



**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET
POPULAIRE**



**MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE
LA RECHERCHE SCIENTIFIQUE**

Université Kasdi Merbah Ouargla

**FACULTE DES NOUVELLES TECHNOLOGIE
Département d'Electroniques et de Télécommunications**

MÉMOIRE ACADÉMIQUE DU MASTER

Domaine : Electroniques et Télécommunications

Filière : Télécommunication

Spécialité: Systèmes de télécommunications.

Thème

**DÉTECTION DE LA VIVACITÉ DES VISAGES
AVEC LES CAMÉRAS PLÉNOPTIQUES**

Réalisé par :

KADRINE Sabrina

AMOUMNE Sana

Membres du jury :

Président : Mr. LATI Abd Elhaye (MCB)

Examineur : Mr. YOUSFA Abd Elmadjid (MCB)

Encadreur : Mme. BENZAOUI Wafa (MAA)

Année universitaire : 2019 /2020

Dédicace

*Je dédie ce travail à celui qui a planté dans mon âme l'ambition et la
persévérance, mon cher père,*

*Je dédie ce travail à celle qui a inséré le goût de la vie et le sens de la
Responsabilité ma chère mère,*

Je dédie ce travail à mes chères sœurs Imane, Hadjer, Zienb

Et mon cher frère Mohammed

Je dédie ce travail à celles que les lignes ne suffisent pas à les nommer, mes amies,

*Je le dédie à tous mes enseignants et à toutes mes enseignantes qui ont éclairci
mon chemin vers la connaissance,*

*A la fin, je le dédie à ceux qui m'ont encouragée pour arriver à ma destination : la
réussite.*

KADRINE Sabrina.

Dédicace

Je dédie ce travail à : Ma très chère mère pour son grand amour et sa patience.

Mon grand père que dieu le protège, qui ne cesse de me soutenir et m'encourager.

*Mon frère « Abdeslam » et ma sœur « Imane », Je leurs souhaite beaucoup de
réussite, ainsi que tous les membres de la famille.*

Ma chère amie « Zahra » qui a toujours pensé à moi.

*A tous ceux que j'aime, et tous ceux qui m'aiment et Tous ceux qui croyaient en
mes capacités.*

AMOUMENE Sana.

Remerciement

Nous tenons tout d'abord à remercier Dieu qui nous a données la puissance et la volonté pour arriver à ce stade.

Nous remercions bien entendu notre encadreur madame **BENZAOUI Wafa** pour ses orientations, ses conseils, ses remarques judicieuses ornées d'encouragements, et sa patience.

Nous adressons nos remerciements aux membres du jury qui nous ont fait l'honneur d'accepter d'évaluer ce travail et pour leur lecture et leur intérêt envers ce sujet.

Résumé

L'imagerie de champ lumineux est apparue comme une technologie permettant de capturer des informations visuelles plus riches de notre monde. Contrairement aux caméras conventionnelles, qui enregistrent la projection 2D des rayons lumineux en intégrant angulairement les rayons à chaque pixel, les caméras à champ lumineux capturent non seulement l'intensité lumineuse dans le plan image mais aussi l'information sur la direction des rayons lumineux. La vulnérabilité des systèmes actuels de reconnaissance faciale aux présences des attaques limite considérablement leur application en biométrie. Dans ce travail, et afin de surmonter ce problème, nous avons utilisé une approche d'utilisation de la caméra plénoptique pour défendre les attaques faciales, comme des photos faciales 2D imprimées et des images de tablette haute définition. La méthode proposée, repose sur trois types de caractéristiques extraites d'une paire d'images recentrées sont extraites pour discriminer les faux visages et les vrais visages. Pour vérifier les performances, nous utilisons des bases de données photographiques en champ lumineux et menons des expériences.

Mots clé : L'imagerie de champ lumineux, systèmes de reconnaissance faciale, la biométrie, la caméra plénoptique, photos faciales 2D, images de tablette haute définition, faux visages, vrais visages.

Abstract

Light field imagery has emerged as a technology to capture richer visual information of our world. Unlike conventional cameras, which record the 2D projection of light rays by angularly integrating the rays into each pixel, light field cameras capture not only the light intensity in the image plane but also information on the direction of the light rays. The vulnerability of current facial recognition systems to the presence of attacks considerably limits their application in biometrics. In this work, and in order to overcome this problem, we had used an approach of using the plenoptic camera to defend facial attacks, such as printed 2D facial photos and high definition tablet images. The proposed method is based on three types of features extracted from a pair of refocused images are extracted to discriminate false faces and real faces. To verify performance, we used light-field photographic databases and conduct experiments.

Keywords: Light field imagery, facial recognition systems, biometrics, plenoptic camera, 2D facial photos, high definition tablet images, fake faces, real faces.

ملخص

ظهرت صور المجال الضوئي كتقنية لالتقاط معلومات مرئية أكثر ثراءً عن عالمنا. على عكس الكاميرات التقليدية، التي تسجل الإسقاط ثنائي الأبعاد لأشعة الضوء من خلال دمج الأشعة بشكل زاوي في كل بكسل، لا تلتقط كاميرات المجال الضوئي شدة الضوء في مستوى الصورة فحسب، بل تلتقط أيضاً معلومات حول اتجاه أشعة الضوء. إن ضعف أنظمة التعرف على هوية الشخص الحالية لوجود هجمات الانتحال يحد إلى حد كبير من تطبيقاتها في مجال القياسات الحيوية. في هذا العمل، ومن أجل التغلب على هذه المشكلة، استخدمنا كاميرا المجال الضوئي للدفاع عن هجمات انتحال الوجه، مثل صور الوجه ثنائية الأبعاد المطبوعة وصور الأجهزة اللوحية عالية الدقة. تعتمد الطريقة المقترحة على ثلاثة أنواع من الخصائص المستخرجة من زوج من الصور المعاد تركيزها التي يتم استخلاصها لتمييز الوجوه الزائفة عن الوجوه الحقيقية. للتحقق من الأداء، استعملنا قواعد بيانات التصوير الفوتوغرافي للمجال الضوئي و أجرينا التجارب.

الكلمات المفتاحية : صور المجال الضوئي، القياسات الحيوية، كاميرات المجال الضوئي، صور الوجه ثنائية الأبعاد ، الصور اللوحية عالية الدقة ، الوجوه المزيفة ، الوجوه الحقيقية.

Table des matières

Remerciements	1
Résumé	
Table des matières	
Table des figures	
Introduction générale.....	1
 Chapitre I : L'imagerie à champs lumineux	
1. Introduction.....	2
2. L'imagerie conventionnelle.....	2
2.1 Les problèmes de la vision conventionnelle	3
3. Le champ lumineux et La fonction plénoptique.....	4
4. L'imagerie plénoptique.....	5
5. Paramétrage du champ lumineux	6
6. Acquisition du champ lumineux	8
6.1. Capteurs multiples	8
6.2. Multiplexage temporel	8
6.3. Multiplexage spatial	9
6.4. Multiplexage en fréquence.....	10
7. Caméra commerciale.....	11
7.1. La caméra Raytrix.....	11
7.2. La caméra Lytro.....	12
8. La formation d'image plénoptique	13
8.1. Équations d'imagerie	14

8.2. La refocalisation numérique	16
8.3. Algorithmes de synthèse d'images	17
9. conclusion	18

Chapitre II : Techniques d'anti-spoofing et présentation d'attaque

1. Introduction	19
2. La détection de la vivacité du visage	19
3. Les Techniques d'anti-spoofing	20
3.1. Les approches basées sur le matériel	21
3.1.1 Caractéristiques du capteur	22
3.1.2 Détection des clignements.....	22
3.1.3 Réponse au défi.....	22
3.2. Les approches basées sur logicielles	22
3.2.1 Approches statiques	23
3.2.1.1 Approches basées sur la texture.....	23
3.2.1.2 Approches basées sur la fréquence.....	23
3.2.1.3. Les approches hybrides.....	23
3.2.2 Approches dynamique	23
3.2.2.1 Approches basées sur le mouvement	24
3.2.2.2 Approches basées sur la texture.....	24
3.2.2.3 Schémas hybrides	24
4. Les type d'attaques	24
4.1 Attaque photo	25

4.2 Attaque vidéo	25
4.3 Attaque avec un masque	25
5. Conclusion	26

Chapitre III : La détection d'attaque par la caméra plénoptique

1. Introduction	27
2. Méthodologie proposée.....	27
2.1. Entropie d'intensité.....	28
2.2. Fonction de mise au point.....	29
2.3. Fonction d'histogramme de puissance	30
3. classificateur SVM	32
3.1 Classification binaire par hyperplan	32
3.2 Cas de données linéairement séparables	32
3.3 Cas des données non-linéairement séparables [Mou 05]	33
4. Conclusion	34

Chapitre IV : Résultats et discussion

1. Introduction	35
2. Base de données	35
3. Expérimentation et discussion.....	36
3.1 Evaluation quantitative.....	36
3.2 La laplacien modifiées de la paire d'image	38
4. Analyse de résultat	40
5. Conclusion	42
Conclusion générale.....	43
Références	

Liste des figures :

Chapitre I : L'imagerie à champs lumineux

Figure (1.1) : Système de premier ordre défocalisé.....	3
Figure (1.2) : Images à différents focus.....	4
Figure (1.3) : La fonction plénoptique décrit les informations disponibles pour un observateur à tout moment dans l'espace et le temps	5
Figure (1.4) : Comparaison du principe d'imagerie de base entre (a) le système d'image traditionnel et (b) le système d'imagerie à champ lumineux.....	6
Figure (1.5) : Représentations du champ lumineux en deux démentions.....	7
Figure (1.6) : Paramétrage des rayons lumineux à deux plans - le paramétrage relatif à deux plans.....	7
Figure (1.7) : Réseau de caméras.....	8
Figure (1.8) : Un multiplexage temporel avec une ouverture programmable	9
Figure (1.9) : Capture des champs lumineux multiplexés spatialement à l'aide d'un réseau de lentilles et de prismes	10
Figure (1.10): Capture des champs lumineux multiplexés en fréquence en plaçant un masque hétérodyne à proximité du capteur	11
Figure (1.11) : La R42, principal modèle de Raytrix à l'heure actuelle.....	12
Figure (1.12) : Le Lytro Illum tel que présenté par ses créateurs	13
Figure (1.13): Dessine en bleu le cône de rayons contribuant à une valeur de pixel de photographie.....	14
Figure (1.14) : Montre une série d'images avec différentes valeurs de profondeur de champ à partir d'un champ de lumière	15

Figure (1.15) : Montre une série d'images avec différentes valeurs de profondeur de champ à partir d'un champ de lumière.....16

Chapitre II : Techniques d'anti-usurpation et présentation d'attaque

Figure (2.1) : Système de reconnaissance faciale comprenant la détection de la vivacité du visage.....20

Figure (2.2) : La classification des algorithmes de PAD.....21

Chapitre III : la détection d'attaque par la caméra plénoptique

Figure (3.1) : Organigramme de la méthode de détection de la vivacité du visage proposée..... 28

Figure (3.2) : Organigramme d'extraction de la fonction de mise au point29

Figure (3.3) : Organigramme d'extraction des caractéristiques de l'histogramme de puissance31

Chapitre IV : Résultats et discussion

Figure (4.1) : Paire Images ainsi que les LAPM correspondantes.....38

Liste des tableaux

Chapitre IV : Résultats et discussion

Tableau (4.1) : Bases de données de champ lumineux.....	36
Tableau (4.2) : Mesurer les fausses erreurs de détection.....	38
Tableau (4.3) : Le taux de précision (%) des trois fonctionnalités dans différentes attaques d'usurpation d'identité.....	40
Tableau (4.4) : HTER et standard déviation de la méthode proposée et comparissant avec les références.....	41

Introduction générale

De nos jours, la caméra plénoptique attire l'attention des ingénieurs. Elle a été largement utilisée dans plusieurs domaines qui dépendent de la sécurité et de la haute précision, par exemple le paiement en ligne, système de l'enregistrement des présences et Les systèmes d'accès basés sur la reconnaissance faciale...etc.

Afin de développer les systèmes de sécurité biométriques d'authentification d'identité, la technologie de reconnaissance faciale (FR) a été appliquée. Dans ce travail, on essaiera de présenter l'un des systèmes biométriques les plus récents et les plus répandus dans le monde, et cela grâce à sa simplicité et son efficacité : il s'agit du système de détection de vivacité de visages pour la reconnaissance faciale avec la caméra plénoptique. En effet, le visage est certainement la caractéristique biométrique que l'on utilise pour s'identifier. Le but de ce système est de contrôler l'accès par les systèmes biométriques et permet de déterminer si un visage saisi est directement capturé à partir d'un humain ou non, en plus de l'automatisation des certaines applications telles que la télésurveillance, l'accès à des endroits sécurisés, etc.

L'utilisation de la caméra à champ lumineux nous permettions d'obtenir un ensemble des informations 3D d'une scène à l'aide d'une seule caméra en une seule exposition. Ceci est fait en ajoutant un réseau de microlentilles à une caméra conventionnelle de telle sorte qu'elle permette d'enregistrer des informations sur les angles des rayons lumineux ainsi que leur intensité et leur position. Ces informations peuvent être traitées par l'ordinateur, par un simple clic, un logiciel permet de faire la mise au point sur une des zones souhaitées. Ces caméras sont récemment devenues disponibles dans le commerce et sont principalement utilisées dans des applications photographiques à grand public.

La mise en œuvre de ce système est basée sur les caractéristiques de la refocalisation numérique. Cette méthode repose sur la différence entre les propriétés des visages réels et les visages usurpés. Nous avons utilisé la fonction l'entropie d'intensité, le laplacien modifié (LAPM) et l'histogramme de puissance comme des descripteurs de caractéristiques et un classificateur SVM pour classifier les visages en termes usurpés ou réels.

Chapitre I

L'imagerie plénoptique

1. Introduction

Au début du XXe siècle, les scientifiques découvrent une nouvelle approche de la photographie quotidienne, la caméra plénoptique ou l'imagerie en champ lumineux. Cette tendance d'imagerie représente l'une des nouvelles générations des méthodes d'imagerie. Cela contribue à résoudre les problèmes liés à la mise au point des images avec précision, et permettre d'utiliser de nombreuses fonctionnalités, qui offre une expérience à l'utilisateur plus réaliste. Ce système peut en effet, capturer et enregistrer la quantité de lumière, quel que soit sa direction. Ces caractéristiques avancées de cette approche sont susceptibles d'être exploitées dans de nombreux domaines de la recherche et de l'industrie, tel que la robotique, la vision par ordinateur, la médecine, la microscopie...etc.

Dans ce chapitre, nous présentons l'imagerie conventionnelle ainsi que les problèmes liés à cette vision. Nous apprendrons également à connaître qu'est-ce que : un champ lumineux, une fonction plénoptique et l'acquisition du champ lumineux. Nous avons aussi présenté la formation d'image, les équations d'imagerie et la refocalisation numérique de la caméra plénoptique. Ensuite nous donnons un aperçu des algorithmes de synthèse d'image et certaines applications qui y sont utilisées cette nouvelle vision.

2. L'imagerie conventionnelle

L'imagerie conventionnelle mesure la lumière le long de chaque faisceau individuel lorsqu'elle se propage au niveau du capteur en combinant les rayons lumineux qui frappent chaque pixel de toutes les directions. Où ce système d'imagerie est formé d'un capteur à des pièces optiques qui permettent de prendre une photo, et une lentille qui capture les faisceaux optiques (rayons lumineux) [5]. Chaque point de l'objet envoie la lumière dans toutes les directions, la lentille de la caméra ne prend que les rayons qui frappent sur elle et les collecte.

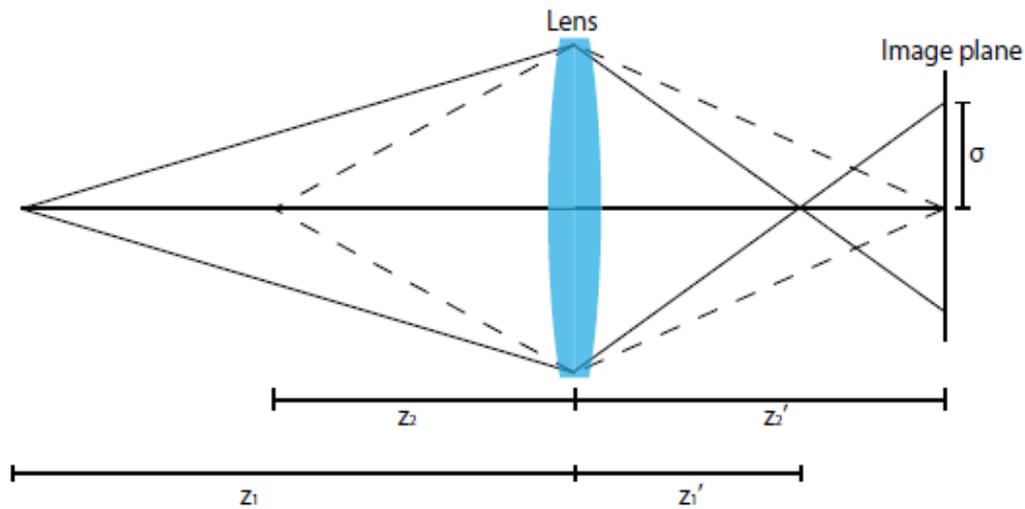


Figure (I.1) : Système de premier ordre défocalisé.

Ces équations peuvent être utilisées pour estimer la profondeur d'un objet.

$$\frac{1}{z'} = \frac{1}{z} = \frac{1}{f} \quad (I.1)$$

Où σ le diamètre de cercle de flou, z est la distance entre l'objet et le plan principale avant, z' est la distance entre le plan principale arrière et le plan image, et f est la distance focale du système d'imagerie ($1/f$ est communément désigné comme la puissance du système optique) [16].

2.1 Les problèmes de la vision conventionnelle

Lorsque nous prenons une image avec un appareil photo conventionnelle, l'objet qui se trouve dans le plan focal semble claire, mais l'objet qui est hors du plan focal semble flou, c'est l'un des problèmes major des caméras conventionnelles. Chaque point de l'objet envoie la lumière dans toutes les directions, la lentille de la caméra ne prend que les rayons frappent sur elle et les collecte ; ainsi l'imagerie conventionnelle n'enregistre pas la plupart des informations sur la lumière entrant dans elle et ne pouvons pas reprendre ces informations sur le hors objet focal après l'acquisition. Ceci indique que ce système fait la somme totale des rayons lumineux frappant chaque point de l'image [5].

Par exemple, si nous voulons faire l'analyse de la lumière déposée sur un pixel, une photographie ne nous dit rien sur la différence entre la lumière provenant d'une partie de l'objectif et la lumière provenant d'une autre. Ils' avère que ces différences sont les éléments cruciaux des informations manquantes qui conduisent au problème de mise au point dans la photographie conventionnelle [5].

Pour cette raison, les scientifiques ont pensé à trouver d'autres moyens qui leur permettraient de capturer tous les rayons lumineux émis par la scène. Pour pouvoir reprendre toutes les informations émises par la scène hors le plan focal après l'acquisition.



Figure (1.2) : Images à différente focus.

3. Le champ lumineux et La fonction plénoptique

Dans un milieu homogène (comme l'espace libre), les rayons lumineux se propagent en lignes droites. Pour les quantifier et les utiliser pour les calculs, nous avons besoin d'une fonction mathématique pour décrire chaque rayon. Pour tenter de comprendre les substances fondamentales de la vision humaine et de la machine. Le champ lumineux est intrinsèquement lié à la fonction plénoptique d'Adelson et Bergen (1991) qui pour chaque point 3D d'une scène, renvoie la valeur de la radiance le long d'un rayon lumineux donné. Cette fonction a initialement sept paramètres : trois pour la position du point dans l'espace ($x ; y ; z$), deux pour la direction du rayon lumineux le traversant ($\theta ; \varphi$), ainsi que la longueur d'onde (λ) et le temps (t).

$$L = P(\theta ; \varphi ; \lambda ; t ; x ; y ; z) \quad (1.2)$$

Il est courant dans l'étude de la formation d'images de ne trouver que les cinq premiers paramètres précédemment cités s'il l'on suppose que cette dernière est indépendante de la longueur d'onde λ (La lumière est monochromatique) et ne varie pas avec le temps t .

$$L = P(\theta ; \varphi ; x ; y ; z) \quad (1.3)$$

Sous l'hypothèse que la radiance est constante le long d'un rayon optique ne rencontrant pas d'obstacle ; donc seulement deux paramètres sont nécessaires pour l'emplacement, en omettant effectivement z , alors on peut réduire la fonction plénoptique à quatre paramètres (Levoy et Hanrahan, 1996), qui est alors entièrement décrit par ces paramètres.

L'espace des paramètres qui caractérisent les rayons optiques est appelé l'espace plénoptique. Les deux paramètres spatiaux $(x ; y)$ caractérisent le centre de projection ou centre optique, pupille de l'observateur au travers de laquelle une infinité de rayons se concentre. L'ensemble de ces rayons, ou plutôt leur direction, est définie par les deux autres paramètres, que l'on qualifie d'angulaires [15].

$$L = P(x, y, \theta, \varphi) \quad (1.4)$$

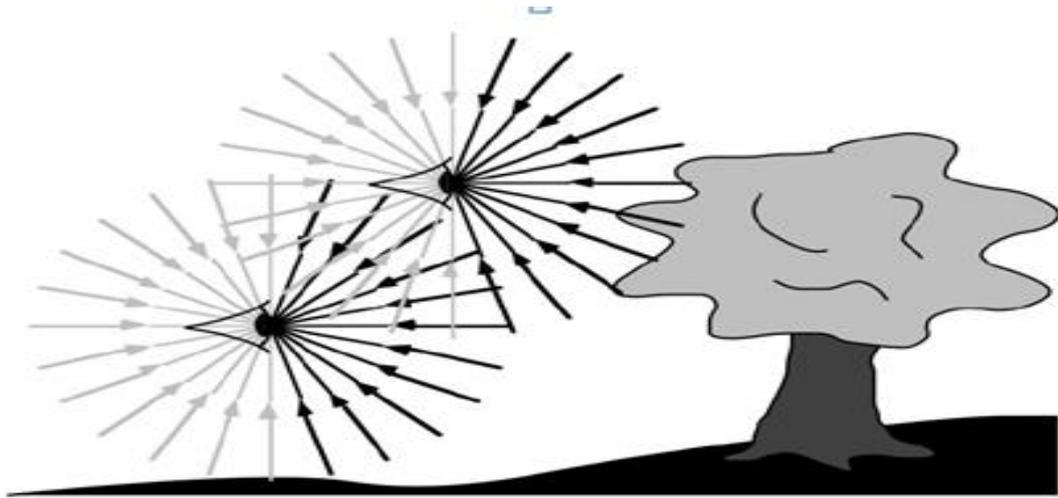


Figure (1.3) : La fonction plénoptique décrit les informations disponibles pour un observateur à tout moment dans l'espace et le temps. Voici deux yeux schématiques - que l'on devrait considérer comme ponctués de pupilles rassemblant des crayons de rayons lumineux. Un véritable observateur ne peut pas voir les rayons lumineux venant de derrière, mais la fonction plénoptique inclut ces rayons [1].

4. L'imagerie plénoptique

Une caméra plénoptique est constituée d'une lentille d'imagerie et un réseau de lentilles captant l'information de profondeur du champ lumineux, pour créer plusieurs images subdivisées de la pupille d'imagerie placé devant un capteur image. L'échantillonnage des rayons est dépendant du nombre de microlentilles du réseau utilisé. Ce système de caméra est capable de capturer des informations spatiales et angulaires sur le réseau de capteurs. Les résultats obtenus permettent d'imaginer des données tridimensionnelles (3D) concernant une scène. Ceci permet en particulier de faire la mise au point par post-traitement (les mesures 3D des objets et le changement de perspective de l'image capturé dans le post-traitement) [19].

Comme nous l'avons mentionné plus tôt, Le système d'imagerie traditionnel illustré sur la Figure (1.4) (a), enregistre uniquement les informations 2D de la scène. Pour surmonter la limitation, les systèmes.

d'imagerie de champ lumineux visent à collecter à la fois la quantité totale de lumière à chaque point de la photo capteur et la quantité de lumière déplaçant le long de chaque rayon qui coupe le capteur. Pour y parvenir, Adelson et Wang ont proposé un système d'imagerie à champ lumineux basé sur une lentille et bien implémenté dans une caméra à main par Ng et al. La figure (1.4) (b) est une illustration simple de l'imagerie basée sur le système microlentille. Un réseau de microlentilles est placé sur le plan image de l'objectif principal de la caméra. Le capteur d'images est positionné dans le plan focal du réseau de microlentilles. Les lumières proviennent des points sur le plan de mise au point [la couleur rouge sur la figure (1.4) (a)] et convergent vers un point sur le capteur d'image dans un système d'imagerie traditionnel ; cependant, dans un système d'imagerie à champ lumineux, les microlentilles à ces points séparent ces rayons lumineux en fonction de leur direction [la couleur rouge sur la figure (1.4) (b)]. Les différentes parties du capteur recevront le rayon lumineux de différentes directions. De cette façon, le plan focal peut être changé après l'exposition et les informations de profondeur peuvent en être obtenues [4].

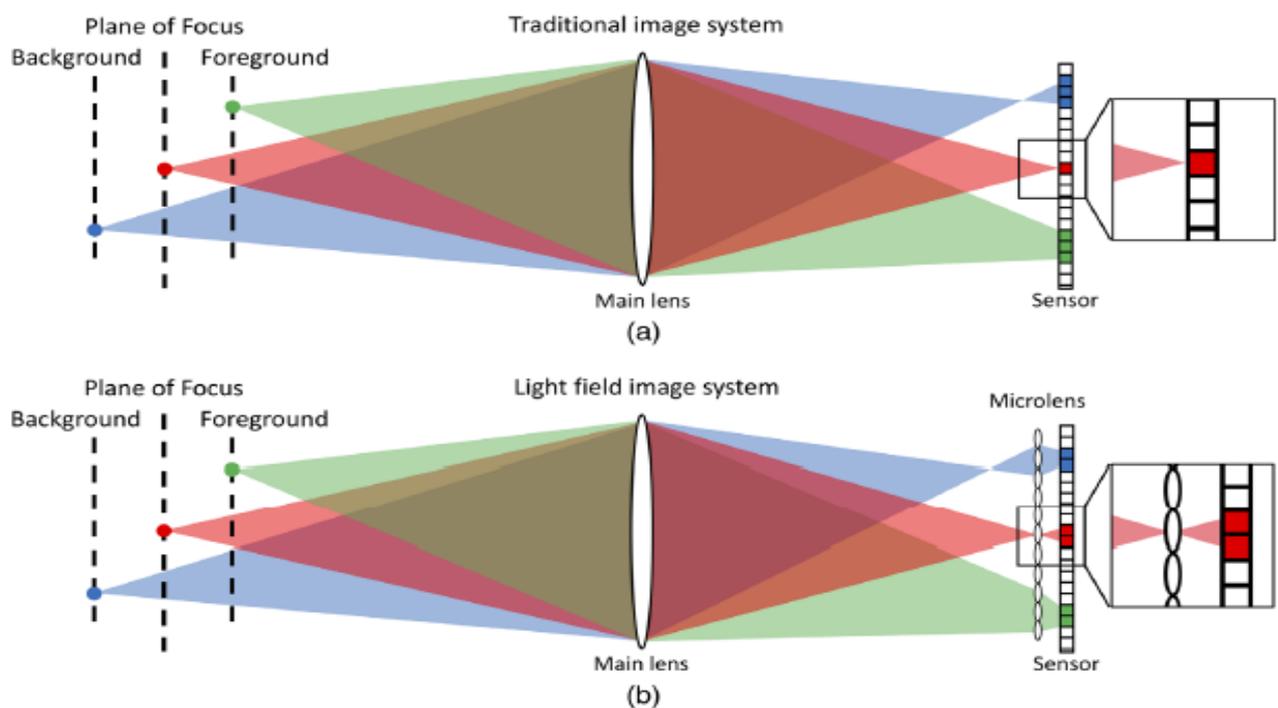


Figure (1.4) : Comparaison du principe d'imagerie de base entre (a) le système d'image traditionnel et (b) le système d'imagerie à champ lumineux.

5. Paramétrage du champ lumineux

Certaines paramétrisations alternatives du champ lumineux 4D, qui représente le flux de lumière à travers une région vide de l'espace à trois dimensions. A gauche : des points sur une surface plane ou

courbe et directions laissant chaque point. Centre : les paires de points sur la surface d'une sphère. A droite : des paires de points sur les deux plans en position générale (ce qui signifie tout) [2].

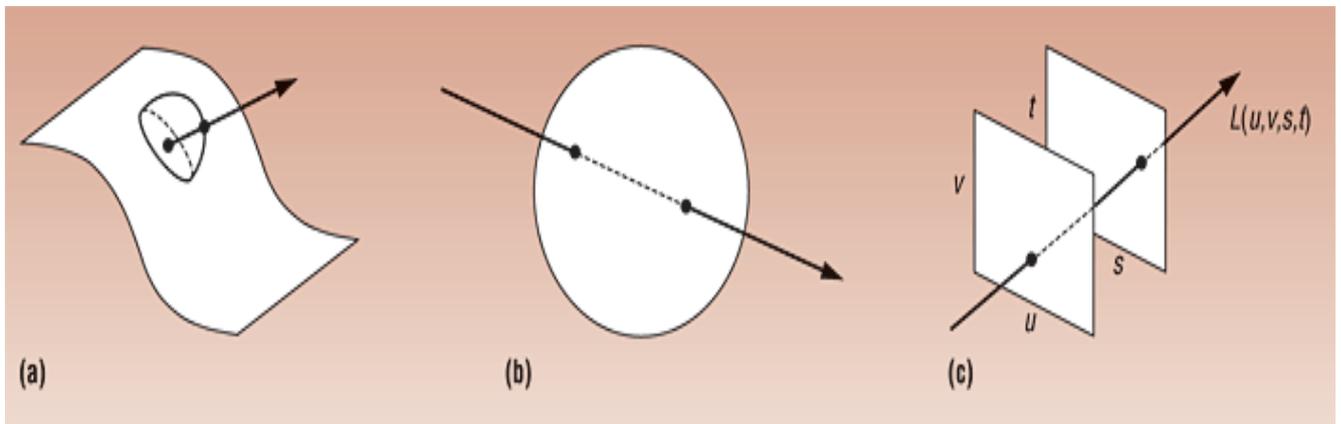


Figure (1.5) : Représentations du champ lumineux en deux démentions

La plus utilisée est la paramétrisation à deux plans, où le point d'intersection d'un rayon avec deux plans parallèles décrit complètement sa position et son orientation dans l'espace. Par convention, le plan s, t est plus proche de la caméra et le plan u, v est plus proche de la scène. Cette réalisation permet de paramétrer les rayons lumineux en quatre dimensions au lieu de cinq. Une manière courante de le faire mesure les points d'intersection de chaque rayon avec deux plans de référence parallèles, ne nécessitant que quatre nombres pour décrire le rayon, deux pour chacun de position et de direction.

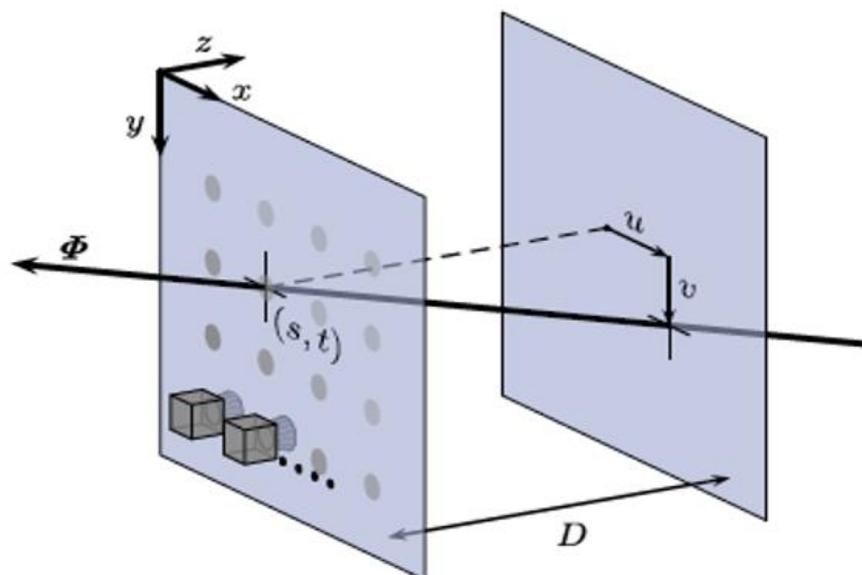


Figure (1.6) : Paramétrage des rayons lumineux à deux plans - le paramétrage relatif à deux plans.

6. Acquisition du champ lumineux

Pour acquérir un champ lumineux 4D, nous devons capturer plusieurs échantillons le long des dimensions angulaires. Les approches existantes d'acquisition de champ lumineux peuvent être divisées en quatre catégories fondamentales [8].

6.1 Capteurs multiples

L'approche de capture multi-capteurs nécessite un réseau de capteurs d'image répartis sur une surface plane ou sphérique pour capturer simultanément des échantillons de champ lumineux à partir de différents points de vue. Les dimensions spatiales (u et v) du champ lumineux sont déterminées par les capteurs, tandis que les dimensions angulaires (s et t) sont déterminées par le nombre de caméras et leur répartition.

Par conséquent, le champ lumineux 4D est enregistré par la combinaison des images capturées. En 2002, Yang et al. ont décrit une conception utilisant un réseau de 88 caméras vidéo pour la capture dynamique du champ lumineux [8].



Figure (1.7): réseau de caméras.

6.2. Multiplexage temporel

Contrairement à l'approche multi-capteurs, une approche de capture séquentielle dans le temps utilise un seul capteur d'image pour capturer des échantillons multiples du champ lumineux à travers de multiples expositions. L'approche typique utilise un capteur monté sur un portique mécanique pour mesurer le champ lumineux à différentes positions. Deux systèmes de portique ont été présentés, l'un était un portique contrôlé par ordinateur avec quatre degrés de liberté, une traduction en X et Y, un signe de tête et un tremblement ;

un autre était un portique Lego Mindstorms dans lequel les moteurs ont des contrôleurs de rotateurs permettant à la caméra de se déplacer le long de trajectoires précises et bien définies. Unger et al. ont utilisé une seule caméra montée sur une platine linéaire motorisée à deux degrés de liberté, translation en X et Y, pour capturer les champs lumineux [8].

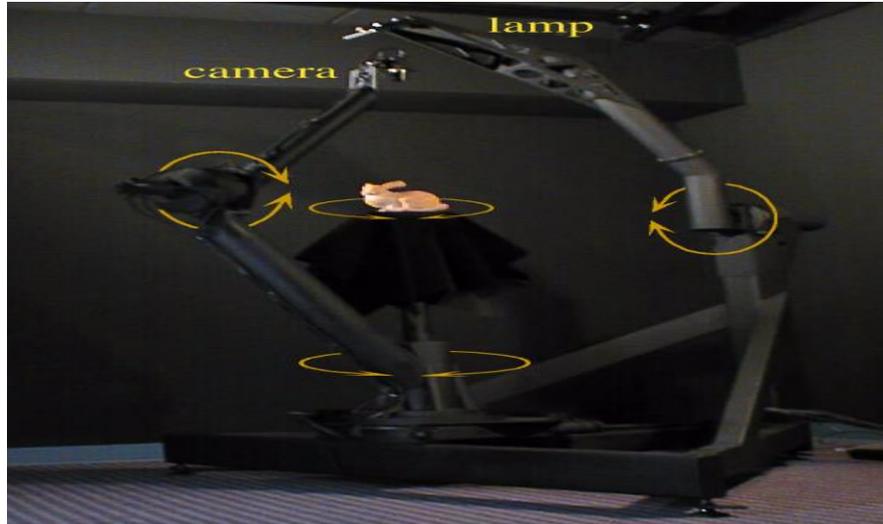


Figure (1.8) : un multiplexage temporel avec une ouverture programmable.

6.3. Multiplexage spatial

Dans le multiplexage spatial, un tableau d'images entrelacées représentant des échantillons de différentes tranches 2D du champ lumineux est capturé par le capteur. La plupart des approches de multiplexage spatial sont mises en œuvre en utilisant un réseau de microlentilles ou un réseau de lentilles monté sur le capteur d'image. Il est intéressant de noter que c'est l'une des premières approches de l'imagerie en champ lumineux. En 1908, Lippmann a utilisé cette approche dans le développement de sa « photographie intégrale » [8].



Figure (1.9) : capture des champs lumineux multiplexés spatialement à l'aide d'un réseau de lentilles et de prismes.

6.4. Multiplexage en fréquence

Contrairement à l'approche de multiplexage spatial, qui entrelace les tranches de champ lumineux 2D sur le plan du capteur, l'approche de multiplexage de fréquence code différentes tranches 2D du champ lumineux dans différentes bandes de fréquences. Typiquement, les approches de multiplexage en fréquence utilisent un masque de modulation pour obtenir une certaine propriété dans le domaine de Fourier. Veeraraghavan et al ont décrit un cadre théorique pour l'acquisition de champ lumineux 4D en utilisant un masque d'atténuation dans le chemin optique d'un capteur d'image conventionnel ; ils ont appelé l'approche « Photographie pommelée ». Plutôt que de mélanger les rayons lumineux, le masque à motifs les a atténués et codés sur le capteur d'image. L'image transformée de Fourier est réorganisée en plans 4D puis une transformée de Fourier inverse est appliquée pour restaurer le champ lumineux [8].



Figure (1.10): capture des champs lumineux multiplexés en fréquence en plaçant un masque hétérodyne à proximité du capteur.

7. Caméra commerciale

À ce jour, de nombreuses cameras plénoptiques commerciale sont utilisées dans plusieurs domaines.

7.1. La caméra Raytrix

Raytrix a été fondée en 2008 à Kiel, au nord de l'Allemagne, par un universitaire et un entrepreneur local. Leur première commercialisation d'une caméra vidéo plénoptique date de 2010, soit avant Lytro. Leurs produits s'adressent à un public averti et professionnel, avec de nombreuses applications industrielles [11].



Figure (I.11) : Caméra de champ lumineux 3D Raytrix R11.

7.2 La caméra Lytro

L'appareil photo d'origine est un tube carré de moins de cinq pouces de long avec une ouverture d'objectif à une extrémité et un écran tactile LCD de 1,52 pouces (38,6 mm) à l'autre. L'appareil photo d'origine dispose d'un capteur de 11 méga ray. L'objectif a un zoom optique 8x et une ouverture $f / 2.0$. La première génération de l'appareil photo est disponible en deux options : une avec 8 Go de mémoire (qui peut contenir 350 photos) et une avec 16 Go (qui peut contenir 750 photos). Méga ray est une mesure que Lytro utilise pour décrire le nombre de méga pixels dans le capteur sous un réseau de microlentilles. Les données brutes sont traitées pour produire des photographies avec une résolution de 1,2 méga pixel [20].



Figure (1.12) : Le Lytro Illum tel que présenté par ses créateurs.

8. La formation d'image plénoptique

Grâce à un réseau de microlentilles situé entre la lentille principale et le capteur, l'appareil photographique plénoptique délivre une image ayant capturé tous les rayons lumineux dans les différents plans de la scène à photographier. C'est comme si on avait pris plusieurs photographies en une seule. Une fois cette image sur l'ordinateur, par un simple clic, un logiciel permet de faire la mise au point sur une des zones souhaitées [21].

Chaque position sur le plan d'imagerie photosensible résume tous les rayons de lumière qui s'y terminent.

Le cône illustré dans la figure ci-dessous correspond (en 2d) à la bande verticale bleue sur le diagramme d'espace de rayon. Les rayons dans le cône partagent la même interception de film x , mais varient sur toutes les positions u sur la lentille [5].

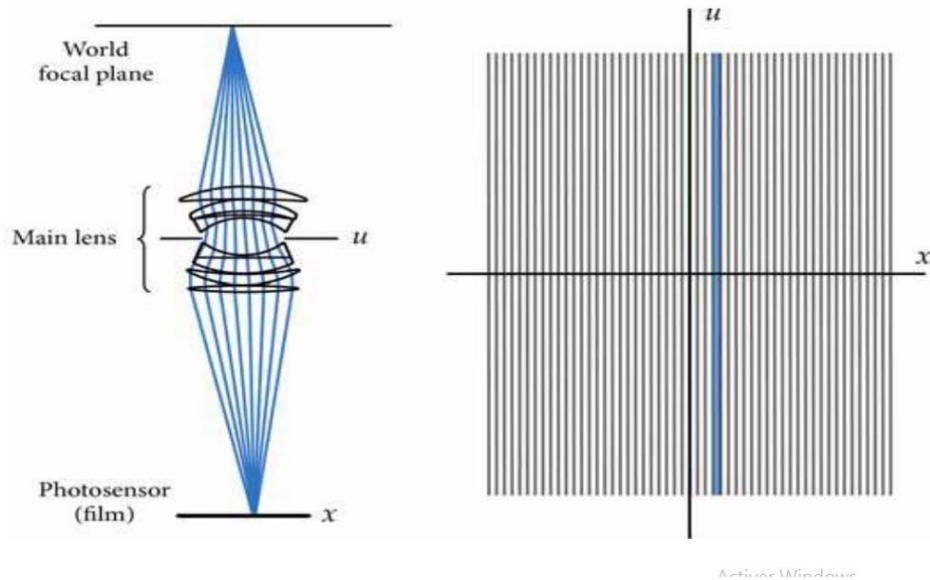


Figure (1.13) : Dessine en bleu le cône de rayons contribuant à une valeur de pixel de photographie.

Les différents pixels de la photographie ont des intersections différentes, ils correspondent donc à différentes lignes verticales sur l'espace des rayons. En fait, l'espace des rayons dessiné sur la figure (2.1) est recouvert de bandes verticales, chaque bande étant l'ensemble des rayons additionnés par un pixel de photographie différent. Ce dessin montre que la formation d'une photographie complète correspond sur le diagramme d'espace de rayons à une projection verticale des valeurs du champ lumineux. La projection préserve l'emplacement spatial x des rayons, mais détruit les informations directionnelles [5].

8.1. Équations d'imagerie

Cette relation mathématique est une base naturelle pour le calcul de photographies focalisées à différentes profondeurs des champs lumineux enregistrés par la caméra [5].

$$E_F = \frac{1}{F^2} \int \int L_F(x, y, U, V) \cos^4 \theta \, dU \, dV \quad (1.5)$$

Où F est la séparation entre la lentille et le film, $E_F(x, y)$ est l'irradiance du film à la position (x, y) , L_F est le champ lumineux paramétré par les plans à la séparation F , et θ est l'angle entre le rayon (x, y, u, v) et le plan du film normal. Par souci de simplicité, l'équation (1.5) suppose également que les plans u, v et x, y sont infinis et que L est simplement nul au-delà des limites physiques de l'objectif et du capteur.

Pour simplifier davantage les équations dans les dérivations tout au long de la thèse, absorbons également le $\cos^4\theta$ dans la définition du champ lumineux lui-même, en redéfinissant

$$L(x, y, u, v) := L(x, y, u, v) \cos^4\theta$$

Les photographies se sont concentrées à des autres profondeurs que le plan de paramétrage x . Comme le montre la figure (1.13), la mise au point à différentes profondeurs correspond à une modification de la séparation entre la lentille et le plan du film, entraînant un cisaillement de la trajectoire des lignes d'intégration sur l'espace des rayons. Si nous considérons la photographie focalisée sur une nouvelle profondeur de film de F , alors dériver son équation d'imagerie consiste à exprimer $L_F(x, u)$ en termes de $L(x, u)$ puis à appliquer l'équation (1.5).

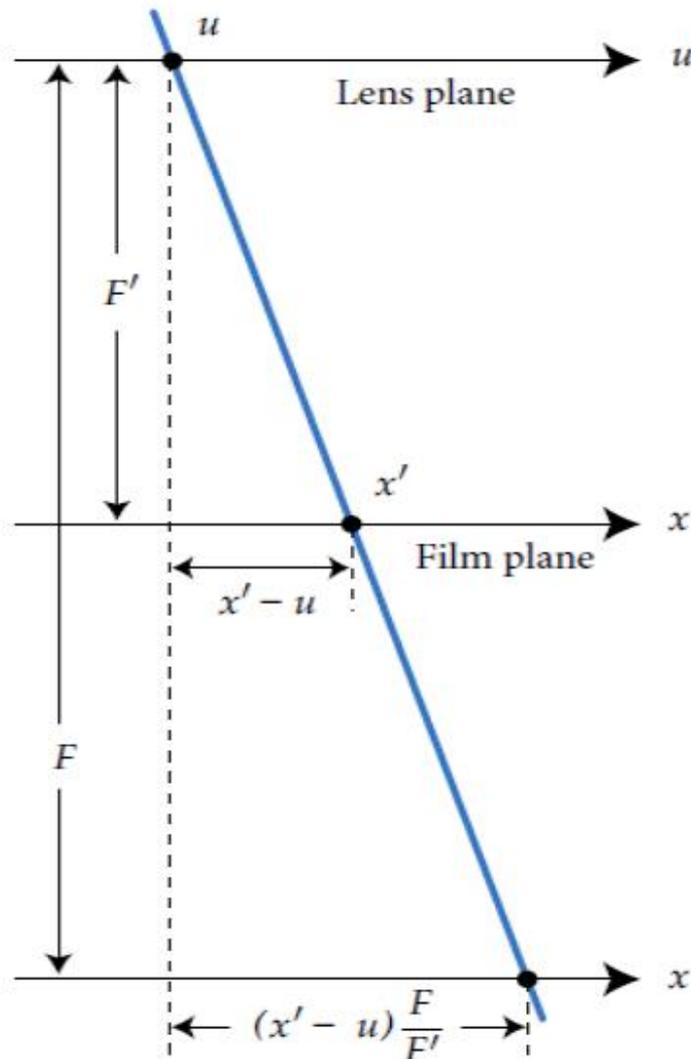


Figure (1.14) : Transformation des coordonnées rayon-espace.

Le diagramme ci-dessus est une construction géométrique qui illustre comment un rayon paramétré par les plans x' et u pour L_F' peut être reparamétriser par son intersection avec les plans x et u pour L_F . Par des triangles similaires, le rayon illustré qui coupe la lentille en u et le plan du film en x' , coupe également le plan x en $u + (x' - u)\frac{F}{F'}$. Bien que le diagramme ne montre que le cas 2d impliquant x et u , ces dimensions et v partagent une relation identique. Par conséquent, si nous définissons $\alpha = F' / F$ comme la profondeur relative du plan du film.

$$E_{(\alpha, F)}(x', y') = \frac{1}{a^2, F^2} \iint L_F \left(U \left(1 - \frac{1}{\alpha} \right) + \frac{x'}{a'} V \right) \quad (1.6)$$

Cette équation formalise le cisaillement 2D du champ de lumière canonique qui résulte de la focalisation à différentes profondeurs. La combinaison des équations (2.1) et (2.2) conduit à l'équation finale pour la valeur de pixel (x, y) dans la photographie focalisée sur un morceau de film à la profondeur $F = \alpha F$ du plan de l'objectif [5].

8.2. La refocalisation numérique

Ng aborde la problématique de la mise au point du point de vue d'un champ de lumière. Puisqu'un champ de lumière possède l'information des directions des rayons lumineux d'une scène, il est alors possible de créer des images de la scène à différentes mises au point. Cette approche permet de ne plus se soucier de la mise au point de la caméra durant la capture.

Avec une caméra plénoptique pouvant capter directement un champ de lumière, il devient possible d'évaluer l'équation d'irradiance pour différentes valeurs de F , et ainsi créer des images à différentes valeurs de profondeur de champ [17].



Figure (1.15) : montre une série d'images avec différentes valeurs de profondeur de champ à partir d'un champ de lumière.

8.3. Algorithmes de synthèse d'images

L'ensemble idéal de rayons contribuant à un pixel dans la refocalisation numérique est l'ensemble de rayons qui convergent vers ce pixel dans une caméra virtuelle conventionnelle focalisée à la profondeur souhaitée, qui spécifie formellement dans son intégrale l'ensemble des rayons lumineux pour le pixel en position (x', y') :

$$E_{(\alpha F)}(x', y') = \frac{1}{\alpha^2 F^2} \iint L_F \left(u \left(1 - \frac{1}{\alpha} \right) + \frac{x'}{\alpha}, v \left(1 - \frac{1}{\alpha} \right) + \frac{y'}{\alpha}, u, v \right) du dv \quad (1.7)$$

En rappelons que dans cette équation le L_F est le champ lumineux paramétré par un plan $x y$ à une profondeur de F à partir du plan de la lentille $u v$, α est la profondeur du plan du film virtuel par rapport à F , et $E(\alpha F)$ est la photographie formée sur un film virtuel à une profondeur de (αF) Une façon d'évaluer cette intégrale est d'appliquer des techniques de quadrature numérique, telles que l'échantillonnage de l'intégrale pour différentes les valeurs de u et v et de les additionner. L'idée est de tracer le rayon $\left(u \left(1 - \frac{1}{\alpha} \right) + \frac{x'}{\alpha}, v \left(1 - \frac{1}{\alpha} \right) + \frac{y'}{\alpha}, u, v \right)$ à travers le réseau de microlentilles jusqu'au le capteur photo. Le point d'intersection est l'endroit où le rayon a déposé son énergie dans l'appareil photo pendant l'exposition, et la valeur de L_F est estimée à partir des valeurs du capteur photo à proximité de ce point [5].

Cependant, une méthode plus efficace est suggérée par la linéarité de l'intégrale par rapport au champ lumineux sous-jacent. Révèle l'observation importante selon laquelle le recentrage est conceptuellement une somme de versions dilatées et décalées des images de sous-ouverture sur toute l'ouverture $u v$. Ce point est clarifié en définissant explicitement l'image de sous-ouverture à la position de l'objectif (u, v) dans le champ lumineux L_F . Nous représentons cette image de sous-ouverture par la fonction 2d, telle que le pixel à la position (x, y) dans l'image de sous-ouverture soit donné par

$$L_F^{(u,v)}(x, y).$$

Avec cette notation, nous pouvons réécrire l'équation :

$$E_{(\alpha F)}(x', y') = \frac{1}{\alpha^2 F^2} \iint L_F^{(u,v)} \left(U \left(1 - \frac{1}{\alpha} \right) + \frac{x'}{\alpha}, V \left(1 - \frac{1}{\alpha} \right) + \frac{y'}{\alpha} \right) dU dV \quad (1.8)$$

Où $L_F^{(U,v)} \left(U \left(1 - \frac{1}{a} \right) + \frac{x'}{a}, v \left(1 - \frac{1}{a} \right) + \frac{y'}{a} \right)$ est une image de sous-ouverture $L_F^{(U,v)}$, dilatée par un facteur α et décalé d'un facteur $U \left(1 - \frac{1}{a} \right), v \left(1 - \frac{1}{a} \right)$. En d'autres termes, la refocalisation numérique peut être mise en œuvre en décalant et en ajoutant les images de sous-ouverture du champ lumineux [5].

9. Conclusion

Ce chapitre a décrit d'abord le paradigme de l'imagerie conventionnel et présente ses limites. Ensuite, nous avons présenté la nouvelle technologie d'imagerie à champ lumineux qui permet de capturer toutes les informations lumineuses envoyé par une scène. Nous avons expliqué le concept de la fonction plénoptique et le paramétrage du champ lumineux puis L'acquisition du champ lumineux qui peut être effectuée par de nombreuses techniques. On peut donc dire que dans une caméra plénoptique, les photos sont générées après l'acquisition du champ lumineux, dans ce cas l'utilisateur peut les modifier avant les afficher (changer de focus, faire une rotation ...). Nous présentons aussi en détail comment on peut changer la focalisation dans l'imagerie de champ lumineux. Enfin nous donnons un aperçu des algorithmes de synthèse d'image.

Chapitre II

Techniques d'anti-spoofing et présentation d'attaque

1. Introduction

Les systèmes de reconnaissance faciale sont vulnérables à différents types de présentation d'attaques. L'objectif principal d'un système PAD (présentation de détection d'attaque) facial est de faire automatiquement la distinction entre une présentation normale, un visage réel, et un visage usurpé à l'aide d'un PAI (présentation de l'Instrument d'attaque) au capteur d'imagerie, en essayant d'imiter les traits biométriques d'un visage réel [6].

Dans ce chapitre on a présenté la détection de vivacité de visage et les techniques d'anti-spoofing qui peuvent être globalement classés en deux types, le 1^{er} type basé sur le matériel et le 2^{ème} type basé sur un logiciel ainsi que les différents types d'attaque.

2. Détection de la vivacité du visage

La détection de la vivacité du visage est un processus permettant de déterminer si un visage détecté est réel ou non avant qu'un système de reconnaissance faciale identifie le visage, comme illustré sur la figure (2.1). Cela empêche le système de reconnaissance faciale de prendre une mauvaise décision. Il existe plusieurs types d'usurpation des visages, tels que des photos imprimées en 2D, les vidéos, des tablettes à hautes définition (HD) et les masques 3D etc. Parmi ces attaques, les photos 2D sont largement utilisées car elles sont faciles et peu coûteuses à obtenir [13].

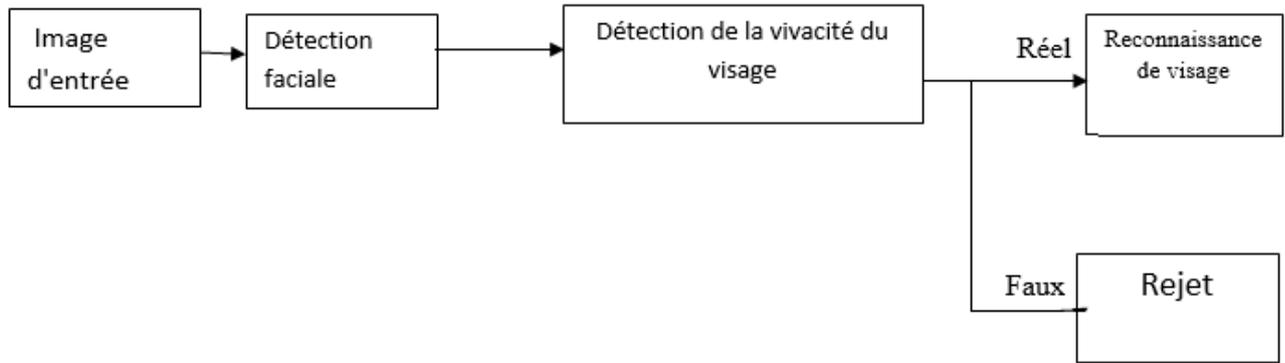


Figure 2.1. Système de reconnaissance faciale comprenant la détection de la vivacité du visage.

3. Les Techniques d'anti-spoofing

Les systèmes de reconnaissance faciale sont vulnérables à divers types d'attaques. Cela exige la nécessité de détecter et d'atténuer ces attaques afin d'améliorer à la fois la sécurité et la fiabilité des systèmes de reconnaissance faciale.

Une méthode PAD peut être définie comme une détermination automatique de la présence d'attaque. Le PAD est également appelé une contre-mesure ou une technique d'anti-spoofing. Dans la plupart des travaux existants, le PAD est également appelé détection de vivacité ; cependant, la détection de la vivacité est définie comme la mesure et l'analyse de caractéristiques anatomiques ou de réactions involontaires ou volontaires, afin de déterminer si un échantillon biométrique est capturé à partir d'un sujet vivant présent au point de capture. Suivant cette définition du terme normalisée, la détection de la vivacité peut être considérée comme un sous-ensemble de PAD mais pas comme un synonyme de PAD lui-même [6].

La figure (2.2) montre la classification des algorithmes de PAD des visages existants, qui peuvent être globalement classés en deux types, le 1^{er} type basé sur le matériel et le 2^{ème} type basé sur un logiciel.

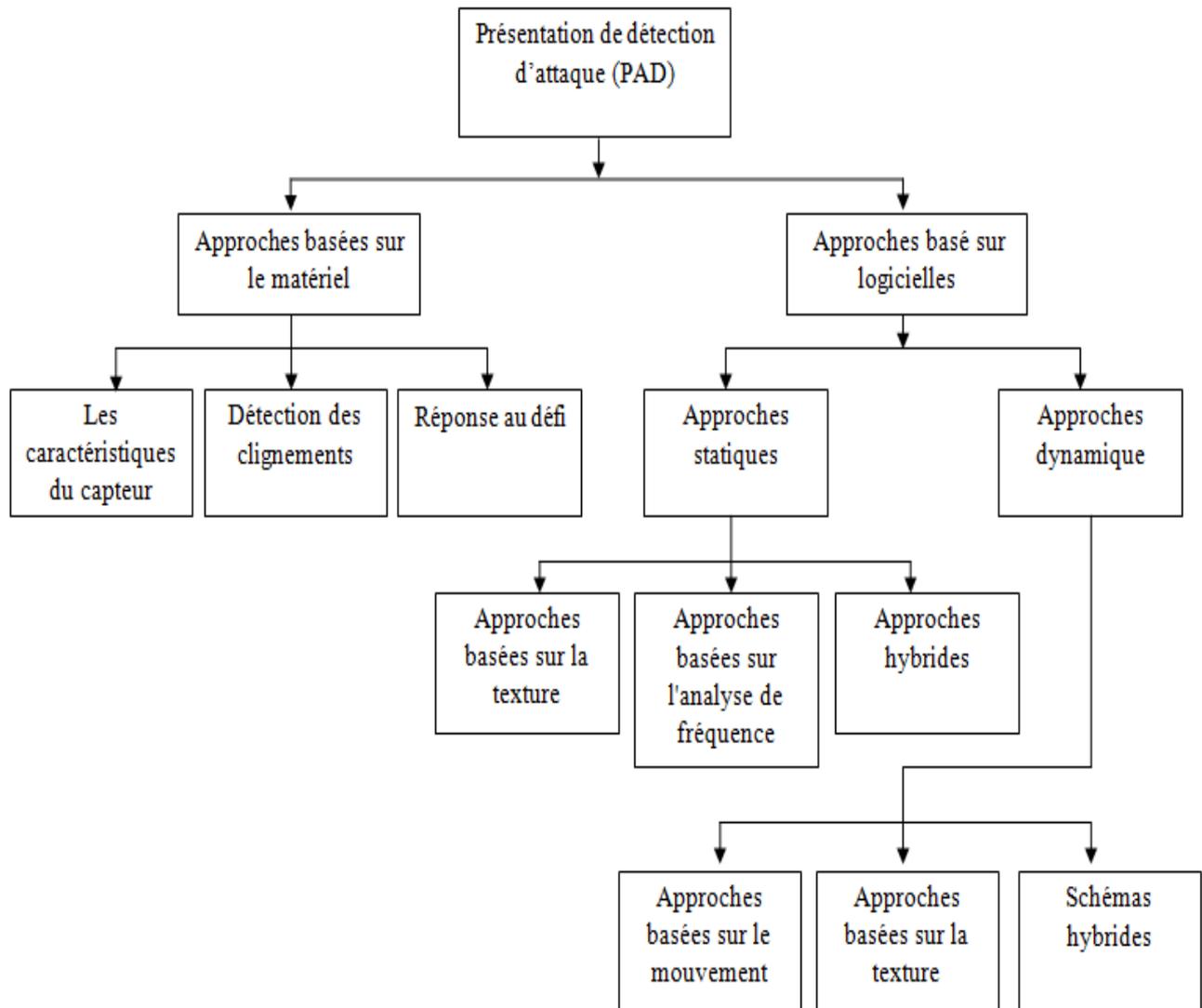


Figure (2.2) : la classification des algorithmes de PAD.

3.1. Les approches basées sur le matériel

Les approches basées sur le matériel explorent les caractéristiques du visage humain à l'aide des composants matériels supplémentaires qui fonctionnent en association avec le capteur de reconnaissance faciale. Ces approches peuvent également nécessiter une interaction avec le matériel ou un capteur de visage (comme le clignotement des yeux), qui utilisera également un logiciel interne pour traiter les données de visage capturées. Les approches basées sur le matériel peuvent être globalement classées en trois types : les caractéristiques du capteur, la détection des clignotements et la réponse à la stimulation, qui sont toutes décrites en détail comme suit [6].

3.1.1. Caractéristiques du capteur

Les techniques développées dans cette approche sont basées sur l'exploration des caractéristiques de la caméra (du capteur) utilisée pour capturer l'image du visage. Les caractéristiques du capteur exploré dépendent du type de capteur utilisé pour capturer les données du visage, par exemple, mesurer la variation de la mise au point avec une caméra à champ lumineux (LFC) ou la mesure de la réflexion d'un capteur facial proche-infrarouge / thermique / multi-spectral ou la mesure de la réflexion dans un scan 3D [6].

3.1.2. Détection des clignements

La détection des clignements est une mesure de vivacité largement utilisée pour atténuer la présentation d'attaques contre les systèmes de reconnaissance faciale. L'idée derrière la détection des clignements est de suivre en continu l'action spontanée des clignements oculaires qui sont exécutés inconsciemment. La détection des yeux-clignements peut être effectuée soit à l'aide d'un matériel dédié, soit d'une technique logicielle [6].

3.1.3. Réponse au défi

L'idée principale de la détection de présentation d'attaque basée sur un défi-réponse est de fournir une interface utilisateur séparée dans laquelle la réponse à un défi est enregistrée et traitée pour identifier une présentation d'un visage réel, par exemple, en suivant le regard de l'utilisateur vers un stimulus prédéfini [6].

3.2. Les approches basé sur logicielles

Les approches logicielles impliquent un algorithme qui peut déterminer si un échantillon de visage capturé provient d'une présentation d'attaque ou d'une présentation d'un visage réel (également appelée présentation réelle ou en direct). Il a été démontré que ce type de schéma PAD a une grande précision et un coût relativement faible. De plus, ces schémas ne nécessitent pas la coopération de l'utilisateur et évitent également le besoin de matériel spécialisé. Les méthodes existantes dans cette famille peuvent être divisées en deux types principaux : (1) les méthodes statiques et (2) les méthodes dynamiques [6].

3.2.1. Approches statiques

Les approches statiques sont conçues pour travailler sur une seule image sans avoir besoin d'informations temporelles. Généralement, ces approches sont connues pour leurs bonnes performances, leur faible calcul et leur faible coût. Les approches statiques de pointe disponibles peuvent être divisées en trois groupes principaux selon la nature de l'algorithme, à savoir des approches basées sur la texture, des approches basées sur la fréquence et des approches hybrides [6].

3.2.1.1. Approches basées sur la texture

Les approches basées sur la texture sont basées sur l'analyse des motifs micro texturaux de l'échantillon d'image de visage. Ce type d'approche est très efficace dans la détection des photos et de l'affichage des attaques, car cette méthode peut distinguer efficacement les caractéristiques des attaques comme la présence de pigments (due à des défauts d'impression), la réflexion spéculaire, et de l'ombre (en raison d'une attaque d'affichage) [6].

3.2.1.2. Approches basées sur la fréquence

Le deuxième type de méthode statique comprend des techniques basées sur l'analyse de fréquence pour la PAD de visage. Les premiers travaux dans cette catégorie étaient basés sur l'analyse du spectre de Fourier et a été utilisé avec succès pour détecter l'attaque de la photo du visage [6].

3.2.1.3. Les approches hybrides

Le troisième type d'approche statique comprend des schémas hybrides qui combinent plus d'un attribut à l'utilisation des informations d'identité du client, la caractérisation de la propriété de défocalisation de l'image de visage capturée, qui combine des informations temps-fréquence avec un descripteur de texture, qui allie forme et texture, ou l'utilisation d'informations contextuelles [6].

3.2.2 Approches dynamique

L'idée principale d'une approche dynamique est d'exploiter les informations temporelles de la relecture vidéo présentée au capteur de reconnaissance faciale. Les approches dynamiques ont tendance à modéliser ces informations temporelles en exploitant le mouvement relatif à travers les images vidéo [6].

Par conséquent, une approche dynamique nécessitera plus de temps et plus d'efforts de calcul par rapport à une approche statique. Les approches dynamiques peuvent être globalement classées en trois types : approches basées sur le mouvement, approches basées sur la texture et schémas hybrides [6].

3.2.2.1. Approches basées sur le mouvement

Les méthodes basées sur le mouvement capturent le mouvement inconscient présenté par les muscles du visage en raison du mouvement de la tête. Le mouvement capturé est notamment dû aux mouvements de la tête, la bouche ou les yeux [6].

3.2.2.2. Approches basées sur la texture

Le deuxième type de schéma basé sur le mouvement explore le changement de texture dynamique à travers la vidéo capturée. Les premiers travaux dans cette direction sont basés sur des modèles binaires locaux de trois plans orthogonaux (LBP-TOP) et a démontré une performance raisonnable sur la base de données d'attaque de relecture [6].

3.2.2.3. Schémas hybrides

La troisième approche explore à la fois les fonctionnalités basées sur le mouvement et la texture pour obtenir des performances très précises d'identification de la présentation des attaques de visage. L'utilisation de plusieurs repères scéniques a été introduite en [2012], explorant à la fois les deux approches présidant pour identifier une attaque de relecture vidéo sur les systèmes de reconnaissance faciale [6].

4. Les types d'attaques

Comme tout élément ayant une utilité de sécurisation, les systèmes de reconnaissance faciale sont sujets à des attaques visant à les tromper en exploitant des failles. Trois principaux types d'attaques sont alors connus dans ce domaine pour tromper les systèmes de reconnaissance faciale sont attaque photo, attaque vidéo et attaque avec un masque [18].

4.1. Attaque photo

Le premier type d'attaque est l'attaque par photo. En effet, montrer une photo au système semble être une bonne méthode pour tromper la caméra. L'efficacité de ce type d'attaque a été aidée par l'arrivée d'imprimantes ayant une résolution élevée, aussi nommée haute définition. Ces imprimantes permettent en effet d'imprimer une image très qualitative, trompant aisément la vigilance des systèmes.

Une autre possibilité consiste à utiliser des écrans à haute définition disponibles dans le commerce (écrans, Smartphones) pour présenter une photo très détaillée et ainsi tromper le système.

Ce type d'attaque est devenu encore plus simple à effectuer avec l'apparition des réseaux sociaux sur internet et les très nombreuses photos postées par les utilisateurs [18].

4.2 Attaque vidéo

Ce deuxième type d'attaque consiste à présenter une vidéo à l'objectif de la caméra afin de tromper le système de reconnaissance faciale.

Des logiciels disponibles sur internet, appelés caméras virtuelles, tels que la Virtual Caméra ou Manycam permettent en effet de tromper le système de reconnaissance en lui présentant une vidéo comme si celle-ci provenait de sa propre caméra. Il est alors possible de présenter la vidéo d'une personne au système de reconnaissance afin de tromper celui-ci.

Ce type d'attaque s'est répandu avec l'arrivée des reconnaissances 3D, car il est très compliqué de tromper une reconnaissance d'un modèle 3D avec une simple photo, cette dernière présentant toujours la même posture d'un individu.

En revanche, la vidéo présente du mouvement et permet d'obtenir plusieurs postures du même individu, le système de reconnaissance 3D peut donc modéliser un modèle réel et être trompé [18].

4.3 Attaque avec un masque

La récente mise à disposition au grand public de l'impression tridimensionnelle crée également une faille pour les systèmes de reconnaissance faciale. En effet, cela permet de créer un masque correspondant au visage de l'utilisateur. Il est alors facile d'imaginer l'utilisation d'un masque présentant un individu pour tromper un système de reconnaissance faciale [18].

5. Conclusion

Dans ce chapitre, nous avons présenté les techniques d'anti-spoofing en montrons la classification des algorithmes de PAD des visages existants. Nous avons également vu les types d'usurpation des visages existants.

Chapitre III

La détection d'attaque par la caméra plénoptique

1. Introduction

La détection de vivacité est une technique de plus en plus populaire pour la vérification d'identification des utilisateurs. Cependant, cette technique est susceptible de provoquer des attaques. Par conséquent, un moyen fiable de détecter les attaques malveillantes est essentiel à la robustesse du système de reconnaissance faciale [9].

Dans ce chapitre, nous présentons une approche d'utilisation de la caméra plénoptique pour défendre les attaques des visages usurpés, comme des photos faciales 2D_s imprimées et des images de tablette haute définition. Dans le procédé proposé, trois types de caractéristiques extraites d'une paire d'image recentrées sont extraites pour discriminer les faux visages et les vrais visages. Pour vérifier les performances, nous utilisons des bases de données photographiques en champ lumineux et menons des expériences.

2. Méthodologie proposée

L'anti-usurpation de visage consiste à déterminer si le visage capturé est un vrai visage ou un faux visage. Cette technique est souvent associée à un système de reconnaissance d'authentification faciale, visant à vérifier si l'utilisateur est bien lui-même.

La détection de vivacité est utilisée comme un pré test du système de sécurité basé sur la reconnaissance de visage humain, qui vise à distinguer le faux visage et le vrai visage devant la caméra. Pour que le visage soit reconnu, le système prend des photos avec l'appareil photo à champ lumineux en une seule exposition, puis deux images spécifiques correspondant à différentes profondeurs de mise au point peuvent être acquises. Sur la base de cette paire d'images, les variations de mise au point sont calculées en considérant l'entropie d'intensité des informations, la fonction de mise au point (focus) et la fonction d'histogramme de puissance. Ensuite, les faux visages peuvent être détectés en fonction des variations de mise au point (deux images à différente focus $D1$, $D2$) [9].

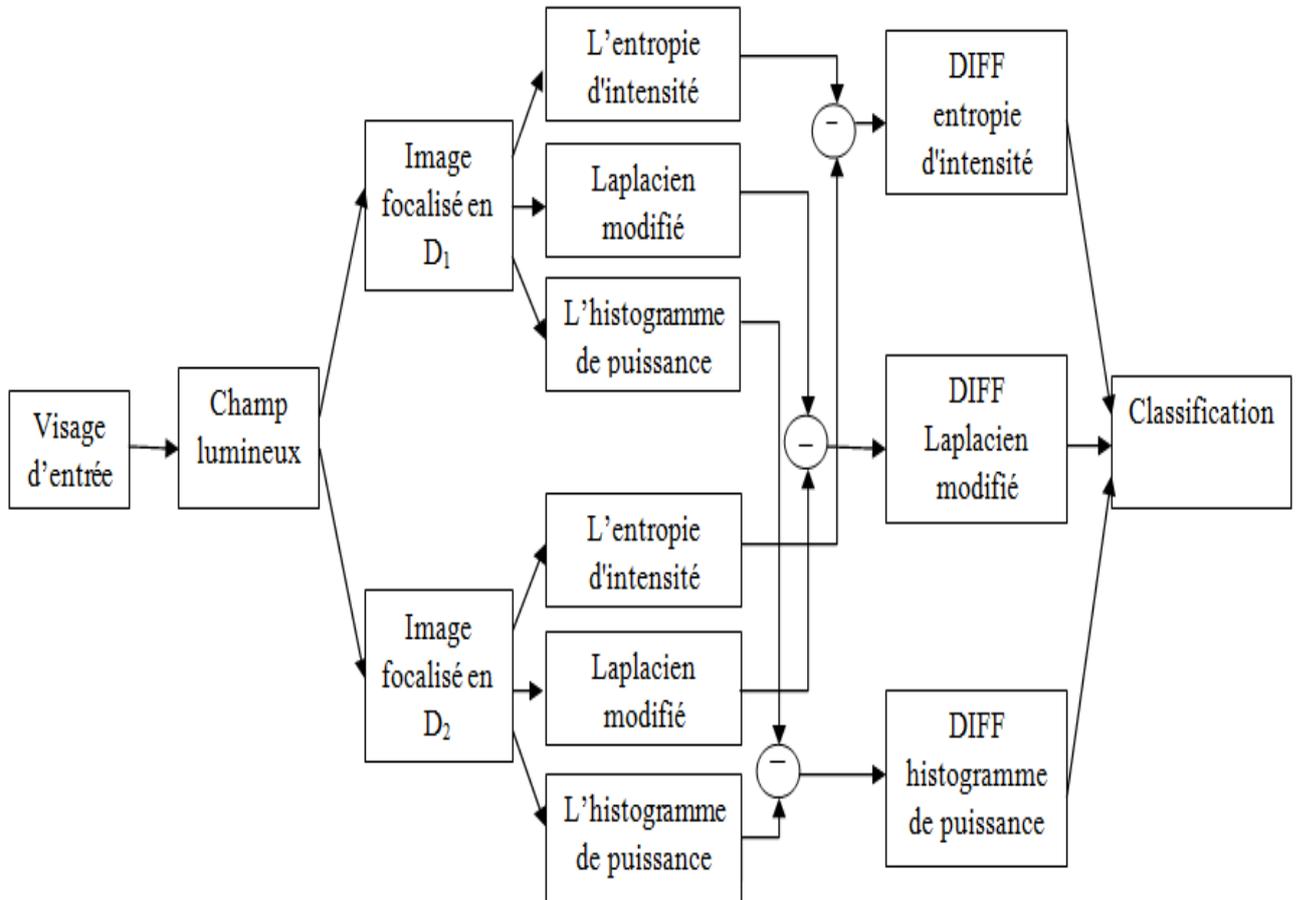


Figure (3.1) : Organigramme de la méthode de détection de la vivacité du visage proposée.

L'organigramme de la méthode de détection de la vivacité du visage proposée est illustré à la figure (3.1). Les détails techniques de notre méthode sont présentés comme suit :

2.1 Entropie d'intensité

L'entropie d'intensité (IE) est une forme statistique d'entités qui représentent la quantité d'informations contenues dans l'entité agrégée de la distribution d'intensité dans une image. Soit p_i représente la proportion des pixels avec la valeur de gris de i dans une image en niveaux de gris I , alors l'entropie d'intensité unaire de I est définie comme :

$$IE = \sum_{i=0}^{255} p_i \log(p_i) \tag{3.1}$$

Désigne les entropies d'intensité de deux images d'un visage avec des focus différents comme IE_1 et IE_2 respectivement. La différence entre les entropies d'intensité des deux images est utilisée comme fonction de détection de la vivacité, $DIFF_{EI} = EI_1 - EI_2$ [9].

2.2 Fonction de mise au point

Le laplacien modifié (LAPM) est également utilisé pour calculer la mesure de la mise au point. Plus précisément, le LAPM est présenté comme la somme des filtres laplacien. Pour une image $2D$, le processus d'extraction de cette fonctionnalité est présenté dans la figure (3.1) ci-dessous [14, 9].

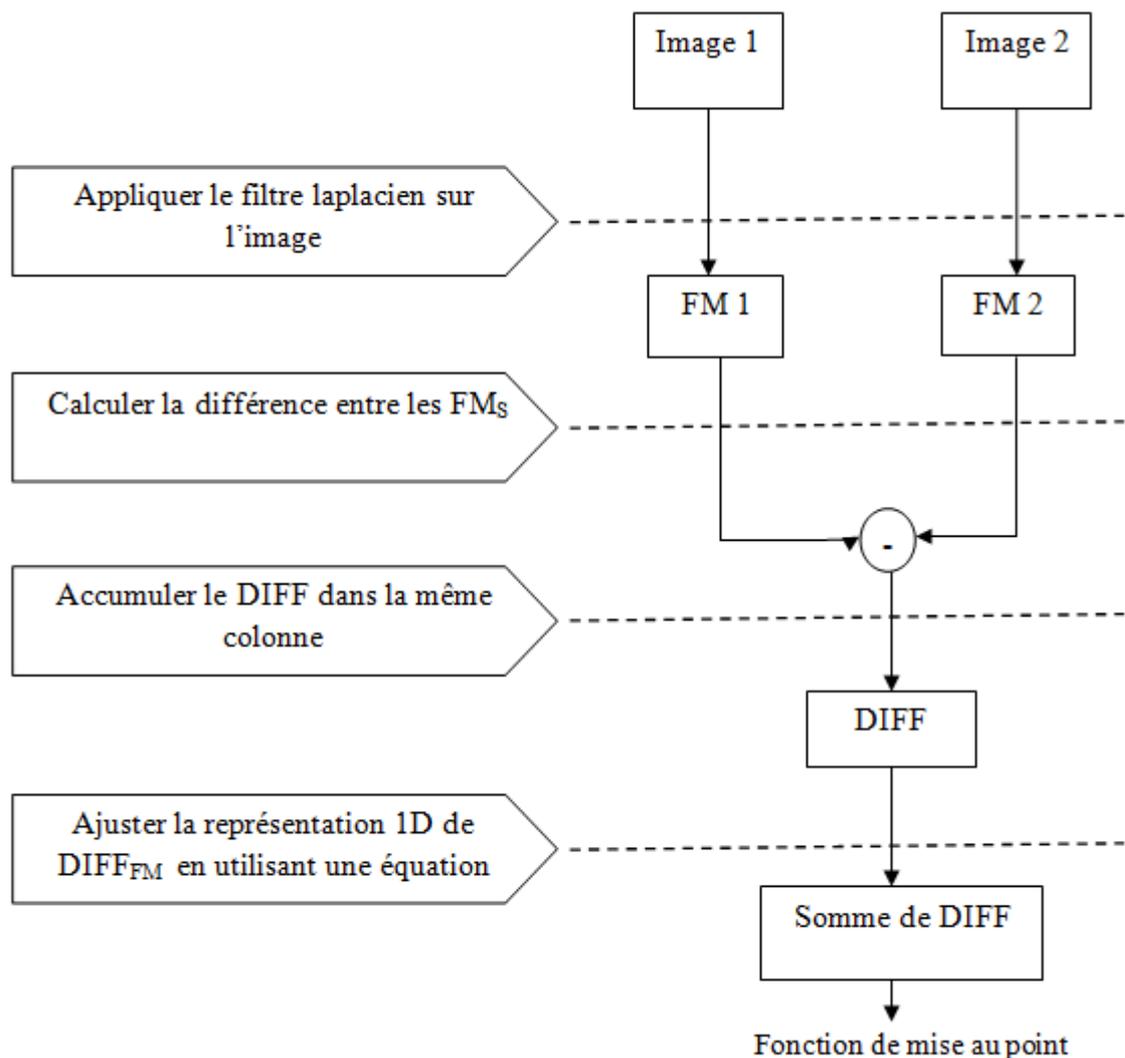


Figure (3.2) : Organigramme d'extraction de la fonction de mise au point.

Le LAPM peut être calculé par l'équation (3.2)

$$LAPM(i, j) = \sum_{(x,y) \in \Omega(i,j)} \Delta_L^2 I(x, y) \cdot t$$

$$\Delta_L^2 I(x, y) = \left| \frac{\partial^2 I}{\partial x^2} \right| + \left| \frac{\partial^2 I}{\partial y^2} \right| \quad (3.2)$$

$$t = \begin{cases} 1 & \text{LAPM}(x, y) \geq T \\ 0 & \text{LAPM}(x, y) < T \end{cases}$$

Où $\Omega(i, j)$ signifie le voisinage de (i, j) et T est un seuil empirique.

Étant donné pour que les régions de mise au point semblent plus nettes et correspondent à des valeurs de LAPM plus élevées que celle qui sont hors du plan focale, le LAPM est lié aux degrés de mise au point.

Les variations LAPM entre deux images recentrées peuvent être utilisées pour discriminer le vrai visage du faux visage. Pour une face d'entrée, notez LAPM1 et LAPM2 comme les mesures de mise au point de deux images se concentrant sur différentes profondeurs D1 et D2, respectivement. La fonction de mise au point finale peut être calculée selon les étapes suivantes :

- 1- Calculer la différence de FM, c'est-à-dire $DIFF_{FM} = LAPM_1 - LAPM_2$;
- 2- Accumuler tous les $DIFF_{FM}$ dans la même colonne, c'est-à-dire $DIFF_{FM}^p(i) = \sum_j DIFF_{FM}^p(i, j)$;
- 3- Adaptez la représentation à une seule dimension de $DIFF_{FM}$ en utilisant une équation quadratique, $y = ax^2 + bx + c$

Ensuite, les coefficients [a, b, c] sont exploités comme une fonction de classification.

2.3 Fonction d'histogramme de puissance

Les variations de mis au point peuvent également être détectées par informations de fréquence. Dans notre méthode, la fonction d'histogramme de puissance est utilisée. Le processus d'extraction de cette fonctionnalité est présenté dans la figure (3.2) ci-dessous [14, 9].

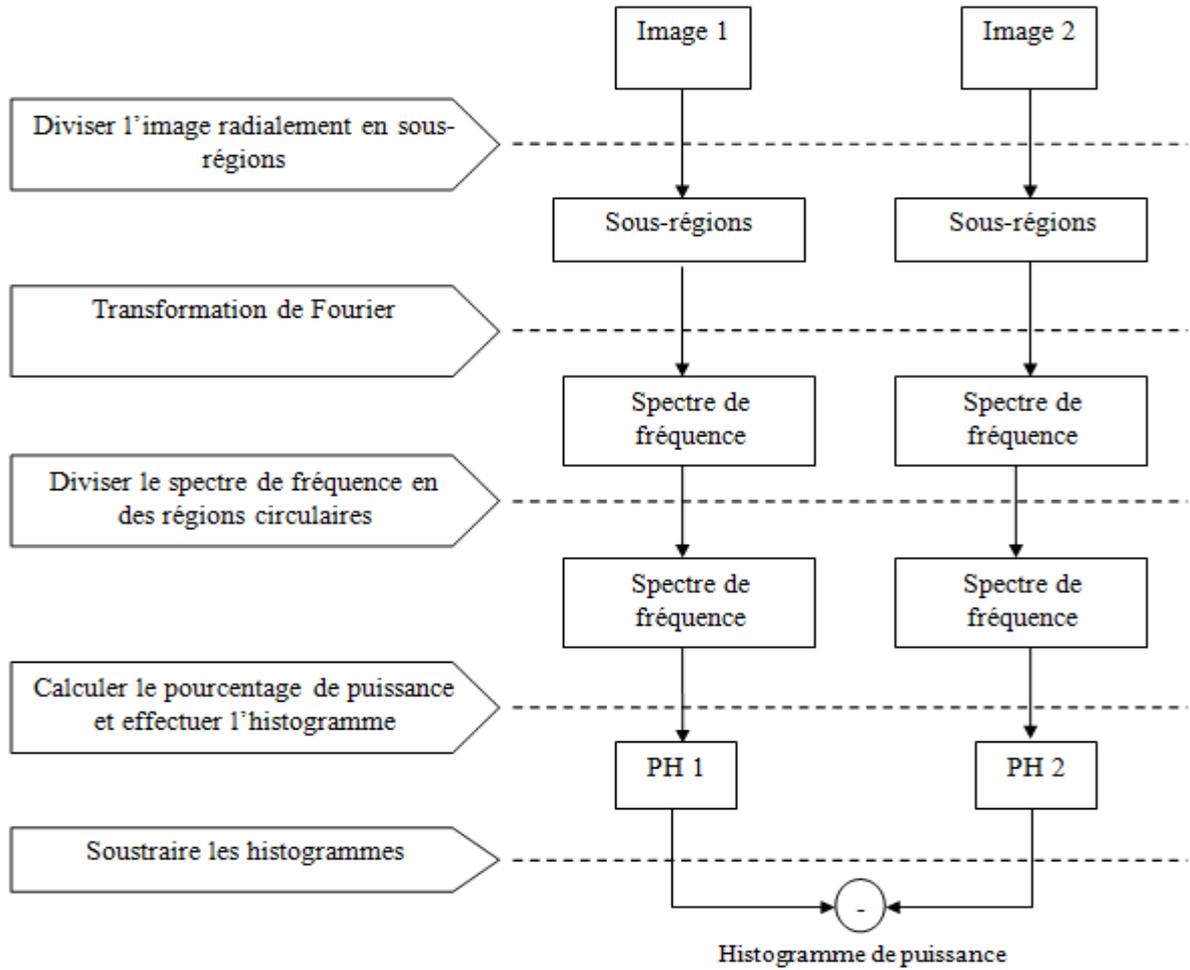


Figure (3.3) : Organigramme d'extraction des caractéristiques de l'histogramme de puissance.

- 1- Diviser I radialement en K sous-régions $\{S_k\}_{k=1}^K$;
- 2- Effectuer la FFT décalée au centre pour chaque sous-région ;
- 3- Pour chaque sous-région $S_k, k = 1, k$, diviser le spectre de fréquences en C cercles $\{C_k^c\}_{c=1}^{c=C}$ en le superposant ;
- 4- Calculer le pourcentage de puissance dans chaque région circulaire S_k par :

$$\text{Perc}_k^c = \frac{\sum_{(u,v) \in C_k^c} P(u,v)}{\sum_{(u,v) \in S_k} P(u,v)} \text{ OÙ } P(u,v) = \text{real}(u,v)^2 + \text{imag}(u,v)^2 \text{ avec } \text{real}(u,v) \text{ et } \text{imag}(u,v) \text{ comme parties réelle et imaginaire de la composante fréquentielle, respectivement.}$$

- 5- Concaténer les histogrammes K par $PH = [PH_1, \dots, PH_K]$, où $PH_k = [\text{Perc}_k^1, \dots, \text{Perc}_k^C]$

3. Classificateur SVM

Les SVM constituent une classe d'algorithmes basée sur le principe de minimisation du « risque structurel » décrit par la théorie de l'apprentissage statistique de Vapnik et Chervonenkis qui utilise la séparation linéaire. Cela consiste à séparer par hyperplan des individus représentés dans un espace de dimension égale au nombre de caractéristiques, les individus étant alors séparés en deux classes. Cela est possible quand les données à classer sont linéairement séparables. Dans le cas contraire, les données seront projetées sur un espace de plus grande dimension afin qu'elles deviennent linéairement séparables [12].

3.1. Classification binaire par hyperplan

Considérons maintenant I points $f(x) = \text{sgn}(\langle w, x \rangle + b)$, avec $i = 1, \dots, I$ et $y_i \in \{\pm 1\}$

Classons ces points en utilisant une famille de fonctions linéaires définies par $\langle w, x \rangle + b = 0$ avec $w \in \mathfrak{R}^N$ et $b \in \mathfrak{R}$ de telle sorte que la fonction de décision concernant l'appartenance d'un point à l'une des deux classes soit donnée par :

$$f(x) = \text{sgn}(\langle w, x \rangle + b) \quad (3.3)$$

3.2. Cas de données linéairement séparables

Nous allons construire l'hyperplan H d'équation : $\langle w, x \rangle + b = 0$ qui sépare au mieux les deux classes et se trouvant à mi-distance des deux hyper plans H_1 et H_2 parallèles à H , d'équations respectives :

$$H_1 : \langle w, x \rangle + b = +1$$

$$H_2 : \langle w, x \rangle + b = -1$$

Telle que les deux conditions suivantes soient respectées :

Condition 1

- il n'y a aucun point qui se situe entre H_1 et H_2 .

Cette contrainte se traduit par les inégalités :

$$\langle w, x \rangle + b \leq +1 \text{ pour } y_i = +1$$

Et

$$\langle w, x \rangle + b \leq -1 \text{ Pour } y_i = -1$$

Ces deux inégalités peuvent être combinées en une seule :

$$y_i(\langle w, x \rangle + b) \geq +1$$

Condition 2

- La distance ou la marge entre H_1 et H_2 est maximale.

Dans ce cas, la distance entre H_1 et H_2 est donné par : $M = \frac{2}{\|w\|}$

Maximiser M revient à minimiser $\|w\|$ ou à minimiser $\|w\|^2$ avec :

$\|w\|^2 = w^T w$ (Carré de la norme euclidienne du vecteur).

3.3. Cas des données non-linéairement séparables [Mou 05]

En pratique, il est assez rare d'avoir des données linéairement séparables. Afin de traiter également des données bruitées ou non linéairement séparable, les SVM ont été généralisées grâce à deux outils : la marge souple (soft margin) et les fonctions noyau (kernel fonctions).

Le principe de la marge souple est d'autoriser des erreurs de classification.

Le nouveau problème de séparation optimale est reformulé comme suit :

L'hyperplan optimal séparant les deux classes est celui qui sépare les données avec du minimum d'erreurs, et satisfait donc les deux conditions suivantes :

Condition 1

- la distance entre les vecteurs bien classés et l'hyperplan doit être maximale.

Condition 2

- la distance entre les vecteurs mal classés et l'hyperplan doit être minimale.

4. Conclusion :

Notre travail repose sur une nouvelle approche d'utilisation de la caméra plénoptique pour la détection de la vivacité du visage. Sans prendre les séquences d'images, nous pouvons capturer le changement de l'apparence du visage lors de la refocalisation, qui peut être utilisée pour détecter les faux visages et les vrais visages.

Chapitre IV

Implémentation et résultats

1. Introduction

La détection de la vivacité du visage est un processus permettant de déterminer si un visage détecté est réel ou non avant qu'un système de reconnaissance faciale identifie le visage [13]. Dans ce chapitre, la mise en œuvre de notre système est illustrée en montrant la méthode utilisée et les résultats obtenus. Pour que le visage soit reconnu, la photo prise par la caméra plénoptique comme un visage d'entrée on s'y applique des algorithmes pour détecter si le visage détecté est un visage réel ou usurpé.

Notre objectif dans cette section est de trouver une solution analytique de la refocalisations numérique calculée à partir du champ lumineux pris par la caméra plénoptique, pour avoir si la l'image à été prise est-elle liée à la photographie exacte ou non.

2. Base de donné

La base de données *IST Lenslet Light Field Face Spoofing* est une base de données qui contient des images de champ lumineux pour tester les systèmes de la présentation de détection d'attaque facial. Elle est composée des éléments suivants : Images de champ de lumière brute *Light Field Raw (LFR)*, Images rendues 2D pour la vue centrale du champ lumineux et une Cartes de profondeur pour chaque image rendue en 2D [7].

L'acquisition IST LLFFSD a été réalisée à l'intérieur, en utilisant une caméra plénoptique, le Lytro ILLUM, pour capturer des images de visages authentiques et des tentatives d'attaque. La base de données comprend les données de cinquante volontaires, Chaque volontaire a participé à deux sessions d'acquisition distinctes. Une image rendue en vue centrale 2D, d'une résolution de 2022×1404 pixels, correspondant à chaque véritable champ lumineux, permet de générer les six types d'attaques considérés dans la base de données proposée : Attaque du papier imprimé, Attaque du papier imprimé enveloppé, Attaque d'un ordinateur portable, Attaque de tablette, Attaque mobile 1, Attaque mobile 2 [7].

Les données importées sont des fichiers LFP (*Light Field Picture*). LFP est un format d'image qui contient le champ lumineux capturé. Chaque fichier LFP enregistre 11 *méga rays*, mais il est impossible de visualiser et d'accéder directement au fichier car il s'agit de données 4D [13]. Afin de visualiser le champ lumineux 4D, nous projetons le champ lumineux sur une image 2D. Nous utilisons le programme lecteur LFP et la boîte d'outils (*Light Field Toolbox for MATLAB*). En exécutant le programme lecture LFP pour obtenir des images brutes et leurs métadonnées associées. Afin d'analyser ces données dans MATLAB, nous les décodons à l'aide de *Light Field Toolbox*, pour extraire le champ lumineux (LF).

3. Expérimentation et discussion

Notre objectif dans cette section est de réaliser un système anti attaque en utilisant la caméra plénoptique qui sert à nous fournir deux images à différents plans focaux. Sur la base de cette paire d'images, les variations de mise au point sont calculées en considérant l'entropie d'intensité des informations, la fonction de mise au point (focus) et la fonction d'histogramme de puissance. Ensuite, les faux visages peuvent être détectés en fonction des variations de mise au point (deux images à différente focus $D1, D2$).

3.1. Evaluation quantitative

Le nombre de photographies en champ lumineux est indiqué dans le tableau (4.1). Le nombre total de sujets est de cinquante. Les photographies en champ lumineux sont décodées avec *Light Field Toolbox for Matlab*.

Contexte	Vrais visages	Faux visages		
		Impression normale	Impression déformée	HD Tablette
	100	100	100	100

Tableau (4.1) : Bases de données de champ lumineux.

Nous évaluons les performances de notre méthode proposée avec nos propres bases de données. Les bases de données sont classées au hasard en trois groupes : ensembles d'entraînement, de développement et de test.

- Ensemble d'entraînement (30%): à utiliser pour la formation du classificateur.
- Ensemble de développement (30%): à utiliser pour estimer le seuil du classificateur.
- Ensemble de test (40%): à utiliser pour évaluer la performance.

Trente pour cent des sujets sont utilisés pour la formation et le développement, et quarante pour cent des sujets sont utilisés pour les tests. Trois groupes sont disjoints. Autrement dit, si les images du sujet A sont utilisées pour la formation, elles ne peuvent pas être utilisées pour le développement ou les tests. Pour les résultats numériques, les mesures suivantes sont utilisées. Les mesures sont exprimées en termes dans le tableau (4.2) [13].

- Taux de fausse acceptation (FAR): la proportion de fausses images mal classées comme réelles.

$$FAR = \frac{FP}{FP+TN} \quad (4.1)$$

- Taux de faux rejets (FRR): la proportion d'images réelles mal classées comme fausses.

$$FRR = \frac{FN}{TP+FN} \quad (4.2)$$

- Taux d'erreur total (TER): somme de FAR et FRR.

$$TER = FAR + FRR \quad (4.3)$$

- Demi taux d'erreur total (HTER): moitié du TER.

$$HTER = TER / 2 \quad (4.4)$$

- Précision : le rapport entre le nombre d'images de test correctement classées et le nombre total d'images de test.

$$\text{Précision} = \frac{TP+TN}{TP+TN+FP+TN} \quad (4.5)$$

	Objet de la cible (visages réels)	Objet de valeur aberrante (fausses faces)
classifié comme cible (visages réels)	Vrai positif (TP)	Faux positif (FP)
classifié comme aberrant (fausses faces)	Faux négatif (FN)	Vrai négatif (TN)

Tableau (4.2) : Mesurer les fausses erreurs de détection.

3.2 La laplacien modifiés de la paire d'image

Les variations LAPM entre deux images focalisées peuvent être utilisées pour discriminer le vrai visage du faux visage. Les résultats obtenus comme suite:



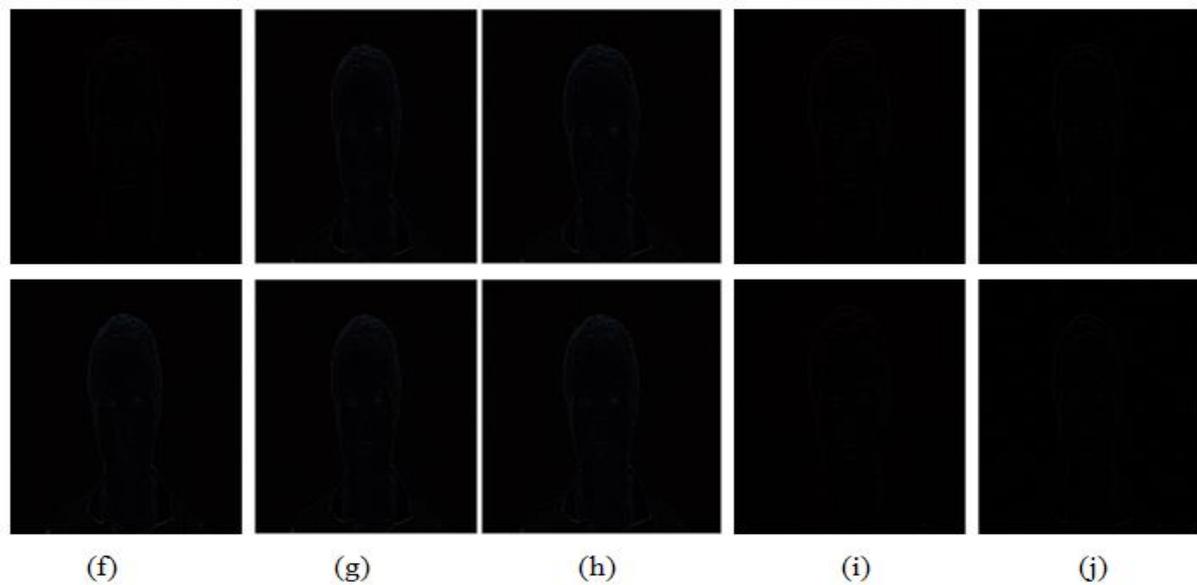


Figure (4.1) : Paire Images ainsi que les LAPM correspondantes. (a), (b), (c), (d) et (e) sont les images d'un visage réel, d'une visage d'écran de l'ordinateur portable, visage HD tablette, visage mobile 1 et visage imprimé, respectivement, focalisées sur plans de profondeur différents. (f), (g), (h), (i) et (j) sont des les LAPM correspondant à (a), (b), (c), (d) et (e) respectivement.

Les images recentrées ainsi que les LAPM correspondants du visage réel, du visage d'écran de l'ordinateur portable, du visage imprimé sur le papier, visage HD sur Tablette et du visage sur l'écran de mobile respectivement, sont représentés. Pour les LAPM, les pixels clairs représentent des valeurs élevées de LAPM. À travers les yeux humains, il est difficile d'examiner la différence entre deux images sous une mise au point différente, indépendamment du vrai visage ou du faux visage. Cependant, pour la paire d'images de visage réel, leurs LAPM sont évidemment discriminants les uns des autres, tandis que la paire de LAPM de fauxvisage a peu de différences.

4. Analyse des résultats

Les résultats expérimentaux sont présentés dans le tableau (4.3). Concernent notre méthode, différentes fonctionnalités semblent plus efficaces pour déterminer les différents types d'attaques. Par exemple, l'entropie d'intensité obtient un taux de précision élevé pour détecter les visages mobile, mais obtient un taux de précision relativement faible pour détecter les visages d'ordinateur portable et d'imprimé ainsi que la tablette HD. Au contraire, la fonction de mise au point a un meilleur effet de reconnaissance sur les visages de papier imprimé, les visages de la tablette HD, et d'Ordinateur portable que sur les visages mobile. L'histogramme de puissance ne semble pas préjugé contre les schémas d'attaque, mais ne peut pas atteindre les meilleures performances pour tout type de mode d'attaque.

La méthode utilisant plusieurs fonctionnalités obtient généralement de meilleurs résultats que celle utilisant un seul type de fonctionnalité, mais ce n'est pas toujours.

	Ordinateur portable	Tablette	Mobile1	Papier imprimé
L'entropie d'intensité	72.5	69.0	84.1	75.3
Fonction de mis au point	87.8	85.3	76.4	92.6
Fonction d'histogramme de puissance	88.8	91.3	94.1	96.2

Tableau (4.3) : Le taux de précision (%) des quatre fonctionnalités dans différentes attaques d'usurpation d'identité.

Année	Type	Ordinateur portable		Tablette		Mobile1		Papier imprimé	
		HTER	% STD	HTER	%STD	HTER	%STD	HTER	%STD
[7] 2014	LF	10.12	0.028	12.39	0.027	12.79	0.033	12.91	0.044
[7] 2015	LF	19.78	0.072	26.36	0.056	29.98	0.042	32.43	0.058
[7] 2016	LF	11.00	0.034	10.77	0.035	8.12	0.034	7.27	0.030
[7] 2017	LF	1.83	0.012	3.70	0.010	1.50	0.008	1.67	0.009
La solution proposée		1.23	0.010	3.24	0.015	0.91	0.006	1.10	0.009

Tableau (4.4) : HTER et standard déviation de la méthode proposée et comparaisant avec les références.

Par rapport aux solutions existantes de détection d'attaque de visage par champ lumineux [2017, 2016, 2015, 2014], la solution proposée atteint toujours de meilleures performances. Il est intéressant de noter que les expériences réalisées on [2016, 2015] ont été menées sur des images de visage capturées dans des environnements moins contrôlés, à différentes distances, qui ne sont pas les scénarios d'attaque les plus plausibles. L'exploration de la variation de mise au point pour ces images apporte plus d'informations par rapport aux images acquises dans des conditions contrôlées, qui sont presque entièrement au point. Par rapport à [2017], les résultats montrent qu'utiliser uniquement les informations angulaires de champ lumineux fournit la meilleure description de l'anti-usurpation de visage.

Conclusion

Dans ce travail, quatre types de fonctionnalités sont utilisés. Pour évaluer notre méthode proposée, nous avons collecté une base de données de champ lumineux avec des attaque du papier imprimé, attaque du papier imprimé, attaque d'un ordinateur portable, attaque de tablette, attaque mobile 1. Les expériences révèlent que différentes caractéristiques sont habiles à déterminer différents types d'attaques. En plus en comparant notre méthode à des autres méthodes présidant, nous montrons que la méthode que nous proposons présente de bonnes performances.

Conclusion Générale

Bien que les dispositifs de prise de vue plénoptiques soient maintenant fonctionnels et offerts au grand public, Le choix de l'appareil, du point de vue et du cadrage étant faits le seul réglage sur lequel peut ensuite jouer le photographe pour contrôler la profondeur de champ est l'ouverture du diaphragme. Les développements récents de la photographie computationnelle ont permis la variation de la mise au point optique d'une caméra plénoptique après l'exposition de l'image, également connue sous le nom de refocalisations. Il est maintenant possible de fixer la photo même après sa capture.

Depuis 2015, les gens ont déjà commencé à utiliser l'apprentissage en profondeur pour étudier l'anti-usurpation faciale. Cependant, en raison des échantillons de formation limités sur les ensembles de données ouverts, les performances d'utiliser l'apprentissage en profondeur ne peut aller au-delà de la méthode traditionnelle.

Dans un monde où la sécurité des individus est devenue un souci majeur, le besoin de se protéger augmente jour après jour. Pour cette raison on a présenté une méthode biométrique de reconnaissance d'identité faciale qui nous permettent de déterminer si un visage saisi est directement capturé à partir d'un humain ou non, en plus de l'automatisation des certaines applications telles que la télésurveillance, l'accès à des endroits sécurisés, etc. [12].

Pour vérifier les performances, nous utilisons une base de données photographique en champ lumineux et menons des expériences. Les résultats expérimentaux révèlent que les fonctionnalités utilisées peuvent atteindre une précision de l'anti-usurpation de visage remarquable sous différents types d'attaques d'usurpation faciale.

Bibliographie

- [1] E. H. Adelson and J. R. Bergen, The plenoptic function and the elements of early vision vol. 2: Vision and Modeling Group, Media Laboratory, Massachusetts Institute of ..., 1991.
- [2] M. Levoy, "Light fields and computational imaging," *Computer*, vol. 39, pp. 46-55, 2006.
- [3] C.-K. Liang, Y.-C. Shih, and H. H. Chen, "Light field analysis for modeling image formation," *IEEE Transactions on Image Processing*, vol. 20, pp. 446-460, 2010.
- [4] M. Liu, H. Fu, Y. Wei, Y. A. U. Rehman, L.-m. Po, and W. L. Lo, "Light field-based face liveness detection with convolutional neural networks," *Journal of Electronic Imaging*, vol. 28, p. 013003, 2019.
- [5] R. Ng, *Digital light field photography vol. 7: stanford university Stanford*, 2006.
- [6] R. Ramachandra and C. Busch, "Presentation attack detection methods for face recognition systems: A comprehensive survey," *ACM Computing Surveys (CSUR)*, vol. 50, pp. 1-37, 2017.
- [7] A. Sepas-Moghaddam, L. Malhadas, P. L. Correia, and F. Pereira, "Face spoofing detection using a light field imaging framework," *IET Biometrics*, vol. 7, pp. 39-48, 2017.
- [8] G. Wetzstein, "Computational plenoptic image acquisition and display," *University of British Columbia*, 2011.
- [9] X. Xie, Y. Gao, W.-S. Zheng, J. Lai, and J. Zhu, "One-snapshot face anti-spoofing using a light field camera," in *Chinese Conference on Biometric Recognition*,
- [10] F. Pérez Nava, A. Pérez Nava, M. Rodríguez Valido, and E. Magdaleno Castellò, "Plenoptic cameras," *Biologically Inspired Computer Vision: Fundamentals and Applications*, pp. 175-200, 2015.
- [11] L. Morillon, "Reconstruction du flou d'objectif à partir de données issues de caméras plénoptiques."
- [12] Mr. GHALI Ahmed. "Amélioration de la reconnaissance par le visage " Département d'informatique, UNIVERSITE DES SCIENCES ET DE LA TECHNOLOGIE D'ORAN MOHAMED BOUDIAF, 2014.

- [13] S. Kim, Y. Ban, and S. Lee, "Face liveness detection using a light field camera," *Sensors*, vol. 14, pp. 22471-22499, 2014.
- [14] S. Kim, Y. Ban, and S. Lee, "Face liveness detection using defocus," *Sensors*, vol. 15, pp. 1537-1563, 2015.
- [15] G. g. Nieto, "Algorithmes de traitement et de visualisation pour la vision plénoptique à grande distance," 2017.
- [16] G. C. Birch, "Computational and design methods for advanced imaging," 2012.
- [17] Y. Rousseau, "Analyse des espaces de lignes pour la capture de cartes d'environnement," 2007.

2. Sites internet

- [18] Système de reconnaissance faciale 10 octobre 2020 https://fr.wikipedia.org/wiki/Syst%C3%A8me_de_reconnaissance_faciale#_Attaques_photo
- [19] https://fr.wikipedia.org/wiki/Appareil_photographique_pl%C3%A9noptique#Notes_et_r%C3%A9f%C3%A9rences
- [20] https://en.wikipedia.org/wiki/Lytro#Lytro_Illum
- [21] Pascal MARTIN, « PLÉNOPTIQUE SYSTÈME, photographie », Encyclopædia Universalis [en ligne], consulté le 20 juin 2020. URL: <http://www.universalis.fr/encyclopedie/systeme-plenoptique-photographie/>)

