , la matrice \mathbf{V} est définie par :

$$\mathbf{V} = (\mathbf{D}^{-1/2}\mathbf{E}^T)_{1:n} \quad (4.8)$$

où la diagonale principale de \mathbf{D} contient les valeurs propres de \mathbf{C} par ordre décroissant, et $(\cdot)_{1:n}$ désigne le premier n lignes de la matrice entre parenthèses.

Ensuite, compte tenu de la valeur moyenne nulle des échantillons de données blanchies \mathbf{z} , on peut utiliser des algorithmes d'analyse en composantes indépendantes standards ICA pour estimer une matrice orthogonale \mathbf{U} avec laquelle on produit les composantes indépendantes \mathbf{s} des données d'apprentissage [202]. En d'autres termes, étant donné que $\mathbf{z} = \mathbf{U}^{-1}\mathbf{s}$, les composantes indépendantes permettent de représenter les échantillons \mathbf{z} comme une superposition linéaire de vecteurs de base définis par les colonnes de \mathbf{U}^{-1} de données. Enfin, compte tenu de \mathbf{U} et \mathbf{V} , on obtient la matrice de filtre $\mathbf{W}=\mathbf{UV}$, qui peut être directement utilisée pour le calcul de BSIF. La **Figure 4.4** illustre un exemple d'extraction de caractéristiques par le descripteur BSIF à différente taille de filtres pour une image d'intensité et profondeur.

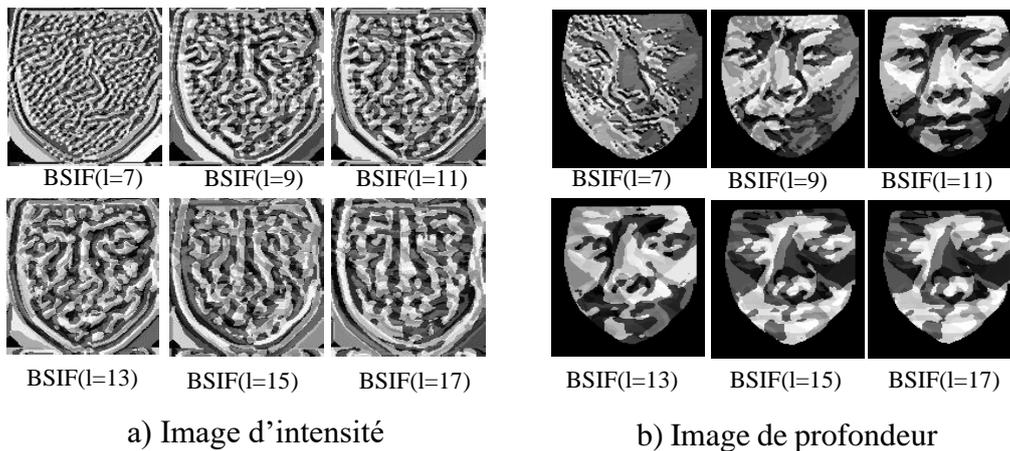


Figure 4.4 Extraction des caractéristiques par le descripteur BSIF.

4.2.3 Réduction de la dimensionalité, comparaison et décision

Les caractéristiques des images faciale 2D (texture) et 3D (profondeur) extraites par LPQ et BSIF sont organisées dans un tenseur $\mathbf{Y} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ d'ordre 3. La formation de ce tenseur est planifiée à partir trois modes comme suite :

- Mode I_1 : représente l'extraction de traits d'images faciales 2D et 3D à différentes échelles.
- Mode I_2 : représente les histogrammes concaténés de différents descripteurs locaux.
- Mode I_3 : représente les différentes personnes issues de différentes bases de données.

Durant la phase d'apprentissage, le tenseur de données $\mathbf{Y} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ est projeté de l'espace d'origine en un sous-espace inférieur. D'abord, MWPCA est appliquée pour capturer la plupart

de variations existant dans les objets tensoriels d'origine en déterminant les matrices projections multilinéaires dans chaque mode. Par conséquent, la dimension de données est réduite seulement dans le deux modes I_1 et I_2 , tandis que le mode I_3 est ne plus réduit car elle représente les différentes personnes des bases de données. Ensuite, notre approche MDA+WCCN est appliquée pour la réduction de dimension et la classification où les données de tenseur d'apprentissage sont projetées dans un nouveau sous espace plus faible. Par conséquent, un nouveau tenseur réduit $\mathbf{Z} \in \mathbb{R}^{d_1 \times d_2 \times I_3}$ est obtenu, où $d_1 \times d_2 \ll L_1 \times L_2$.

Dans la phase de test, l'image du visage du candidat $Y_i \in \mathbb{R}^{I_1 \times I_2}$ passe par les mêmes étapes lors de la phase d'apprentissage. Après des projections successives dans le sous-espace inférieur, les caractéristiques obtenues $Z_i \in \mathbb{R}^{d_1 \times d_2}$ ($d_1 \times d_2 \ll I_1 \times I_2$) sont comparées à l'ensemble du tenseur \mathbf{Z} des échantillons d'apprentissage en utilisant la similarité de cosinus. La décision finale est obtenue par la fusion au niveau du score en utilisant la règle de somme pondérée.

4.3 Résultats expérimentaux

Dans cette section, l'évaluation de l'approche proposée est effectuée sur trois grandes bases de données FRGC v2.0 [203], Bosphorus [204] et CASIA 3D [205]. D'abord, nous représentons ces bases de données ainsi que ses protocoles d'apprentissage et de test. Ensuite, nous donnons la configuration de paramètres de notre approche. Après, nous évaluons l'approche proposée pour la vérification de visage 2D, 3D et 2D+3D. finalement, nos résultats seront commentés et comparés avec ceux de l'état de l'art.

4.3.1 Les bases de données

4.3.1.1 FRGC v2.0

La base de données FRGC v2.0 est l'une des plus grands ensembles de données disponibles [203]. Six différentes expériences standards sont définies avec FRGC v2.0 pour évaluer les systèmes de reconnaissance faciale dans différentes conditions utilisant des visages 2D ou 3D ainsi que leur combinaison. Dans notre travail, nous concentrons sur l'expérience 3 destinée à évaluer les systèmes de reconnaissance faciale multimodale. L'ensemble de données d'expérience 3 de FRGC v2.0 subdivise en deux partitions principales apprentissage et validation. L'ensemble de données d'apprentissage contient 943 visages 2D/3D de 266 sujets tandis que l'ensemble de validation contient 4007 visages 2D/3D de 466 sujets de différentes races (68% blancs, 22% asiatiques et 10% autres), différents sexes (57% d'hommes et 43% de femmes) et d'âges différents. Le protocole de test définit également trois courbes de

caractéristiques de fonctionnement ROC 1, ROC 2 et ROC 3 correspondants aux images acquises au cours des semestres, dans l'année et entre les semestres respectivement. La **Figure 4.5** représente un ensemble d'images de la base de données FRGCv2.0.



Figure 4.5 Ensemble d'images faciales de la base de données FRGC v2.0.

4.3.1.2 Bosphorus

La base de données Bosphorus [204] est destinée à la recherche sur le visage humain 2D et 3D. Les données faciales sont acquises à l'aide d'un scanner 3D à lumière structurée. Cette base de données comprend 4666 scans de visage de 105 sujets recueillis dans diverses conditions de pose, d'expression et d'occlusion. Nous avons utilisé un protocole d'évaluation comprenant 1097 scans de 105 sujets pour la séparation des personnes en deux phases, apprentissage et test. La phase d'apprentissage contient 58 sujets avec un total de 580 images de visage, cependant, la phase de test comprend 470 images de 47 sujets, 10 images par personne. La plupart des sujets sont des Caucasiens âgés de 25 à 35 ans. Des échantillons d'images faciales de cette base de données sont illustrés par la **Figure 4.6**.

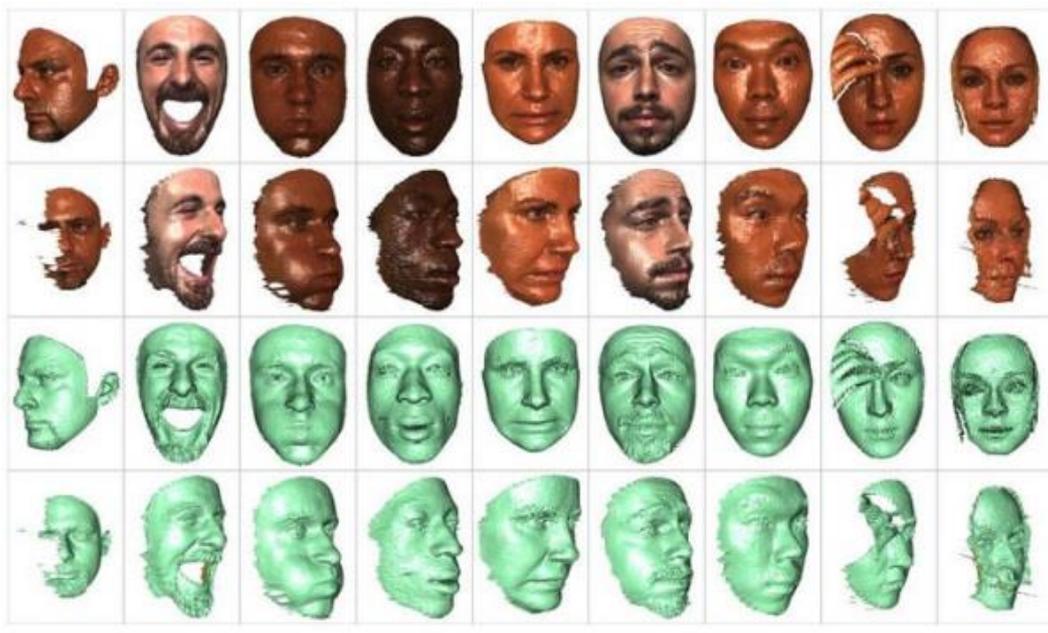


Figure 4.6 Ensemble des échantillons d'images faciales extraites de base de données Bosphorus [206].

4.3.1.3 CASIA 3D

La base de données CASIA 3D [33] (Chinese Academy of Sciences, Institute of Automation) a été collectée entre août et septembre 2004 à l'aide du scanner Minolta Vivid 910. La base de données de visages 3D comprend 4624 scans de 123 personnes, chaque personne a 37 ou 38 images avec différentes expressions et variations de pose. Les 123 sujets sont répartis selon le protocole de test en clients et imposteurs : 100 clients et 13 imposteurs pour l'évaluation et 10 imposteurs pour le test. La **Figure 4.7** représente des exemples des images faciales sont extraite de base de données CASIA 3D.



Figure 4.7 Exemples des images faciales extraites de base de données CASIA 3D.

4.4 Paramètres de configuration

Notre système de vérification de visage 2D et 3D passe par plusieurs étapes. Tout d'abord, les images faciales 2D et 3D sont détectées et recadrées, puis les visages 2D sont

converties en niveaux de gris. Les images faciales 2D et 3D sont redimensionnées à 150×130 pixels. Après cela, chaque image faciale est subdivisée en 30 blocs de taille 25×26 pixels. Dans chaque bloc, les caractéristiques locales sont extraites par les descripteurs LPQ et BSIF à multi-échelles. Les histogrammes de tous ces blocs sont concaténés en un vecteur de taille 30×256 . Ensuite, l'algorithme MDA + WCCN est appliqué afin de réduire les vecteurs de caractéristiques multidimensionnelles. Ce processus est terminé lorsque le nombre d'itérations est atteint. Le nombre maximum d'itérations est ajusté empiriquement à $itr_{max} = 5$ correspondants au sous-espace final de dimension inférieure. Pour déterminer la performance de notre approche de vérification de visages unimodaux 2D/3D et multimodaux 2D+3D, nous rapportons les critères principaux suivants : le taux de vérification TV à 0.001 TFA (taux de fusse acceptation), le taux d'égale erreur (TEE) et les courbes ROC (Receiver Operating Characteristic).

L'algorithme de notre système de vérification de visage est implémenté à l'aide d'un logiciel de Matlab 2016a. Par conséquent, le programme est exécuté sur un ordinateur ASUS R540L a un processeur Intel i5 2.7Ghz avec une carte graphique NVIDIA GEFORCE 920M et RAM de 8Go.

4.5 Discussion et Comparaison

Notre système de reconnaissance biométrique par la fusion multimodale de visage a été performé sur trois bases de données FRGC v2.0, Bosphorus et CASIA 3D. Les résultats expérimentaux nous conduisent à discuter les performances suivantes :

- 1) La performance des descripteurs locaux LPQ et BSIF individuellement ainsi que leur fusion.
- 2) La performance de l'approche proposée par rapport à la méthode MDA.
- 3) Les performances de la modalité 2D et 3D.
- 4) Les performances de la vérification multimodale de visage 2D + 3D.
- 5) Enfin, l'approche proposée est comparée avec l'état de l'art.

4.5.1 Descripteurs locaux LPQ et BSIF

Nous avons évalué les performances des descripteurs locaux LPQ et BSIF, où le meilleur descripteur donne un taux de vérification TV plus élevé et un taux d'égale erreur TEE plus petit. A travers les résultats résumés dans le tableau 4.1 qui concerne la base de données FRGC V2.0, le descripteur local BSIF donne une petite plage de supériorité par rapport au LPQ

avec une différence de 0,67%, 054% et 0,27% de taux de vérification pour 2D, 3D et 2D + 3D respectivement. Alors que les résultats expérimentaux sur les bases Bosphorus et CASIA 3D rapportés dans les tableaux 4.3 et 4.4, démontrent que le descripteur LPQ est plus efficace que le descripteur BSIF à l'extraction de caractéristiques. Dans le tableau 4.3, LPQ a surmonté BSIF avec différence de taux de vérification égale à 0,47%, 1,07% et 0,6% pour les images faciales 2D, 3D et 2D + 3D respectivement. En addition, la supériorité de LPQ est facilement notée sur la base de données Bosphorus où la différence de taux de vérification égale à 4,89%, 5,96 et 5,32% pour les image faciales 2D, 3D et 2D + 3D respectivement. Par conséquent, les résultats expérimentaux rapportées dans le tableau 4.1 montrent que la fusion de deux descripteurs LPQ et BSIF peut apporter des petites améliorations plus de 0,32% que d'utiliser chaque descripteur individuellement.

Nous pouvons conclure que la fusion de deux descripteurs tel que LPQ et BSIF ne conduit pas nécessairement à meilleurs résultats que leur utilisation individuellement. Cette conclusion est prouvée dans les tableaux 4.3 et 4.4 où nous avons utilisé un seul descripteur LPQ qui a donné un résultat plus satisfaisant que la fusion de deux descripteurs LPQ + BSIF. Ceci est dû au fait que chaque descripteur a une méthodologie d'extraction des caractéristiques différente, sans oublier la différence entre les bases de données (conditions de collection d'images, protocoles, divers dispositifs de scan, ... etc.) qui peut limiter l'efficacité du descripteur.

Tableau 4.1 Taux de vérification 2D, 3D et 2D + 3D à 0,001 FAR (%) sur la base de données FRGC v2.0.

Réduction de la dimensionnalité du tenseur	Descripteur	2D			3D			2D+3D		
		ROC1	ROC2	ROC3	ROC1	ROC2	ROC3	ROC1	ROC2	ROC3
MDA	LPQ	87.81	77.78	67.88	92.76	91.64	90.05	97.10	95.05	93.03
	BSIF	88.45	78.55	68.55	92.84	91.85	90.74	97.37	95.54	93.74
	LPQ+BSIF	87.99	77.49	67.17	92.84	92.01	90.97	97.12	95.24	93.25
MDA+WCCN	LPQ	88.34	79.45	70.19	95.54	94.60	93.58	98.04	96.83	95.81
	BSIF	89.01	79.59	70.21	96.08	95.44	94.70	98.28	97.27	96.25
	LPQ+BSIF	90.50	81.90	73.23	96.70	95.90	95.02	98.60	97.51	96.60

Tableau 4.2 Taux d'erreur égal (%) 2D, 3D et 2D + 3D sur la base de données FRGC v2.0.

Réduction de la dimensionnalité du tenseur	Descripteur	2D			3D			2D+3D		
		ROC1	ROC2	ROC3	ROC1	ROC2	ROC3	ROC1	ROC2	ROC3
MDA	LPQ	3.12	6.14	8.30	1.6	1.71	1.87	0.81	1.25	1.55
	BSIF	2.77	5.82	7.87	1.56	1.68	1.84	0.77	1.18	1.43
	LPQ+BSIF	3.04	6.36	8.63	1.52	1.59	1.68	0.80	1.17	1.41
MDA+WCCN	LPQ	3.14	5.50	7.16	1.10	1.21	1.31	0.62	0.89	1.02
	BSIF	2.82	5.49	7.41	0.98	1.09	1.23	0.57	0.82	1.03
	LPQ+BSIF	2.58	4.98	6.59	0.84	0.96	1.03	0.51	0.78	0.96

4.5.2 L'approche proposée par rapport à la méthode MDA

Les tableaux 4.1, 4.3 et 4.4 représentent les résultats expérimentaux obtenus lorsque nous avons utilisé MDA seule et MDA combinée avec WCCN. Les résultats obtenus montrent clairement un avantage significatif de notre méthode par rapport au MDA. D'après les résultats des tableaux 4.1, 4.3 et 4.4, nous pouvons affirmer que la MDA combinée au WCCN a clairement amélioré le taux de vérification du visage d'une pourcentage de 0,54%, 0,94% et 0,54% pour les images faciales 2D, 3D et 2D + 3D respectivement en utilisant la base de données CASIA 3D. Quand nous utilisons la base de données FRGC v2.0, nous pouvons obtenir une amélioration pour ROC1 plus de 2,05%, 3,86% et 1,23% pour les images faciales 2D, 3D et 2D + 3D respectivement. Pour la base de données Bosphorus, l'amélioration est fortement augmentée et peut atteindre respectivement 5,95%, 6,39% et 6,39% pour les images faciales 2D, 3D et 2D + 3D. De plus, les résultats présentés dans le tableau 4.2 montrent que notre approche obtient un taux égal d'erreur moins que la MDA. En général, l'approche proposée fonctionne mieux que la MDA dans les trois bases de données populaires FRGC v2.0, Bosphorus et CASIA 3D. La **Figure 4.8** illustre les courbes ROC1 de la base de données FRGC v2.0 lorsqu'on applique MDA et MDA + WCCN pour les images faciales 2D, 3D et 2D + 3D.

Tableau 4.3 Taux de vérification 2D, 3D et 2D + 3D à 0,001 FAR (%) sur la base de données CASIA 3D.

Réduction de la dimensionnalité du tenseur	Descripteur	2D	3D	2D+3D
MDA	LPQ	98.56	98.36	98.98
	BSIF	97.78	97.71	97.80
	LPQ+BSIF	98.35	97.28	97.76
MDA+WCCN	LPQ	99.10	99.30	99.52
	BSIF	98.63	98.23	98.92
	LPQ+BSIF	98.62	98.68	98.90

Tableau 4.4 Taux de vérification 2D, 3D et 2D + 3D à 0,001 FAR (%) sur la base de données Bosphorus.

Réduction de la dimensionnalité du tenseur	Descripteur	2D		3D		2D+3D	
		ERR	VR	ERR	VR	ERR	VR
		(%)	(%)	(%)	(%)	(%)	(%)
MDA	LPQ	2.55	87.02	2.08	89.14	2.12	90.63
	BSIF	2.45	82.97	3.98	82.76	2.55	86.80
	LPQ+BSIF	2.98	82.34	3.56	84.89	2.55	85.53
MDA+WCCN	LPQ	1.49	92.97	1.07	95.53	1.2	97.02
	BSIF	1.89	88.08	2.59	89.57	2.07	91.70
	LPQ +BSIF	1.74	90.85	1.91	92.55	1.46	94.89

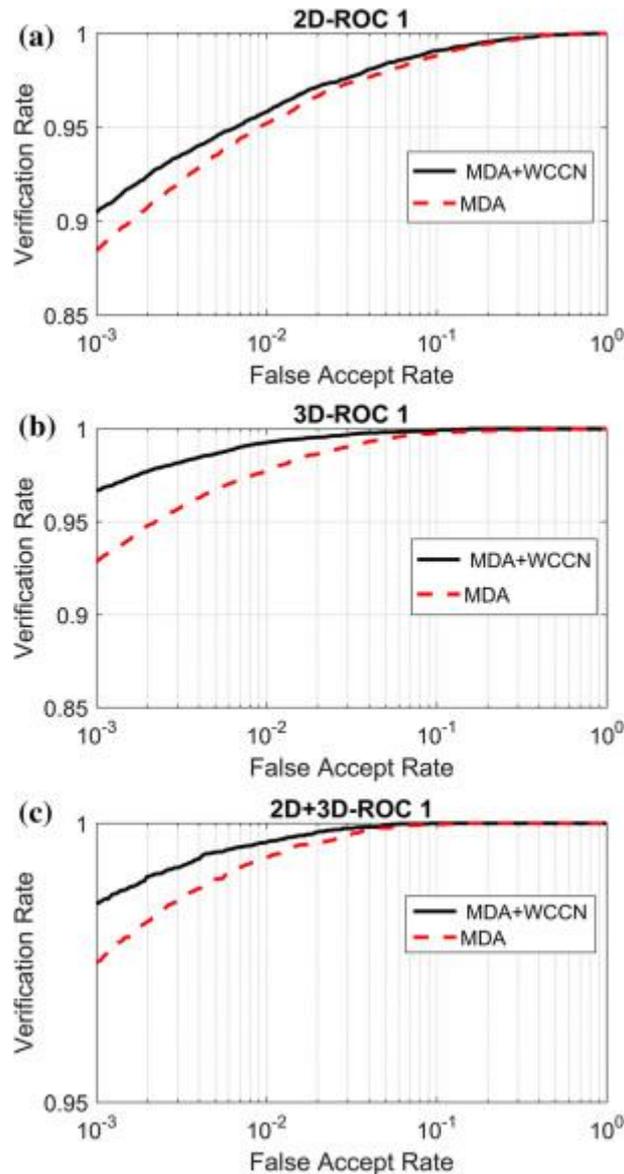


Figure 4.8 Courbes ROC1 de MDA et MDA + WCCN pour les images faciales 2D, 3D et 2D + 3D sur bases de données FRGC v2.0.

4.5.3 Les performances de la modalité 2D et 3D

Les résultats rapportés dans tableaux 4.1, 4.3 et 4.4 montrent que l'utilisation des scans 3D est plus performante que celle des scans 2D. De plus, les tableaux 4.2 et 4.4 montrent que le taux d'erreur égal ERR résultant des scans 3D est inférieur à celui résultant des scans 2D. Le meilleur taux de vérification obtenu par l'approche proposée en utilisant des scans 3D de visage sur les trois bases de données est 96,70%, 95,53% et 99,30% pour FRGC v2.0, Bosphorus et CASIA 3D respectivement. Cela est dû au pouvoir discriminant des images faciales en 3D contre des défis tels que les expressions, les illuminations et les variations de pose, contrairement aux images de visage 2D qui sont sensibles à ces défis.

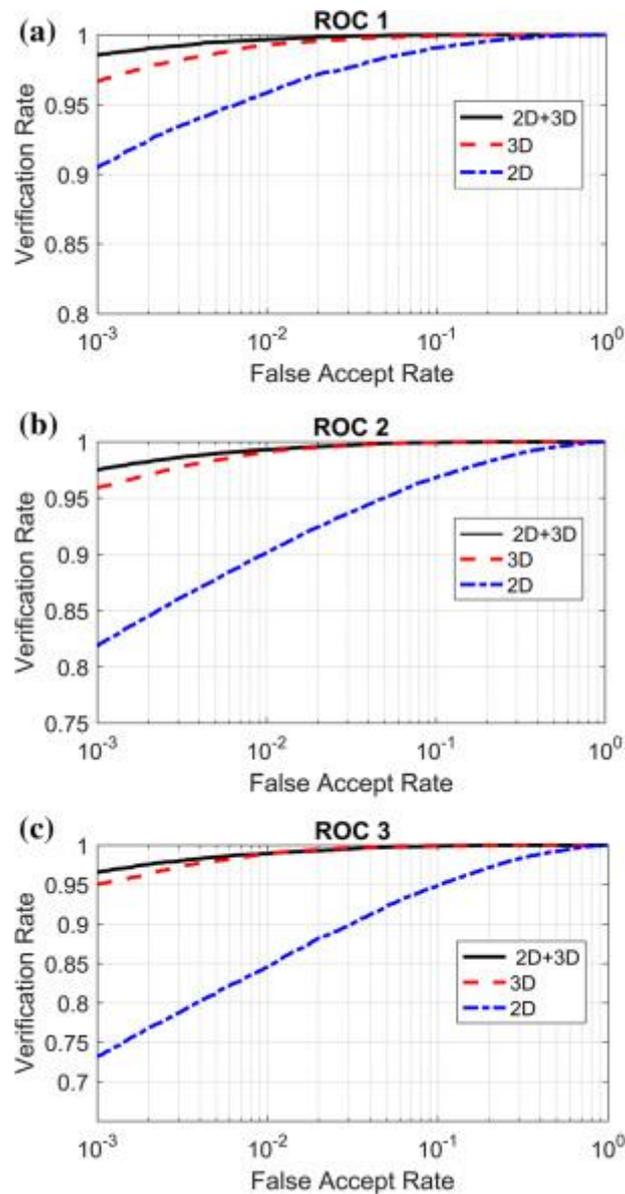


Figure 4.9 Courbes ROC de la vérification de visage 2D, 3D et 2D + 3D sur la base de données FRGC v2.0.

4.5.4 Les performances de la vérification multimodale de visage 2D + 3D

Les résultats fournis dans les tableaux 4.1, 4.3 et 4.4 montrent que le système de vérification de visage basé sur des images de visage 2D + 3D fonctionne mieux qu'un système qui utilise une image faciale 2D ou 3D individuellement. Cette conclusion est plus évidente sur toutes les bases de données où le meilleur taux de vérification est obtenu par fusion d'images faciales 2D + 3D est de 98,60%, 97,02% et 99,52% sur les bases de données FRGC v2.0,

Bosphorus et CASIA 3D respectivement. Par conséquent, les résultats obtenus prouvent clairement que 2D et 3D sont complémentaires et que leur fusion basée sur une transformation multilinéaire surpasse leur fusion en utilisant des méthodes classiques basées sur la représentation matricielle ou vectorielle. La **Figure 4.9** représente les courbes ROC de la vérification de visage 2D, 3D et 2D + 3D sur la base de données FRGC v2.0.

4.5.5 L'approche proposée VS l'état de l'art

Nous avons évalué les performances de l'approche proposée pour la vérification du visage en utilisant trois grandes bases de données FRGC v2.0, CASIA 3D et Bosphorus. Les tableaux 4.5, 4.6 et 4.7 présentent les meilleurs taux de vérification de notre approche et des travaux récents utilisant des images de visage 3D et 2D + 3D. Les meilleurs taux de vérification obtenus par notre approche en utilisant la base de données FRGC V2.0 sont de 96,70% pour le visage 3D et de 98,60% pour le 2D + 3D. Avec la base de données CASIA 3D, les meilleurs taux de vérification sont de 99,30% en utilisant uniquement les images faciales 3D et de 99,52% en utilisant à la fois la fusion des images faciales 2D + 3D. Pour la base de données Bosphorus, nous avons atteint des taux de vérification de 95,53% pour la modalité de visage 3D et de 97,02% pour la multimodalité des images 2D + 3D. Notre approche est comparée à des travaux récents dans l'état de l'art qui sont mentionnés dans les tableaux 4.5, 4.6 et 4.7. Dans un premier temps, nous comparons l'approche proposée avec des méthodes basées sur la réduction de dimensionnalité des tenseurs. A titre d'exemple, les meilleurs taux de vérification des visages 3D obtenus dans [207] et [198] rapportés dans le tableau 4.5 sont de 95,94% et 96,98% alors que notre approche atteint 96,70% en utilisant la même base de données FRGCv2.0. Pour les visages 2D + 3D, les résultats ne montrent pas de différence notable entre les trois méthodes. Les résultats décrits dans les tableaux 4.6 et 4.7 prouvent que notre méthode est plus performante que celles de l'état de l'art [207] and [198] aussi bien en 3D qu'en 2D + 3D en utilisant les bases de données CASIA 3D et Bosphorus; sauf le résultat obtenu dans [207] pour la base de données Bosphorus (2D + 3D) qui est similaire à notre résultat 97,02%.

Lorsque nous comparons notre approche aux résultats de [208], qui est un réseau de neurones convolutif basé sur un système de reconnaissance faciale 2D et 3D. Nous montrons clairement une grande différence de plus de 11,55% pour la vérification faciale entre l'approche proposée et [208]. Ces résultats sont rapportés dans le tableau 4.5. La comparaison avec [209] met l'accent sur l'efficacité de l'approche proposée, où nos expériences réalisées sur CASIA 3D dépassent les résultats obtenus en [209] avec une différence de taux égale à 13,37% pour le visage 3D et

Tableau 4.5 Taux de vérification (%) de l'approche proposée et de l'état de l'art sur la base de données FRGC v2.0.

Auteurs	Méthode	Modalité	ROC1	ROC2	ROC3
Ouamane [210]	EDA+WCCN	3D	94.17	93.15	92.06
L.Spreewers [211]	PCA-LDA		94.6	94.6	94.6
Al-Osaimi [114]	PCA		94.55	94.12	94.05
Drira [135]	Elastic shape analysis		/	/	97.14
Hu [208]	CNN2		85.15	/	/
Ouamane [198]	TEDA+WCCN		96.98	95.93	95.10
Bessaoudi [207]	MEFDA		95.94	95.25	94.40
Notre méthode [192]	MDA+WCCN		96.70	95.90	95.02
Ouamane [210]	EDA+WCCN	2D+3D	98.28	98.22	98.12
Wang [212]	CSDC		97.97	98.01	98.04
Kakadiaris [108]	Haar+Pyramid Transforms		97.3	97.2	97.00
Ouamane [198]	TEDA+WCCN		98.95	98.50	98.44
Bessaoudi [207]	MEFDA		99.00	98.34	97.81
Notre méthode [192]	MDA+WCCN		98.60	97.60	96.70

1,08% pour 2D + 3D. Les résultats présentés dans le tableau 4.7 rendent la comparaison plus compétitive entre l'approche proposée et [213]. Les résultats dans [213] montrent un petit avantage par rapport à nos résultats pour le visage 3D en utilisant la base de données Bosphorus. Pour clarifier davantage la vision, notre approche nécessite qu'une base de données suffisante pour l'apprentissage, tandis que les systèmes de reconnaissance faciale d'apprentissage en profondeur basés sur « CNN » et « deep face » ont besoin pour l'apprentissage plusieurs base de données. De plus, notre approche ne dépend d'aucun autre système tel que MobileNetV2 qui est utilisé dans [213]. Néanmoins, l'approche proposée a obtenu d'excellents résultats.

Tableau 4.6 Taux de vérification (%) de l'approche proposée et l'état de l'art sur la base de données CASIA 3D.

Auteurs	Méthode	3D	2D+3D
Chouchane [214]	PCA+EFM	98.22	/
Ouamane [210]	EDA+WCCN	98.21	99.12
Bessaoudi [207]	MEFDA	99.06	99.36
Guo [215]	LGP-WHGO	/	98.61
Feng [209]	DCNN	85.93	98.44
Ouamane [198]	TEDA+WCCN	98.97	99.47
Chouchane [216]	HOSVD+EFM	98.60	/
Ming [217]	ROSR	96.25	/
Notre méthode [192]	MDA+WCCN	99.30	99.52

Tableau 4.7 Taux de vérification (%) de l'approche proposée et l'état de l'art sur la base de données Bosphorus.

Auteurs	Méthode	3D	2D+3D
Ming [217]	ROSR	93.95	-
Guo [215]	LGP-WHGO	-	96.75
Bessaoudi [207]	MEFDA	95.31	97.02
Olivetti [213]	MobileNet V2 NN	97.56	-
Ouamane [198]	TEDA+WCCN	92.97	96.17
Bessaoudi [218]	MPCA+EDA	92.12	-
Elaiwat [219]	Curvelet local features	77.50	91
Notre méthode [192]	MDA+WCCN	95.53	97.02

4.6 Conclusion

Dans ce chapitre, Nous avons présenté l'approche proposée en détail. Cette approche concerne la fusion multimodale de visage 2D + 3D pour l'identification des personnes en biométrie basée sur la représentation des tenseurs. Notre système comprend quatre étapes principales qui sont : le pré-traitement, l'extraction des caractéristiques, la réduction de la dimensionnalité et la comparaison. Dans la première étape, les visage 2D et 3D sont passés par une phase de prétraitement pour supprimer les effets de bruits et améliorer la qualité de données 2D et 3D. Après cette étape, chaque image faciale 2D ou 3D est subdivisée en des petits blocks en utilisant deux descripteurs locaux LPQ et BSIF. Les caractéristiques locales de chaque block sont enregistrées dans un histogramme. Ce dernier avec les histogrammes des autres blocks sont réunis et concaténés pour former le vecteur de caractéristiques. Dans la troisième étape, la réduction de dimension du vecteur de caractéristiques et la classification des données est effectuée par l'approche proposée MDA associée à WCCN. Dans la dernière étape, les scores sont calculés et comparés pour apprendre la décision finale. Les résultats obtenus sont évalués sur trois populaires bases de données qui sont FRGC v2.0, CASIA 3D et Bosphorus. L'approche proposée a donnée des performances satisfaisantes par rapport l'état de l'art.

Chapitre 5

Conclusion générale et perspectives

5.1 Conclusion générale.....	94
5.2 Perspectives.....	99

Chapitre 5 : Conclusion générale et perspectives

5.1 Conclusion générale

La reconnaissance faciale est une technologie de plus en plus répandue, basée sur la vision par ordinateur et l'intelligence artificielle, permettant d'identifier une personne sur une photo ou une vidéo en comparant son visage avec ceux sauvegardés dans une base de données. De nombreuses applications sont proposées dans la recherche et le marché industriel. La plupart de ces applications rencontrent le problème de la grande dimensionnalité des données multidimensionnelles. Les systèmes de la reconnaissance faciale qui se basent sur leur fonctionnement sur les algorithmes traditionnels d'apprentissage de sous espace tels que PCA, LDA donnent des performances insatisfaisantes car ces méthodes de réduction de la dimensionnalité projettent les données multidimensionnelles en matrice ou un vecteur à grande dimension. Par conséquent, la structure originale des données peut être changée et difficile à récupérer grâce à la linéarité de ces méthodes. Néanmoins, elles ont besoin d'un temps de calcul plus long et un grand espace de stockage. Les méthodes multilinéaires tels que MPCA et MDA représentent une solution prometteuse pour les systèmes de reconnaissance faciale. La multilinéarité permet d'appliquer ces méthodes directement sur les données dans leur forme tensorielle sans les transformer en matrice ou vecteur. À travers cette vision, nous avons proposé notre système de reconnaissance biométrique par la fusion multimodale de visage. Nous avons consacré à proposer ou développer des nouvelles solutions uni-modales et multimodales pour identifier les personnes en biométrie. Notre travail se réalise à travers deux côtés. Dans le premier côté, nous avons tenté de trouver des solutions au niveau de caractéristiques. Par conséquent, nous avons utilisé deux descripteurs LPQ et BSIF qui sont très efficaces à extraire les caractéristiques locales de visages 2D et 3D. Les caractéristiques locales sont riches par les informations utiles par rapport à les caractéristiques globales. Afin de trouver des résultats satisfaisants, nous avons fusionné les caractéristiques locales de deux descripteurs dans les deux cas unimodaux 2D, 3D et multimodale 2D+3D. Dans le deuxième côté, nous avons proposé d'utiliser deux approches multilinéaires pour la réduction de dimension. D'abord, la méthode MWPCA est appliquée pour projeter le vecteur de caractéristiques 2D et 3D dans sa forme tensorielle d'un espace de grande dimension à un sous espace de dimension inférieure. Cette projection est effectuée selon trois modes (caractéristiques, histogrammes et

personnes). Pour la classification, nous avons proposé d'associer la méthode MDA à la métrique WCCN car cette dernière améliore la performance de MDA à travers la minimisation de taux d'erreurs de classification attendu sur les échantillons d'apprentissage en utilisant l'inverse de la matrice de covariance intra-classe.

Dans ce travail, nous avons présenté quatre contributions principales :

- Premièrement, nous avons proposé une étude et un concept d'un nouveau système de vérification automatique de visage uni-modale 2D, 3D et multimodale 2D+3D basé sur la fusion multimodale des caractéristiques locales.
- Deuxièmement, nous avons proposé d'utiliser l'analyse multidimensionnelles comme une solution prometteuse pour le problème de la grande dimensionnalité de données. Pour ce but, deux approches d'analyses multilinéaires sont proposées qui sont MWPCA MDA
- Troisièmement, nous avons proposé d'améliorer la performance de MDA en minimisant l'erreurs de classification. Donc, l'approche MDA est associée avec WCCN pour assurer une meilleure classification de données.
- Quatrièmement, nous avons étudié les performances de descripteurs locaux LPQ et BSIF sur les visages 2D et 3D d'une manière individuelle, ainsi que dans leur fusion.

Nous avons évalué notre approche par la comparaison avec les approches récentes de l'état de l'art sur les bases de données FRGC v2.0, Bosphorus et CASIA 3D.

La conception d'un système de reconnaissance faciale automatique robuste et efficace demande beaucoup des efforts d'un côté de matériel et le côté d'algorithmes. Le but principal de cette thèse concentre sur le côté algorithmique sur un thème « Fusion Multimodale pour l'identification des personnes en Biométrie ». Notre thèse est articulée sur cinq chapitres principales.

Dans **le premier chapitre**, nous avons présenté le contexte et les motivations du sujet de thèse, ainsi que les objectifs et les contributions. D'abord, nous avons montré la place de la biométrie de visage parmi les autres biométries existantes. Ensuite, nous avons introduit les notions de base sur les systèmes biométriques en général et les outils permettant leur évaluation.

Dans **le deuxième chapitre**, nous avons présenté brièvement les différentes techniques proposées dans la reconnaissance faciale unimodale 2D, 3D et multimodale 2D+3D. La biométrie faciale 2D a été reconnue avec sa succès dans nombreuses techniques de

reconnaissance faciale. Ces techniques peuvent être généralement classées en trois catégories : les méthodes globales ou holistiques, les méthodes locales et les méthodes hybrides. Chacune de ces méthodes a des avantages et des inconvénients. Comme indique son nom, les méthodes globales permettent d'analyser que les caractéristiques globales de visage. Ces méthodes telles que PCA, LDA, ICA représentent les images faciales comme des matrices ou des vecteurs plus faciles à manipuler. Les approches globales ou les approches de sous espace sont rapides à mettre en œuvre mais généralement elles sont généralement sensibles aux conditions d'éclairage, de pose et d'expressions faciale. Tandis que, les approches locales ne traitent que certaines zones du visage (le nez, la bouche, les yeux, ...etc). Nous remarquons deux types pour les approches locales, chacune basées sur l'apparence locale du visage comme LBP, LPQ, BSIF, HOG, ... et d'autres basées sur les points d'intérêt dans l'image du visage telles que BRIEF, SURF et SIFT, ...etc. Elles sont généralement plus sensibles aux expressions faciales, aux occlusions et à la pose. Il existe aussi les approches hybrides qui combinent les caractéristiques globales et locales d'images faciales pour améliorer la performance et la précision de la reconnaissance.

L'éclairage et les variations de pose sont deux problèmes majeurs non résolus de reconnaissance faciale 2D. Récemment, la reconnaissance faciale 3D a été largement étudié par la communauté scientifique pour surmonter ces problèmes non résolus de la reconnaissance faciale 2D et pour obtenir une précision nettement plus élevée en mesurant la géométrie des éléments rigides sur le visage. Plusieurs systèmes récents basés sur des données 3D ont été développés. Comme la reconnaissance faciale 2D, les approches de la reconnaissance faciale 3D peuvent être classer en deux catégories : globales et locales. Les informations géométriques 3D correspondant aux approches globales travaillent directement sur toute l'image de profondeur ou nuage de points du visage pour calculer la mesure de similarité. Les approches globale 3D peuvent être réduits à trois méthodes principales : les méthodes de sous espace basées sur PCA et LDA, les méthodes basées sur l'alignement de surfaces à l'aide de l'algorithme ICP et méthodes basées sur le modèle morphable. Dans l'autre côté, les approches locale 3D, Les techniques de correspondance basées sur les caractéristiques locales se concentrent sur des points descriptifs locaux, des courbes ou des régions de scans faciaux 3D (les coordonnées originales, les zones, les distances, les angles, les courbures, ...etc). Ces approches ont été largement étudiées dans la reconnaissance faciale 3D, puisqu'elles donnent généralement de meilleurs résultats que l'appariement holistiques. Les approches de reconnaissance faciale 3D existantes peuvent être classifiées selon la fonctionnalité qu'elles

utilisent : (1) méthodes basées sur le calcul des points clés (keypoint), (2) méthodes basées sur le maillage, (3) méthodes basées sur le calcul de courbure et (4) méthodes basées sur la surface. Les méthodes de reconnaissance faciale 3D sont généralement robustes aux variations d'éclairage et de pose. Cependant, elles sont plus sensibles aux variations d'expression faciale.

Il s'agit de techniques qui combinent des données 2D et 3D sur le visage peuvent améliorer les performances et la robustesse de la reconnaissance. Récemment, les approches de la reconnaissance faciale basée sur la fusion multimodale des données 2D et 3D sont plus étudiées. Les résultats obtenus par ces méthodes ont affirmé la complémentarité de l'image faciale de profondeur et de l'intensité.

Dans **le troisième chapitre**, nous avons présenté les propriétés des tenseurs pour la modélisation de données multidimensionnelles, ainsi que les techniques multilinéaires de sous espace explorées dans cette thèse. D'abord, nous avons présenté les notations et les concepts nécessaires pour la définition du tenseur. Puis, nous avons expliqué les opérations de bases d'algèbre multilinéaire. Ensuite, nous avons déterminé les projections multilinéaires possibles, notamment la projection vecteur à vecteur, la projection tenseur à vecteur, et la projection tenseur à tenseur, et leurs relations sous-jacentes. Dans la dernière partie, nous avons détaillé les approches multilinéaires utiliser dans notre travail pour la réduction de dimension MWPCA, MDA et MDA+WCCN.

Dans **le quatrième chapitre**, nous avons résumé le travail de cette thèse. Nous avons donné un aperçu de l'approche proposée à l'aide d'un schéma fonctionnel. Notre système a déroulé en deux phases, vérification et test. Chaque phase comprend de quatre étapes essentielles : prétraitement, extraction des caractéristiques, réduction de la dimensionnalité et la comparaison. Dans la première étape, les visages 2D et 3D sont détectés, recadrés et alignés pour les préparées à l'étape suivante. Dans la deuxième étape, les caractéristiques locales de visages 2D et 3D sont détectées et analysées par les descripteurs locaux LPQ et BSIF. Chaque image faciale a été divisée en 30 petits blocs. Les histogrammes appartiennent de tous les petits blocs sont concaténées pour former le vecteur de caractéristiques. Dans la troisième étape, nous avons proposé d'utiliser les méthodes multilinéaires de sous espace MWPCA pour réduire la dimension du vecteur de caractéristiques et MDA associée à WCCN pour la classification. Dans la dernière étape, les scores de similarité sont calculés et normalisés puis fusionnés. Nous avons présenté les bases de données FRGC v2.0, Bosphorus et CASIA 3D qui sont utilisées pour valider notre approche. En plus, nous avons expliqué les paramètres de configuration

choisis pour atteindre les meilleurs résultats. Enfin, Nos résultats sont comparés aux résultats de méthodes de l'état de l'art et nous avons résumé les commentaires et remarques rédigés ci-dessus :

(1) au niveau d'extraction de caractéristiques :

- La stratégie multi-échelles est toujours meilleure que celle d'une seule échelle pour tous les descripteurs binaires locaux utilisés et pour la vérification uni-modale du visage 2D et 3D et multimodale 2D+3D.
- La fusion de deux descripteurs tel que LPQ et BSIF ne conduit pas nécessairement à meilleurs résultats que leur utilisation individuellement.
- Les résultats obtenus par LPQ sont mieux que ceux obtenus par BSIF presque pour toutes les bases de données utilisées.

(2) au niveau de réduction de la dimensionnalité et classification :

- L'approche proposée MDA associée à WCCN surpasse l'approche MDA dans tous les résultats pour la vérification uni-modale du visage 2D et 3D et multimodale 2D+3D sur les trois bases de données utilisées.

(3) au niveau de modalité 2D et 3D :

- La vérification de visage 3D donne des meilleurs résultats par rapport de la vérification de visage 2D car cette dernière est très sensible à l'éclairage, aux variations de pose et aux expressions faciales.

(4) au niveau de multimodalité 2D+3D :

- La fusion multimodale de visage 2D+3D donne des meilleurs résultats par rapport à l'utilisation unimodale individuelle 2D ou 3D.

(5) au niveau de la comparaison avec l'état de l'art :

- L'approche MDA+WCCN a donné des meilleurs taux de vérification dans les deux cas unimodale et multimodale surtout sur les bases de données Bosphorus et CASIA 3D.

- MDA+WCCN surpasse plusieurs récentes méthodes d'états de l'art, ce qu'il affirme que l'approche proposée est une méthode robuste et efficace pour la reconnaissance faciale.

5.2 Perspectives

La proposition d'un système de la reconnaissance faciale robuste et efficace est l'objectif de notre travail. Par conséquent, ce système a encore besoin de modifications ou d'améliorations. Pour continuer la recherche dans le travail présenté dans cette thèse. Nous proposons à développer la recherche selon les points suivants :

1) pour le visage 3D

- Nous pensons à utiliser les méthodes basées sur le profil (courbes) avec l'algorithme d'apprentissage automatique AdaBoost (machine learning). L'apprentissage automatique basée sur la méthode de sélection des caractéristiques fournit une signature très compacte d'un visage 3D et une approche de classification très rapide pour la reconnaissance faciale. En plus, les méthodes basées sur le profil sont robustes sous les expressions faciales.
- Pour l'extraction de caractéristiques de visage 3D, nous proposons d'utiliser l'opérateur 3DLBP au lieu de l'opérateur LBP original car ce dernier a été proposé pour analyser la texture et ne peut pas traiter le comportement des valeurs de profondeur.
- L'utilisation des techniques de détection les points et régions d'intérêts dans le visage 3D.

2) pour la multimodalité

- La fusion par la somme de l'information locale parfois conduit à diminuer la performance de système de la reconnaissance car beaucoup des informations locales combinées sont similaires et redondantes. Nous proposons d'utiliser à la fusion de l'information locale par algorithme génétique ou PSO (Particle Swarm Optimization) pour la sélection des caractéristiques faciales.
- Proposer un système multimodal de la reconnaissance facial basé sur la représentation tensorielle multi descripteurs et multi classificateur.

3) pour la classification

Comme nous avons vu dans cette thèse, la performance d'un système de reconnaissance faciale est directement liée par la robustesse ou la faiblesse des méthodes de classification et discrimination de données. Par conséquent, nous proposons de :

- L'utilisation des méthodes d'apprentissage de la métrique (Metric Learning) comme LMNN (large margin nearest neighbor) [220], Cross-view Quadratic Discriminant Analysis (XQDA) [221], Null Foley-Sammon Transfer (NFST) [222], keep-it-simple-and-straightforward metric (KISSME) [223] et kernel Local Fisher Discriminant Analysis [224]. Ces méthodes ont dernièrement reçu beaucoup d'attention dans la vision par ordinateur et l'apprentissage de la machine. Nous proposons d'utiliser la fusion de ces méthodes.
- Le traitement de données multidimensionnelles dans leur forme tensorielle évite plusieurs contraintes comme le temps et la complexité de la calcul, le problème de taille d'échantillon (le nombre de pixel dans l'image surpasse le nombre d'échantillons). Il est possible de développer des nouvelles extensions de méthodes d'apprentissage de sous espace ou de méthodes d'apprentissage de la métrique pour augmenter la discrimination de données multidimensionnelles.
- Approfondir la recherche sur l'utilisation des tenseurs d'ordre 4 ou plus dans la reconnaissance faciale.

Bibliographie

1. Jain, A.K., R. Bolle, and S. Pankanti, *Biometrics: personal identification in networked society*. Vol. 479. 2006: Springer Science & Business Media.
2. Jain, A.K., A. Ross, and S. Prabhakar, *An introduction to biometric recognition*. IEEE Transactions on circuits and systems for video technology, 2004. **14**(1): p. 4-20.
3. Heitmeyer, R., *Biometric identification promises fast and secure processing of airline passengers*. ICAO journal., 2000. **55**(9).
4. Phillips, P.J., et al., *The FERET evaluation methodology for face-recognition algorithms*. IEEE Transactions on pattern analysis and machine intelligence, 2000. **22**(10): p. 1090-1104.
5. Kolda, T.G. and B.W. Bader, *Tensor decompositions and applications*. SIAM review, 2009. **51**(3): p. 455-500.
6. Law, M.H. and A.K. Jain, *Incremental nonlinear dimensionality reduction by manifold learning*. IEEE transactions on pattern analysis and machine intelligence, 2006. **28**(3): p. 377-391.
7. Turk, M.A. and A.P. Pentland. *Face recognition using eigenfaces*. in *Proceedings. 1991 IEEE computer society conference on computer vision and pattern recognition*. 1991. IEEE Computer Society.
8. Belhumeur, P.N., J.P. Hespanha, and D.J. Kriegman, *Eigenfaces vs. fisherfaces: Recognition using class specific linear projection*. IEEE Transactions on pattern analysis and machine intelligence, 1997. **19**(7): p. 711-720.
9. Lu, H., K.N. Plataniotis, and A.N. Venetsanopoulos. *Multilinear principal component analysis of tensor objects for recognition*. in *18th International Conference on Pattern Recognition (ICPR'06)*. 2006. IEEE.
10. Yan, S., et al., *Multilinear discriminant analysis for face recognition*. IEEE Transactions on Image Processing, 2006. **16**(1): p. 212-220.
11. Hatch, A.O., S. Kajarekar, and A. Stolcke. *Within-class covariance normalization for SVM-based speaker recognition*. in *Ninth international conference on spoken language processing*. 2006.
12. Gutschoven, B. and P. Verlinde. *Multi-modal identity verification using support vector machines (SVM)*. in *Proceedings of the Third International Conference on Information Fusion*. 2000. IEEE.
13. Zhao, W., et al., *Face recognition: A literature survey*. ACM computing surveys (CSUR), 2003. **35**(4): p. 399-458.

14. Huang, D., *Robust Face Recognition based on Three Dimensional Tata*. 2011, Central School of Lyon, Écully: france.
15. Yang, J., et al., *Two-dimensional PCA: a new approach to appearance-based face representation and recognition*. IEEE transactions on pattern analysis and machine intelligence, 2004. **26**(1): p. 131-137.
16. Li, B. and K.-K. Ma. *Fisherface vs. eigenface in the dual-tree complex wavelet domain*. in *2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*. 2009. IEEE.
17. Bartlett, M.S., J.R. Movellan, and T.J. Sejnowski, *Face recognition by independent component analysis*. IEEE Transactions on neural networks, 2002. **13**(6): p. 1450-1464.
18. Comon, P., *Independent component analysis, a new concept?* Signal processing, 1994. **36**(3): p. 287-314.
19. Jutten, C. and J. Herault, *Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture*. Signal processing, 1991. **24**(1): p. 1-10.
20. Draper, B.A., et al., *Recognizing faces with PCA and ICA*. Computer Vision and Image Understanding, 2003. **91**(1-2): p. 115-137.
21. Pentland, A., B. Moghaddam, and T. Starner, *View-based and modular eigenspaces for face recognition*. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 1994, pp. 84-91, doi: 10.1109/CVPR.1994.323814.
22. Lu, J., K.N. Plataniotis, and A.N. Venetsanopoulos, *Face recognition using LDA-based algorithms*. IEEE Transactions on Neural networks, 2003. **14**(1): p. 195-200.
23. Vasilescu, M.A.O. and D. Terzopoulos. *Multilinear analysis of image ensembles: Tensorfaces*. in *European conference on computer vision*. 2002. Springer.
24. Cox, M.A. and T.F. Cox, *Multidimensional scaling*, in *Handbook of data visualization*. 2008, Springer. p. 315-347.
25. Lee, D.D. and H.S. Seung, *Learning the parts of objects by non-negative matrix factorization*. Nature, 1999. **401**(6755): p. 788-791.
26. Huang, Z.-H., et al., *Non-uniform patch based face recognition via 2D-DWT*. Image and Vision Computing, 2015. **37**: p. 12-19.
27. Friedman, J.H., *Regularized discriminant analysis*. Journal of the American statistical association, 1989. **84**(405): p. 165-175.
28. Hastie, T., A. Buja, and R. Tibshirani, *Penalized discriminant analysis*. The Annals of Statistics, 1995: p. 73-102.

29. Liu, W., et al. *Null space approach of fisher discriminant analysis for face recognition*. in *International Workshop on Biometric Authentication*. 2004. Springer.
30. Wang, X. and X. Tang. *Dual-space linear discriminant analysis for face recognition*. in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004*. 2004. IEEE.
31. Sufyanu, Z., et al., *Enhanced Face Recognition Using Discrete Cosine Transform*. *Engineering Letters*, 2016. **24**(1).
32. Howland, P. and H. Park, *Generalizing discriminant analysis using the generalized singular value decomposition*. *IEEE transactions on pattern analysis and machine intelligence*, 2004. **26**(8): p. 995-1006.
33. Lu, J., K.N. Plataniotis, and A.N. Venetsanopoulos. *Boosting linear discriminant analysis for face recognition*. in *Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429)*. 2003. IEEE.
34. Yang, Q., X. Ding, and Z. Chen. *Discriminant local feature analysis of facial images*. in *Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429)*. 2003. IEEE.
35. Nhat, V.D.M. and S. Lee. *Block LDA for face recognition*. in *International Work-Conference on Artificial Neural Networks*. 2005. Springer.
36. Zhou, D. and X. Yang. *Face recognition using enhanced fisher linear discriminant model with facial combined feature*. in *Pacific Rim International Conference on Artificial Intelligence*. 2004. Springer.
37. Zhang, W., et al. *Component-based cascade linear discriminant analysis for face recognition*. in *Chinese Conference on Biometric Recognition*. 2004. Springer.
38. Cevikalp, H., et al., *Discriminative common vectors for face recognition*. *IEEE Transactions on pattern analysis and machine intelligence*, 2005. **27**(1): p. 4-13.
39. Visani, M., C. Garcia, and J.-M. Jolion. *Normalized radial basis function networks and bilinear discriminant analysis for face recognition*. in *IEEE Conference on Advanced Video and Signal Based Surveillance, 2005*. 2005. IEEE.
40. Shah, J.H., et al., *A Survey: Linear and Nonlinear PCA Based Face Recognition Techniques*. *Int. Arab J. Inf. Technol.*, 2013. **10**(6): p. 536-545.
41. Shawe-Taylor, J. and N. Cristianini, *Kernel methods for pattern analysis*. 2004: Cambridge university press.
42. Hoffmann, H., *Kernel PCA for novelty detection*. *Pattern recognition*, 2007. **40**(3): p. 863-874.

43. Lima, A., et al., *On the use of kernel PCA for feature extraction in speech recognition*. IEICE TRANSACTIONS on Information and Systems, 2004. **87**(12): p. 2802-2811.
44. Arashloo, S.R. and J. Kittler, *Class-specific kernel fusion of multiple descriptors for face verification using multiscale binarised statistical image features*. IEEE Transactions on Information Forensics and Security, 2014. **9**(12): p. 2100-2109.
45. Lauer, F. and G. Bloch. *Méthodes SVM pour l'identification*. in *Journées Identification et Modélisation Expérimentale (JIME'2006)*. 2006.
46. Bach, F.R. and M.I. Jordan, *Kernel independent component analysis*. Journal of machine learning research, 2002. **3**(Jul): p. 1-48.
47. Zhang, T., et al., *Generalized discriminant analysis: A matrix exponential approach*. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2009. **40**(1): p. 186-197.
48. He, X., et al. *Learning a locality preserving subspace for visual recognition*. in *Proceedings Ninth IEEE International Conference on Computer Vision*. 2003. IEEE.
49. Socolinsky, D.A. and A. Selinger. *Thermal face recognition in an operational scenario*. in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004*. 2004. IEEE.
50. Yang, J., A.F. Frangi, and J.-y. Yang, *A new kernel Fisher discriminant algorithm with application to face recognition*. Neurocomputing, 2004. **56**: p. 415-421.
51. Wang, Y., et al., *Face recognition using nonlinear locality preserving with deep networks*, in *Proceedings of the 7th International Conference on Internet Multimedia Computing and Service*. 2015, Association for Computing Machinery: Zhangjiajie, Hunan, China. p. Article 66.
52. Wang, Q. and J. Li, *Combining local and global information for nonlinear dimensionality reduction*. Neurocomputing, 2009. **72**(10-12): p. 2235-2241.
53. Weinberger, K.Q. and L.K. Saul, *Unsupervised learning of image manifolds by semidefinite programming*. International journal of computer vision, 2006. **70**(1): p. 77-90.
54. !!! INVALID CITATION !!! [56, 57].
55. Sivasathya, M. and S.M. Joans, *Image Feature Extraction using Non Linear Principle Component Analysis*. Procedia Engineering, 2012. **38**: p. 911-917.
56. Vankayalapati, H.D. and K. Kyamakya. *Nonlinear feature extraction approaches with application to face recognition over large databases*. in *2009 2nd International Workshop on Nonlinear Dynamics and Synchronization*. 2009.

57. Li, S., et al., *Face Recognition Based on Nonlinear DCT Discriminant Feature Extraction Using Improved Kernel DCV*. IEICE Trans. Inf. Syst., 2009. **92-D**: p. 2527-2530.
58. Javidi, B., J. Li, and Q. Tang, *Optical implementation of neural networks for face recognition by the use of nonlinear joint transform correlators*. Applied Optics, 1995. **34(20)**: p. 3950-3962.
59. Ming-Hsuan, Y. *Face recognition using extended isomap*. in *Proceedings. International Conference on Image Processing*. 2002.
60. Liao, S., A.K. Jain, and S.Z. Li, *Partial face recognition: Alignment-free approach*. IEEE Transactions on pattern analysis and machine intelligence, 2012. **35(5)**: p. 1193-1205.
61. Kortli, Y., et al. *A novel face detection approach using local binary pattern histogram and support vector machine*. in *2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET)*. 2018. IEEE.
62. Brunelli, R. and T. Poggio, *Face recognition: Features versus templates*. IEEE transactions on pattern analysis and machine intelligence, 1993. **15(10)**: p. 1042-1052.
63. Viola, P. and M.J. Jones, *Robust real-time face detection*. International journal of computer vision, 2004. **57(2)**: p. 137-154.
64. Ahonen, T., A. Hadid, and M. Pietikäinen. *Face recognition with local binary patterns*. in *European conference on computer vision*. 2004. Springer.
65. Ojansivu, V. and J. Heikkilä. *Blur insensitive texture classification using local phase quantization*. in *International conference on image and signal processing*. 2008. Springer.
66. Chen, J., et al., *WLD: A robust local image descriptor*. IEEE transactions on pattern analysis and machine intelligence, 2009. **32(9)**: p. 1705-1720.
67. Kannala, J. and E. Rahtu. *Bsif: Binarized statistical image features*. in *Proceedings of the 21st international conference on pattern recognition (ICPR2012)*. 2012. IEEE.
68. Karaaba, M., et al. *Robust face recognition by computing distances from multiple histograms of oriented gradients*. in *2015 IEEE Symposium Series on Computational Intelligence*. 2015. IEEE.
69. Ojala, T., M. Pietikäinen, and D. Harwood, *A comparative study of texture measures with classification based on featured distributions*. Pattern recognition, 1996. **29(1)**: p. 51-59.
70. Khoi, P., L.H. Thien, and V.H. Viet, *Face Retrieval Based On Local Binary Pattern and Its Variants: A Comprehensive Study*. Int. J. Adv. Comput. Sci. Appl, 2016. **7**: p. 249-258.

71. Xi, M., et al. *Local binary pattern network: A deep learning approach for face recognition*. in *2016 IEEE international conference on Image processing (ICIP)*. 2016. IEEE.
72. Kambi Beli, I.L. and C. Guo, *Enhancing face identification using local binary patterns and k-nearest neighbors*. *Journal of Imaging*, 2017. **3**(3): p. 37.
73. Bonnen, K., B.F. Klare, and A.K. Jain, *Component-based representation in automated face recognition*. *IEEE Transactions on Information Forensics and Security*, 2012. **8**(1): p. 239-253.
74. Ren, J., X. Jiang, and J. Yuan. *Relaxed local ternary pattern for face recognition*. in *2013 IEEE International Conference on Image Processing*. 2013.
75. Hongliang, J., et al. *Face detection using improved LBP under Bayesian framework*. in *Third International Conference on Image and Graphics (ICIG'04)*. 2004.
76. Hussain, S.U., T. Napoléon, and F. Jurie. *Face recognition using local quantized patterns*. 2012.
77. Lenc, L. and P. Král, *Automatic face recognition system based on the SIFT features*. *Computers & Electrical Engineering*, 2015. **46**: p. 256-272.
78. Du, G., F. Su, and A. Cai. *Face recognition using SURF features*. in *MIPPR 2009: Pattern Recognition and Computer Vision*. 2009. International Society for Optics and Photonics.
79. Işık, Ş. and K. Özkan, *A comparative evaluation of well-known feature detectors and descriptors*. *International Journal of Applied Mathematics, Electronics and Computers*, 2014. **3**(1): p. 1-6.
80. Calonder, M., et al., *BRIEF: Computing a Local Binary Descriptor Very Fast*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012. **34**(7): p. 1281-1298.
81. Alahi, A., R. Ortiz, and P. Vandergheynst. *FREAK: Fast Retina Keypoint*. in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. 2012.
82. Leutenegger, S., M. Chli, and R.Y. Siegwart. *BRISK: Binary Robust invariant scalable keypoints*. in *2011 International Conference on Computer Vision*. 2011.
83. Lades, M., et al., *Distortion invariant object recognition in the dynamic link architecture*. *IEEE Transactions on computers*, 1993. **42**(3): p. 300-311.
84. Wiskott, L., et al., *Face recognition by elastic bunch graph matching*. *IEEE Transactions on pattern analysis and machine intelligence*, 1997. **19**(7): p. 775-779.
85. Lanitis, A., C.J. Taylor, and T.F. Cootes, *Automatic face identification system using flexible appearance models*. *Image and vision computing*, 1995. **13**(5): p. 393-401.

86. Lee, T.S., *Image representation using 2D Gabor wavelets*. IEEE Transactions on pattern analysis and machine intelligence, 1996. **18**(10): p. 959-971.
87. Fathima, A.A., et al. *Hybrid approach for face recognition combining Gabor Wavelet and Linear Discriminant Analysis*. in *2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS)*. 2015.
88. Barkan, O., et al. *Fast High Dimensional Vector Multiplication Face Recognition*. in *2013 IEEE International Conference on Computer Vision*. 2013.
89. Kim, D.-J., S.-H. Lee, and M.-K. Sohn, *Face recognition via local directional pattern*. International Journal of Security and Its Applications, 2013. **7**(2): p. 191-200.
90. Cho, H., et al., *An efficient hybrid face recognition algorithm using PCA and GABOR wavelets*. International Journal of Advanced Robotic Systems, 2014. **11**(4): p. 59.
91. Nefian, A.V. and M.H. Hayes. *Face detection and recognition using hidden Markov models*. in *Proceedings 1998 international conference on image processing. icip98 (cat. no. 98cb36269)*. 1998. IEEE.
92. Edwards, G.J., T.F. Cootes, and C.J. Taylor. *Face recognition using active appearance models*. in *European conference on computer vision*. 1998. Springer.
93. Cardinaux, F., C. Sanderson, and S. Marcel. *Comparison of MLP and GMM classifiers for face verification on XM2VTS*. in *International Conference on Audio-and Video-based Biometric Person Authentication*. 2003. Springer.
94. Nefian, A.V. and M.H. Hayes III. *Face recognition using an embedded HMM*. in *IEEE Conference on Audio and Video-based Biometric Person Authentication*. 1999. Citeseer.
95. Ouerhani, Y., M. Jridi, and A. Alfalou. *Fast face recognition approach using a graphical processing unit "GPU"*. in *2010 IEEE International Conference on Imaging Systems and Techniques*. 2010. IEEE.
96. Alfalou, A. and C. Brosseau, *Understanding Correlation Techniques for Face Recognition: From Basics to Applications*. Face Recognition, 2010.
97. Chihaoui, M., et al., *A survey of 2D face recognition techniques*. Computers, 2016. **5**(4): p. 21.
98. Heshner, C., A. Srivastava, and G. Erlebacher. *A novel technique for face recognition using range imaging*. in *Seventh International Symposium on Signal Processing and Its Applications, 2003. Proceedings*. 2003. IEEE.
99. Achermann, B., X. Jiang, and H. Bunke. *Face recognition using range images*. in *Proceedings. International Conference on Virtual Systems and MultiMedia VSMM'97 (Cat. No. 97TB100182)*. 1997. IEEE.

100. Heseltine, T., N. Pears, and J. Austin. *Three-dimensional face recognition: an eigensurface approach*. in *2004 International Conference on Image Processing, 2004. ICIP '04*. 2004.
101. Heseltine, T., N. Pears, and J. Austin. *Three-Dimensional Face Recognition: A Fishersurface Approach*. 2004. Berlin, Heidelberg: Springer Berlin Heidelberg.
102. Besl, P.J. and N.D. McKay, *A method for registration of 3-D shapes*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992. **14**(2): p. 239-256.
103. Medioni, G. and R. Waupotitsch, *Face modeling and recognition in 3-D*. 2003. 232-233.
104. Cook, J., et al. *Face recognition from 3d data using iterative closest point algorithm and gaussian mixture models*. in *Proceedings. 2nd International Symposium on 3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004*. 2004. IEEE.
105. Irfanoglu, M.O., B. Gokberk, and L. Akarun. *3D shape-based face recognition using automatically registered facial surfaces*. in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004*. 2004.
106. Wang, Y., et al. *Exploring Facial Expression Effects in 3D Face Recognition Using Partial ICP*. 2006. Berlin, Heidelberg: Springer Berlin Heidelberg.
107. Kyong, I.C., K.W. Bowyer, and P.J. Flynn. *Adaptive Rigid Multi-region Selection for Handling Expression Variation in 3D Face Recognition*. in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*. 2005.
108. Kakadiaris, I.A., et al., *Three-dimensional face recognition in the presence of facial expressions: An annotated deformable model approach*. IEEE Transactions on pattern analysis and machine intelligence, 2007. **29**(4): p. 640-649.
109. Li, X., J. Tao, and H. Zhang. *Expression-insensitive 3D face recognition using sparse representation*. in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 2009.
110. Ben Amor, B., *Contributions à la modélisation et à la reconnaissance faciales 3D*. 2006.
111. Blanz, V. and T. Vetter. *A morphable model for the synthesis of 3D faces*. in *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*. 1999.
112. Lu, X. and A. Jain, *Deformation modeling for robust 3D face matching*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008. **30**(8): p. 1346-1357.

113. Amberg, B., R. Knothe, and T. Vetter. *Expression invariant 3D face recognition with a morphable model*. in *2008 8th IEEE International Conference on Automatic Face & Gesture Recognition*. 2008. IEEE.
114. Al-Osaimi, F., M. Bennamoun, and A. Mian, *An expression deformation approach to non-rigid 3D face recognition*. *International Journal of Computer Vision*, 2009. **81**(3): p. 302-316.
115. ter Haar, F.B. and R.C. Veltkamp, *Expression modeling for expression-invariant face recognition*. *Computers & Graphics*, 2010. **34**(3): p. 231-241.
116. Guo, Y., et al., *A comprehensive performance evaluation of 3D local feature descriptors*. *International Journal of Computer Vision*, 2016. **116**(1): p. 66-89.
117. Lowe, D.G., *Distinctive image features from scale-invariant keypoints*. *International journal of computer vision*, 2004. **60**(2): p. 91-110.
118. Mian, A.S., M. Bennamoun, and R. Owens, *Keypoint detection and local feature matching for textured 3D face recognition*. *International Journal of Computer Vision*, 2008. **79**(1): p. 1-12.
119. Inan, T. and U. Halici, *3-D face recognition with local shape descriptors*. *IEEE transactions on Information Forensics and Security*, 2012. **7**(2): p. 577-587.
120. Huang, D., et al. *3D Face recognition using distinctiveness enhanced facial representations and local feature hybrid matching*. in *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. 2010.
121. Soltanpour, S., Q.M.J. Wu, and M. Anvaripour *Multimodal 2D-3D Face Recognition Using Structural Context and Pyramidal Shape Index*. *IET Conference Proceedings*, 2015. 2 (6 .)-2 (6 .).
122. Maes, C., et al. *Feature detection on 3D face surfaces for pose normalisation and recognition*. in *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. 2010. IEEE.
123. Smeets, D., et al., *meshSIFT: Local surface features for 3D face recognition under expression variations and partial data*. *Computer Vision and Image Understanding*, 2013. **117**(2): p. 158-169.
124. Li, H., et al. *Expression robust 3D face recognition via mesh-based histograms of multiple order surface differential quantities*. in *2011 18th IEEE International Conference on Image Processing*. 2011. IEEE.
125. Samir, C., A. Srivastava, and M. Daoudi, *Three-dimensional face recognition using shapes of facial curves*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006. **28**(11): p. 1858-1863.

126. Samir, C., et al., *An intrinsic framework for analysis of facial surfaces*. International Journal of Computer Vision, 2009. **82**(1): p. 80-95.
127. Berretti, S., A. Del Bimbo, and P. Pala, *3D face recognition using isogeodesic stripes*. IEEE transactions on pattern analysis and machine intelligence, 2010. **32**(12): p. 2162-2177.
128. Jahanbin, S., et al. *Three dimensional face recognition using iso-geodesic and iso-depth curves*. in *2008 IEEE Second International Conference on Biometrics: Theory, Applications and Systems*. 2008. IEEE.
129. Mpiperis, I., S. Malasiotis, and M.G. Strintzis, *3D face recognition by point signatures and iso-contours*. Proc. of SPPRA, 2007.
130. Li, L., et al., *3D face recognition by constructing deformation invariant image*. Pattern Recognition Letters, 2008. **29**(10): p. 1596-1602.
131. Mpiperis, I., S. Malassiotis, and M.G. Strintzis, *3-D face recognition with the geodesic polar representation*. IEEE Transactions on Information Forensics and Security, 2007. **2**(3): p. 537-547.
132. Feng, S., H. Krim, and I. Kogan. *3D face recognition using Euclidean integral invariants signature*. in *2007 IEEE/SP 14th Workshop on Statistical Signal Processing*. 2007. IEEE.
133. Smeets, D., et al., *A comparative study of 3-D face recognition under expression variations*. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2011. **42**(5): p. 710-727.
134. Drira, H., et al. *Pose and Expression-Invariant 3D Face Recognition using Elastic Radial Curves*. in *British machine vision conference*. 2010. Aberystwyth, United Kingdom.
135. Drira, H., et al., *3D face recognition under expressions, occlusions, and pose variations*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013. **35**(9): p. 2270-2283.
136. Ballihi, L., et al., *Boosting 3-D-geometric features for efficient face recognition and gender classification*. IEEE Transactions on Information Forensics and Security, 2012. **7**(6): p. 1766-1779.
137. Lei, Y., et al., *An efficient 3D face recognition approach using local geometrical signatures*. Pattern Recognition, 2014. **47**(2): p. 509-524.
138. Berretti, S., A. Del Bimbo, and P. Pala, *Sparse matching of salient facial curves for recognition of 3-D faces with missing parts*. IEEE Transactions on Information Forensics and Security, 2012. **8**(2): p. 374-389.

139. Ojala, T., M. Pietikainen, and T. Maenpaa, *Multiresolution gray-scale and rotation invariant texture classification with local binary patterns*. IEEE Transactions on pattern analysis and machine intelligence, 2002. **24**(7): p. 971-987.
140. Li, S.Z., et al. *Learning to fuse 3D+ 2D based face recognition at both feature and decision levels*. in *International Workshop on Analysis and Modeling of Faces and Gestures*. 2005. Springer.
141. Huang, Y., Y. Wang, and T. Tan. *Combining Statistics of Geometrical and Correlative Features for 3D Face Recognition*. in *BMVC*. 2006. Citeseer.
142. Huang, D., et al., *3-D face recognition using eLBP-based facial description and local feature hybrid matching*. IEEE Transactions on Information Forensics and Security, 2012. **7**(5): p. 1551-1565.
143. Tang, H., et al., *3D face recognition using local binary patterns*. Signal Processing, 2013. **93**(8): p. 2190-2198.
144. Li, H., et al., *Expression-robust 3D face recognition via weighted sparse representation of multi-scale and multi-component local normal patterns*. Neurocomputing, 2014. **133**: p. 179-193.
145. Werghi, N., S. Berretti, and A. Del Bimbo, *The mesh-lbp: a framework for extracting local binary patterns from discrete manifolds*. IEEE Transactions on Image Processing, 2014. **24**(1): p. 220-235.
146. Werghi, N., et al., *Boosting 3D LBP-based face recognition by fusing shape and texture descriptors on the mesh*. IEEE Transactions on Information Forensics and Security, 2016. **11**(5): p. 964-979.
147. Xu, C., et al. *Automatic 3D face recognition combining global geometric features with local shape variation information*. in *Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings*. 2004. IEEE.
148. Li, X. and H. Zhang. *Adapting Geometric Attributes for Expression-Invariant 3D Face Recognition*. in *IEEE International Conference on Shape Modeling and Applications 2007 (SMI '07)*. 2007.
149. Li, Y., et al. *An Efficient 3D Face Recognition Method Using Geometric Features*. in *2010 2nd International Workshop on Intelligent Systems and Applications*. 2010.
150. Li, X., T. Jia, and H. Zhang. *Expression-insensitive 3D face recognition using sparse representation*. in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 2009. IEEE.
151. Chua, C.-S., F. Han, and Y.-K. Ho. *3D human face recognition using point signature*. in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*. 2000. IEEE.

152. Wang, Y. and C.-S. Chua, *Robust face recognition from 2D and 3D images using structural Hausdorff distance*. Image and Vision Computing, 2006. **24**(2): p. 176-185.
153. Chang, K.I., K.W. Bowyer, and P.J. Flynn, *Multiple nose region matching for 3D face recognition under varying facial expression*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006. **28**(10): p. 1695-1700.
154. Faltemier, T.C., K.W. Bowyer, and P.J. Flynn, *A region ensemble for 3-D face recognition*. IEEE Transactions on Information Forensics and Security, 2008. **3**(1): p. 62-73.
155. Mian, A.S., M. Bennamoun, and R.A. Owens. *Matching tensors for pose invariant automatic 3D face recognition*. in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops*. 2005. IEEE.
156. Al-Osaimi, F.R., M. Bennamoun, and A. Mian, *Integration of local and global geometrical cues for 3D face recognition*. Pattern Recognition, 2008. **41**(3): p. 1030-1040.
157. Ming, Y., *Robust regional bounding spherical descriptor for 3D face recognition and emotion analysis*. Image and Vision Computing, 2015. **35**: p. 14-22.
158. Ocegueda, O., et al., *3D face discriminant analysis using Gauss-Markov posterior marginals*. IEEE transactions on pattern analysis and machine intelligence, 2012. **35**(3): p. 728-739.
159. Hadid, A., M. Pietikainen, and T. Ahonen. *A discriminative feature space for detecting and recognizing faces*. in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004*. 2004. IEEE.
160. Bronstein, A.M., M.M. Bronstein, and R. Kimmel. *Expression-invariant 3D face recognition*. in *international conference on Audio-and video-based biometric person authentication*. 2003. Springer.
161. Chang, K.J., K.W. Bowyer, and P.J. Flynn. *Effects on facial expression in 3D face recognition*. in *Biometric Technology for Human Identification II*. 2005. International Society for Optics and Photonics.
162. Lu, X., A.K. Jain, and D. Colbry, *Matching 2.5 D face scans to 3D models*. IEEE transactions on pattern analysis and machine intelligence, 2005. **28**(1): p. 31-43.
163. Lee, Y.-h. and J.-c. Shim. *Curvature based human face recognition using depth weighted hausdorff distance*. in *2004 International Conference on Image Processing, 2004. ICIP'04*. 2004. IEEE.
164. Russ, T.D., M.W. Koch, and C.Q. Little. *A 2D range Hausdorff approach for 3D face recognition*. in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops*. 2005. IEEE.

165. Gordon, G.G. *Face recognition based on depth maps and surface curvature*. in *Geometric Methods in Computer Vision*. 1991. International Society for Optics and Photonics.
166. Moreno, A.B., et al. *Face recognition using 3D surface-extracted descriptors*. in *Irish Machine Vision and Image Processing Conference*. 2003. Citeseer.
167. Gupta, S., et al. *3D face recognition founded on the structural diversity of human faces*. in *2007 IEEE Conference on Computer Vision and Pattern Recognition*. 2007. IEEE.
168. Nagamine, T., T. Uemura, and I. Masuda. *3D facial image analysis for human identification*. in *1992 11th IAPR International Conference on Pattern Recognition*. 1992. IEEE Computer Society.
169. Beumier, C. and M. Acheroy, *Automatic 3D face authentication*. *Image and Vision Computing*, 2000. **18**(4): p. 315-321.
170. Tanaka, H.T., M. Ikeda, and H. Chiaki. *Curvature-based face surface recognition using spherical correlation. principal directions for curved object recognition*. in *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*. 1998. IEEE.
171. Bronstein, A.M., M.M. Bronstein, and R. Kimmel, *Three-dimensional face recognition*. *International Journal of Computer Vision*, 2005. **64**(1): p. 5-30.
172. Beumier, C. and M. Acheroy, *Face verification from 3D and grey level clues*. *Pattern Recognition Letters*, 2001. **22**(12): p. 1321-1329.
173. Chenghua, X., et al. *Automatic 3D face recognition combining global geometric features with local shape variation information*. in *Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings*. 2004.
174. Gökberk, B., A.A. Salah, and L. Akarun. *Rank-Based Decision Fusion for 3D Shape-Based Face Recognition*. 2005. Berlin, Heidelberg: Springer Berlin Heidelberg.
175. Xu, C., et al., *Automatic 3D face recognition from depth and intensity Gabor features*. *Pattern recognition*, 2009. **42**(9): p. 1895-1905.
176. Wang, Y., C.-S. Chua, and Y.-K. Ho, *Facial feature detection and face recognition from 2D and 3D images*. *Pattern Recognition Letters*, 2003. **23**: p. 1191-1202.
177. Bronstein, A., M. Bronstein, and R. Kimmel, *Expression-Invariant 3D Face Recognition*. 2003. 62-70.
178. Chang, J., K. Bowyer, and P. Flynn, *Face recognition using 2D and 3D facial data*. *MMUA*, 2004.

179. Godil, A., S. Ressler, and P. Grother. *Face recognition using 3D facial shape and color map information: comparison and combination*. in *Biometric Technology for Human Identification*. 2004. International Society for Optics and Photonics.
180. Husken, M., et al., *Strategies and Benefits of Fusion of 2D and 3D Face Recognition*. 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2005. **0**: p. 174.
181. Lu, X. and A. Jain, *Integrating Range and Texture Information for 3D Face Recognition*. 2005. 156-163.
182. Lu, X., D. Colbry, and A. Jain, *Matching 2.5D Scans for Face Recognition*. 2004. 30-36.
183. Papatheodorou, T. and D. Rueckert, *Evaluation of automatic 4D face recognition using surface and texture registration*. 2004. 321-326.
184. Al-Osaimi, F.R., M. Bennamoun, and A. Mian, *Spatially Optimized Data-Level Fusion of Texture and Shape for Face Recognition*. *IEEE Transactions on Image Processing*, 2012. **21**(2): p. 859-872.
185. Papatheodorou, T. and D. Rueckert. *Evaluation of automatic 4D face recognition using surface and texture registration*. in *Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings*. 2004.
186. Zhang, G. and Y. Wang, *Robust 3D face recognition based on resolution invariant features*. *Pattern Recognition Letters*, 2011. **32**(7): p. 1009-1019.
187. Shih, P. and C. Liu, *Wavelet Features for 3D Face Recognition*, in *Cross Disciplinary Biometric Systems*. 2012, Springer. p. 93-116.
188. Elaiwat, S., et al., *A curvelet-based approach for textured 3D face recognition*. *Pattern Recognition*, 2015. **48**(4): p. 1235-1246.
189. De Lathauwer, L., B. De Moor, and J. Vandewalle, *A multilinear singular value decomposition*. *SIAM journal on Matrix Analysis and Applications*, 2000. **21**(4): p. 1253-1278.
190. De Lathauwer, L., B. De Moor, and J. Vandewalle, *On the best rank-1 and rank-(r_1, r_2, \dots, r_m) approximation of higher-order tensors*. *SIAM journal on Matrix Analysis and Applications*, 2000. **21**(4): p. 1324-1342.
191. Bader, B.W. and T.G. Kolda, *MATLAB tensor classes for fast algorithm prototyping: source code*. 2004, Sandia National Laboratories.
192. Saoud, A., et al., *Multimodal 2d + 3d multi-descriptor tensor for face verification*. *Multimedia Tools and Applications*, 2020. **79**(31): p. 23071-23092.
193. Lu, H., *Multilinear subspace learning for face and gait recognition*. 2008: Citeseer.

194. He, X., D. Cai, and P. Niyogi. *Tensor subspace analysis*. in *Advances in neural information processing systems*. 2006.
195. Lu, H., K.N. Plataniotis, and A. Venetsanopoulos, *Multilinear subspace learning: dimensionality reduction of multidimensional data*. 2013: CRC press.
196. Lathauwer, L.D., B.D. Moor, and J. Vandewalle, *On the Best Rank-1 and Rank-(R_1, R_2, \dots, R_N) Approximation of Higher-Order Tensors*. SIAM Journal on Matrix Analysis and Applications, 2000. **21**(4): p. 1324-1342.
197. Lu, H., K.N. Plataniotis, and A.N. Venetsanopoulos, *Uncorrelated multilinear discriminant analysis with regularization and aggregation for tensor object recognition*. IEEE Transactions on Neural Networks, 2008. **20**(1): p. 103-123.
198. Ouamane, A., et al., *Efficient tensor-based 2d+ 3d face verification*. IEEE Transactions on Information Forensics and Security, 2017. **12**(11): p. 2751-2762.
199. He, Y., et al., *An Iterative Closest Points Algorithm for Registration of 3D Laser Scanner Point Clouds with Geometric Features*. Sensors, 2017. **17**(8): p. 1862.
200. Ahonen, T., A. Hadid, and M. Pietikainen, *Face description with local binary patterns: Application to face recognition*. IEEE transactions on pattern analysis and machine intelligence, 2006. **28**(12): p. 2037-2041.
201. Hyvärinen, A., J. Hurri, and P.O. Hoyer, *Natural image statistics: A probabilistic approach to early computational vision*. Vol. 39. 2009: Springer Science & Business Media.
202. Hyvärinen, A. and E. Oja, *Independent component analysis: algorithms and applications*. Neural networks, 2000. **13**(4-5): p. 411-430.
203. Phillips, P.J., et al. *Overview of the face recognition grand challenge*. in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. 2005.
204. Savran, A., et al. *Bosphorus Database for 3D Face Analysis*. 2008. Berlin, Heidelberg: Springer Berlin Heidelberg.
205. Xu, C., et al. *3D Face Recognition Based on G-H Shape Variation*. 2005. Berlin, Heidelberg: Springer Berlin Heidelberg.
206. Ouamane, A., *Reconnaissance Biométrique par Fusion Multimodale du Visage 2D et 3D*. 2015.
207. Bessaoudi, M., et al., *Multilinear Enhanced Fisher Discriminant Analysis for robust multimodal 2D and 3D face verification*. Applied Intelligence, 2019. **49**(4): p. 1339-1354.

208. Hu, H., et al. *2D and 3D face recognition using convolutional neural network*. in *TENCON 2017 - 2017 IEEE Region 10 Conference*. 2017.
209. Feng, J., et al., *3D face recognition method based on deep convolutional neural network*, in *Smart Innovations in Communication and Computational Sciences*. 2019, Springer. p. 123-130.
210. Ouamane, A., et al., *A novel statistical and multiscale local binary feature for 2D and 3D face verification*. *Computers & Electrical Engineering*, 2017. **62**: p. 68-80.
211. Spreeuwens, L., *Fast and Accurate 3D Face Recognition*. *International Journal of Computer Vision*, 2011. **93**(3): p. 389-414.
212. Wang, Y., J. Liu, and X. Tang, *Robust 3D face recognition by local shape difference boosting*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010. **32**(10): p. 1858-1870.
213. Olivetti, E.C., et al. *Deep CNN for 3D Face Recognition*. 2020. Cham: Springer International Publishing.
214. Ammar, C., et al., *Evaluation of histograms local features and dimensionality reduction for 3D face verification*. *Journal of information processing systems*, 2016. **12**(3): p. 468-488.
215. Guo, Y., R. Wei, and Y. Liu, *Weighted gradient feature extraction based on multiscale sub-blocks for 3d facial recognition in bimodal images*. *Information*, 2018. **9**(3): p. 48.
216. Chouchane, A., et al., *3D face verification across pose based on euler rotation and tensors*. *Multimedia Tools and Applications*, 2018. **77**(16): p. 20697-20714.
217. Ming, Y., *Rigid-area orthogonal spectral regression for efficient 3D face recognition*. *Neurocomputing*, 2014. **129**: p. 445-457.
218. Bessaoudi, M., et al. *A Novel Hybrid Approach for 3D Face Recognition Based on Higher Order Tensor*. 2019. Cham: Springer International Publishing.
219. Elaiwat, S., et al., *A Curvelet-based approach for textured 3D face recognition*. *Pattern Recognition*, 2015. **48**(4): p. 1235-1246.
220. Weinberger, K.Q. and L.K. Saul, *Distance metric learning for large margin nearest neighbor classification*. *Journal of machine learning research*, 2009. **10**(2).
221. Liao, S., et al. *Person re-identification by local maximal occurrence representation and metric learning*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
222. Zhang, L., T. Xiang, and S. Gong. *Learning a discriminative null space for person re-identification*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

223. Koestinger, M., et al. *Large scale metric learning from equivalence constraints*. in *2012 IEEE conference on computer vision and pattern recognition*. 2012. IEEE.
224. Xiong, F., et al. *Person re-identification using kernel-based metric learning methods*. in *European conference on computer vision*. 2014. Springer.