# Optical and software tools for the design of a new transparent 3D display.

Thomas Crespel

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

# DOCTEUR DE

# L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE MATHEMATIQUES ET INFORMATIQUE

SPÉCIALITÉ INFORMATIQUE

Par Thomas CRESPEL

# OUTILS OPTIQUES ET LOGICIELS POUR LA CONCEPTION D'UN NOUVEL ÉCRAN 3D TRANSPARENT

Sous la direction de : Xavier GRANIER
(co-directeur : Patrick REUTER)

Soutenue le 9 décembre 2019

**Membres du jury :**

| | | |
|---|---|---|
| M. PICART, Pascal | Professeur, Université du Mans | Rapporteur |
| M. MANTIUK, Rafal | *Reader*, Université de Cambridge | Rapporteur |
| M. GUITTON, Pascal | Professeur, Université de Bordeaux | Président |
| M. TRAVIS, Adrian | Ingénieur de recherche, Travoptics | Examinateur |
| M. GEORGIOU, Andreas | Chargé de recherche, Microsoft Research | Examinateur |
| M. GRANIER, Xavier | Professeur, Institut d'Optique | Directeur de thèse |
| M. REUTER, Patrick | Maître de conférences, Université de Bordeaux | Co-directeur de thèse |

**Titre :** Outils optiques et logiciels pour la conception d’un nouvel écran 3D transparent

## Résumé :

Nous vivons une époque exaltante où de nouveaux types d'écrans sont rendus possibles, et la communauté scientifique s'emploie à améliorer l'expérience utilisateur. Nous vivons notamment l'émergence d'écrans courbés, volumétriques, autostéréoscopiques, transparents ou même portés sur la tête, avec des capteurs et algorithmes de plus en plus complexes permettant des interactions toujours plus riches.

Cette thèse vise à contribuer à la création de ces nouveaux types d'afficheurs. À travers trois projets concrets, nous associons l'optique et l'informatique pour répondre à des problématiques spécifiques, avec l'objectif final de créer un nouveau type d'écran 3D. Chacun de ces projets a mené au développement de prototypes basés sur l'utilisation de picoprojecteurs laser, de caméras, d'éléments optiques et de logiciels dédiés.

Dans un premier projet, nous avons étudié les écrans sphériques : ceux-ci sont plus adaptés que les écrans classiques pour visualiser des données sphériques, cependant les solutions existantes sont onéreuses et difficiles à mettre en place. Nous proposons une méthode pour concevoir un écran sphérique tactile à moindre coût en utilisant seulement des optiques commerciales et peu onéreuses ainsi que des éléments créés par impression 3D, dans le but de rendre ces écrans plus accessibles et reproductibles. Notre solution utilise un picoprojecteur laser associé à un système optique permettant de projeter une image nette sur toute la sphère. L'aspect tactile est réalisé par suivi optique de doigts dans l'infrarouge et nous avons développé un logiciel permettant de gérer l'affichage et l'interaction. Nous compensons l'utilisation de matériel peu coûteux par des calibrations et des corrections logicielles.

Nous avons ensuite largement étudié la technologie des guides "wedges" (en forme de "cale"), qui sont devenus des éléments essentiels du reste de la thèse. Les guides wedges ont été initialement développés pour des systèmes de projection plats, mais dans ce projet nous les utilisons dans un contexte d'acquisition. La problématique est la suivante : dans certaines configurations, une zone d'intérêt peut être difficile à imager avec une caméra standard à cause du manque d'espace en face de celle-ci. Nous proposons d'utiliser un guide wedge et un film prismatique afin de replier la distance nécessaire. Nous avons étudié et validé différentes applications dans le domaine spécifique de l'archéologie.

Les compétences que nous avons développées au cours de ces deux projets nous ont permis d'imaginer et de concevoir un nouvel écran autostéréoscopique transparent. Un tel écran peut être vu comme une vitre permettant d'ajouter au monde réel des informations 3D dépendantes du point de vue, et cela sans avoir besoin de porter de lunettes ou de casques. Le principe est d'utiliser un guide wedge avec des picoprojecteurs laser générant chacun un point de vue différent. Les points de vues sont répartis en face de l'écran par un élément optique holographique que nous avons spécialement conçu. Ce nouvel écran ouvre le champ à de nombreuses applications en réalité augmentée.

**Mots clés :** Informatique, optique, écran 3D, holographie, réalité augmentée

**Title:** Optical and software tools for the design of a new transparent 3D display

## Abstract:

We live exciting times where new types of displays are made possible, and current challenges focus on enhancing user experience. As examples, we witness the emergence of curved, volumetric, head-mounted, autostereoscopic, or transparent displays, among others, with more complex sensors and algorithms that enable sophisticated interactions.

This thesis aims at contributing to the creation of such novel displays. In three concrete projects, we combine both optical and software tools to address specific applications with the ultimate goal of designing a three-dimensional display. Each of these projects led to the development of a working prototype based on the use of picoprojectors, cameras, optical elements, and custom software.

In a first project, we investigated spherical displays: they are more suitable for visualizing spherical data than regular flat 2D displays, however, existing solutions are costly and difficult to build due to the requirement of tailored optics. We propose a low-cost multitouch spherical display that uses only off-the-shelf, low-cost, and 3D printed elements to make it more accessible and reproducible.

Our solution uses a focus-free projector and an optical system to cover a sphere from the inside, infrared finger tracking for multitouch interaction, and custom software to link both. We leverage the use of low-cost material by software calibrations and corrections.

We then extensively studied wedge-shaped light guides, in which we see great potential and that became the center component of the rest of our work. Such light guides were initially devised for flat and compact projection-based displays but in this project, we exploit them in a context of acquisition. We seek to image constrained locations that are not easily accessible with regular cameras due to the lack of space in front of the object of interest. Our idea is to fold the imaging distance into a wedge guide thanks to prismatic elements. With our prototype, we validated various applications in the archaeological field.

The skills and expertise that we acquired during both projects allowed us to design a new transparent autostereoscopic display. Our solution overcomes some limitations of augmented reality displays allowing a user to see both a direct view of the real world as well as a stereoscopic and view-dependent augmentation without any wearable or tracking. The principle idea is to use a wedge light guide, a holographic optical element, and several projectors, each of them generating a different viewpoint. Our current prototype has five viewpoints, and more can be added. This new display has a wide range of potential applications in the augmented reality field.

**Keywords:** Computer science, optics, 3D display, holography, augmented reality

## Unité de recherche

Equipe Manao, Inria Bordeaux Sud-Ouest, 200 avenue de la Vieille Tour, 33405, Talence
LP2N, Institut d'Optique d'Aquitaine, Rue François Mitterand, 33400, Talence

# Acknowledgments

The success of this thesis would not have been possible if I was not surrounded by outstanding people that I would like to acknowledge.

I wish to thank first my thesis director Xavier Granier and co-director Patrick Reuter. Since the beginning and the creation of this project, I have always felt supported. From their mentoring, I have appreciated their scientific rigour, their ideas and human qualities. They gave me the liberty to explore my ideas, the motivation of always digging more and the opportunity to discover and trust myself, and I am grateful to them for this. The trips with Patrick for conferences, demonstrations or appointments have been really enjoyable professional and human experiences.

I am grateful to the committee members of my thesis defense, Pascal Picart, Rafal Mantiuk, Pascal Guitton, Andreas Georgiou and Adrian Travis, for the time they have spent reading the manuscript, attending my defense, and for the interesting discussions. I would like to address a special thank to Adrian Travis who have largely contributed to the success of the thesis through his expertise and experience with wedge guides. It has been a real pleasure to collaborate with him in the exploration of their applications.

Thanks also to external people that I have had the chance to meet and that have also contributed to this thesis. I wish to deeply thank Felix Relats Monsterrats and his team for their welcoming in Egypt, allowing us to validate applications of our prototype and widening my horizon to new topics. Thanks also to Yves Gentet and Pierre Picart, for their contribution on recording the holographic optical element for the 3D display. In their laboratory, I have lived awesome and fruitful scientific experiences. Thanks also to Bertrand Simon, Frédéric Burgy, Philippe Coni, Anke Brock, for their interest in my work and their good advices at some points of the thesis.

Then, I also want to specially thank all current and former people of the Manao team for their daily accompaniment: David, Antoine, Charlie, Camille, Mégane, Corentin, Charlotte, Pascal, Pierre, Romain, Gaël, Anne-Laure, Arthur, Loïs, Ivo, Georges, Boris, Brett, Carlos, Thibaud. I am particularly grateful to Antoine, Charlie and Brett who have more directly contributed to some aspects of this work. I also want to thank the interns that have worked with me on parallel topics: Clara, Sofiane, Valentin and Timothée.

Finally, I am very grateful to those who accompanies me outside of work, and through their support, have largely contributed to this thesis. Thanks to my family of course, that has supported and placed their trust in me all my life , and to my friends for making life more spicy. Special thanks to Joséphine, for her daily accompaniment and for boosting me in difficult times. I also have a heartfelt thoughts for my grandmothers, that still help me surpass myself.

# Résumé détaillé

## Introduction

De nos jours, les écrans sont omniprésents dans tous les domaines de la société. Cependant, la majorité d'entre eux partagent les mêmes propriétés: ils sont plats, rectangulaires, et affichent des images 2D. Ces propriétés permettent de représenter une grande diversité d'informations mais limitent en même temps la manière dont ces informations peuvent être présentées et perçues.

Depuis quelques années, nous vivons l'émergence de nouveaux dispositifs d'affichage qui changent progressivement notre rapport au numérique, et c'est dans ce contexte que se place cette thèse. Les écrans peuvent notamment être courbés, flexibles, transparents, interactifs, et cela ouvre le champ à de nouvelles possibilités. En particulier, la réalité augmentée nous permet maintenant d'intégrer des données numériques superposées directement au monde rééel.

Les solutions de réalité augmentée existantes sont soit faites au travers d'un écran standard (par exemple sur un *smart phone*) et dans ce cas l'utilisateur ne voit qu'une projection 2D du monde réel et de son augmentation virtuelle. Cela peut se faire également au travers du port d'un casque de réalité augmentée, et dans ce cas l'utilisateur peut percevoir directement le vrai monde, ainsi qu'une augmentation perçue en 3D. Par contre, le fait de porter un casque constitue une forte contrainte, et nous avons remarqué que très peu de solutions existaient pour percevoir une augmentation 3D du monde réel sans casque ou lunettes.

Nous avons donc développé un écran 3D transparent qui permet de présenter de l'information numérique en 3D superposée au monde réel. Pour rendre cela possible, nous avons tout d'abord étudié en détail les technologies d'écrans 3D existantes (Section ), puis nous avons travaillé sur deux autres projets complémentaires. Ains, dans la Section , nous introduisons notre solution logicielle et matérielle pour développer un écran sphérique tactile basé sur l'utilisation d'un vidéoprojecteur laser. Ensuite, nous décrirons dans la Section  notre prototype de caméra basée sur un guide *wedge* permettant d'imager des zones contraintes. Nous utilisons ensuite des vidéoprojecteurs lasers associés à un guide *wedge* pour développer notre solution d'écran 3D transpar-

ent que nous décrivons dans la Section . Dans tous les aspects de cette thèse, nous avons une approche qui combinent à la fois l'optique et le numérique afin de résoudre des problématiques variées.

# Écrans 3D

Les écrans 3D cherchent à reproduire le maximum d'indices de profondeur présents dans notre perception du monde réél. Deux indices de profondeur son particulièrement importants: la *disparité binoculaire*, qui fait que chaque oeil perçoit une image différente, et la *parallaxe de mouvement*, qui est le mouvement relatif des objets à différentes distances.

La plupart d'entre nous a déjà expérimenté la 3D au cinéma: il est nécessaire de porter des lunettes pour reproduire la disparité binoculaire. Similairement, les casques de réalité augmentée et virtuelle peuvent reproduire le même effet mais intégrent également la parallaxe de mouvement, renforçant l'immersion 3D. Le fait de porter des lunettes ou un casque représentent quand meme une forte contrainte et les technologies d'afficheurs 3D sans lunettes sont intensivement étudiées par la communauté scientifique. On peut séparer les technologies d'affichage 3D sans lunettes en plusieurs catégories:

- Les afficheurs **autostéréoscopiques** consistent à distribuer différentes vues à différents emplacements de l'espace. Cela peut être fait par exemple en utilisant un écran classique et une surface de micro-lentilles cylindriques.
- Les afficheurs **volumétriques** permettent d'afficher de l'information dans un volume, par exemple en bougeant très rapidement un écran 2D dans l'espace.
- Les afficheurs ***light-field*** cherchent à reproduire la configuration des rayons lumineux émis par la scène virtuelle et contrôlant angulairement la valeur de luminance des rayons, par exemple par un empilement de panneaux LCD.
- Les afficheurs **holographiques** cherchent à reproduire un front d'onde par le biais d'interférences. Ils produisent des informations 3D de haute qualité mais ont encore beaucoup de limitations matérielles et logicielles.

Toutes ces technologies sont détaillées dans le Chapitre 3. En particulier, les afficheurs autostéréoscopiques par projection permettent de distribuer des vues sans diviser la résolution spatiale, et la génération des images est relativement compatible avec les techniques de rendu actuelles.

# Écran sphérique tactile

Un écran sphérique tactile est un globe opaque sur lequel est affiché de l'information virtuelle avec laquelle un utilisateur peut interagir directement par le toucher. Ce type d'écran peut être utilisé par exemple pour afficher des données géographiques ou mathématiques à des fins pédagogiques. Il peut également servir de support de travail collaboratif ou de support de communication original, pour de la publicité par exemple. Des solutions commerciales existent mais sont peu démocratisées, car elles restent assez onéreuses. Nous proposons dans le Chapitre 4 une méthode pour fabriquer un écran sphérique à moindre coût, dans le but de démocratiser ce type d'afficheur et permettre aux chercheurs ou passionnés de développer leur propre écran sphérique tactile.

Pour cela, nous utilisons un projecteur LBS (*Laser Beam-Steering*). Ce type de projecteur est *focus-free*, c'est à dire qu'il projette une image nette à toute distance, sans avoir á régler la position d'un plan image. Nous utilisons cette propriété afin de projeter une image nette sur une surface sphérique. Nous utilisons une sphère de diamètre 24cm avec une ouverture permettant de projeter la lumière depuis l'intérieur. Un système optique composé de deux parties est placé devant l'ouverture.

D'un côté, le vidéoprojecteur LBS est imagé à travers un système à deux lentilles. Le flux passe ensuite à travers une lentille *fisheye* qui étend l'image sur toute la sphère. D'un autre côté, une caméra infrarouge permet d'imager l'intérieur de la sphère à travers la même lentille fisheye. Cela est possible grâce à l'utilisation d'un miroir froid qui reflète la lumière visible du projecteur tout en étant transparent aux rayons infrarouges. Des LEDs infrarouges entourent la lentille fisheye, et illuminent la sphère depuis l'intérieur. Ainsi quand l'utilisateur pose le doigt sur la surface, celui-ci est imagé sur le capteur de la caméra: cela permet de réaliser par logiciel un suivi optique de doigts, et donc de rendre la surface tactile.

L'aspect logiciel est d'ailleurs indissociable de la conception matérielle car il permet de donner à l'écran ses propriétés uniques. En effet, le terme d'*écran sphérique tactile* n'a de sens que si l'image affichée est bien adaptée à la sphère et que l'interaction tactile est fonctionnelle, tout cela en prenant en compte la configuration matérielle.

Pour adapter l'image source du projecteur, qui est donc originalement construite sur un écran 2D classique, il est nécessaire de réaliser une transformation spécifique, appelée projection azimuthale équidistante. Il faut imaginer que la texture d'entrée est par exemple une version "cylindrique" classique d'un planisphère, et l'image de sortie est une projection dans un cercle centré sur le pôle Nord. Les paramètres de la projection dépendent de la configuration réelle du dispositif et celles-ci sont estimés par deux étapes de calibration qui doivent être faites une seule fois. D'abord,

un utilisateur règle, sur l'image source du projecteur, la position du pôle Nord, de l'équateur, et du plus grand cercle visible de manière a adapter la transformation. Des effets de décalage de couleur peuvent être introduits par le système optique, et l'utilisateur peut également les corriger à l'aide de notre logiciel. Les images sources subissent donc une projection azimuthale et une correction colorimétrique: ces deux transformations, qui prennent en compte la configuration réélle, sont programmées dans des shaders pour être ex'ecutées en temps réel sur la carte graphique.

Dans le même temps, l'image de la caméra est analysée par des algorithmes de vision par ordinateur pour extraire les coordonnées des doigts de l'utilisateur. A ce stade, ces coordonnées sont en espace caméra, et il est nécessaire de leur appliquer une transformation donnant les coordonnées correspondantes en espace projecteur. Cette transformation est estimée par une autre étape de calibration demandant de toucher une série de points sur la sphère. Cela donne une liste de correspondances qui est ensuite extrapolée pour générer une texture liant chaque pixel de la caméra à un pixel du projecteur. Ainsi, chaque coordonée en espace caméra est transformée en espace projecteur et l'image projetée est modifiée en fonction.

Ce travail a été présenté á la conférence *Display Week 2017* [CRG17]. De plus, nous avons présenté un outil permettant d'accélérer le développement d'applications sur un écran sphérique tactile à *Pervasive Displays 2017* [CRR+17]. Nous pensons que ces travaux permettront de démocratiser ce type d'afficheurs et contribueront ainsi à changer notre rapport aux données numériques.

## Caméra wedge et applications en archéologie

Dans le Chapitre 5, nous présentons d'abord la technologie des guides de lumière *wedge* (en forme de "cale"), qui sont des guides optiques utilisés à la fois pour des applications d'affichage et d'acquisition. Il s'agit d'un élément très important dans la réalisation de l'écran 3D présenté dans le Chapitre 6, mais ici nous les utilisons pour l'acquisition. La problématique est la suivante: une caméra traditionnelle a besoin d'une distance minimale entre l'optique d'entrée et la zone à imager afin que l'image soit nette et de taille raisonnable. Dans certaines configurations, cette distance minimale ne peut être atteinte en raison du manque d'espace en face de la zone à capturer. Nous proposons de rediriger les rayons lumineux dans un guide *wedge* grâce à des micro prismes, et de les guider vers une caméra classique. Nous avons développé un prototype basé sur cette idée et testé des applications dans le domaine de l'archéologie.

Un guide wedge est constitué de deux parties: une partie à faces parallèles où la lumière est guidée par réflexion totale interne comme dans un guide de lumiére classique, et une partie en "cale" où l'une des faces est inclinée par rapport à l'autre.

Cette deuxième partie permet de faire diminuer l'angle de réflexion interne jusqu'à ce que l'angle critique soit atteint et la lumière sorte du guide. Cela permet de réaliser notamment des dispositifs de projection plats, où l'image d'un projecteur peut s'étendre dans la première partie et sortir du guide pour être diffusée sur la surface de sortie. Réciproquement, cela permet d'imager la surface par une caméra en guidant les rayons dans le guide, un peu à la manière d'un scanner à plat. Tel quel, la surface de sortie peut etre imagée mais des objets situés au délà ne sont pas visibles car seuls les rayons arrivant à angle rasant entrent dans le guide.

Afin d'obtenir un dispositif permettant d'imager des objets au delà de la surface, nous utilisons un film prismatique pour rediriger les rayons. Il s'agit d'une surface constituée de micro-prismes étirés dans une direction, et qui permet aux rayons incidents à 45° d'être dirigés dans le guide jusqu'à une caméra située à l'extrémité. Nous proposons également d'illuminer la zone d'intérêt à travers le même guide, à l'aide d'une source de lumière située à côté de la caméra. Le guide et les prismes introduisent des distorsions optiques que nous corrigeons par logiciel pour obtenir l'image finale.

Nous avons développé un prototype intégrant tous ces éléments afin d'étudier des applications en archéologie. En effet, les archéologues cherchent parfois à documenter des zones peu accessibles en étant le moins invasif, c'est à dire en introduisant le minimum de perturbation autour de la zone d'intérêt. En particulier, il n'est pas forcément possible de dégager un espace nécessaire en face de la zone d'intérêt pour y insérer une caméra classique. Avec l'aide d'archéologues, nous avons identifié trois applications concrètes:

- Sondages sous-terrains: Actuellement, si les archéologues veulent étudier ce qu'il y a sous le sol, il est nécessaire de creuser des sondages de plusieurs mètres carrés afin de pouvoir y descendre et inspecter les parois. En plus d'être chronophage, cela n'est pas toujours possible soit par une configuration de terrain particulière soit par des autorisations qu'il peut être difficile d'obtenir. Nous montrons que notre dispositif peut être inséré dans des sondages beaucoup plus étroits pour obtenir plus rapidement des informations sur le sous-sol.
- Sondages sous-marins: Nous montrons qu'en adaptant un peu notre prototype, il peut également fonctionner sous l'eau. Actuellement, la profondeur des sondages peut être limitée par le niveau de la nappe phréatique et notre dispositif peut donc être utilisé pour aller au delà. De plus, certaines zones innondées ont également des limitations d'espace auxquelles il est possible de répondre avec notre système, le rendant ainsi compatible avec des applications en archéologie sous-marine.
- Inspection de blocs: Dans l'Histoire, il n'est pas rare que des pierres soit récupérées d'anciens édifices pour en construire de nouveaux, et dans ce cas des

inscriptions (comme des hiéroglyphes) peuvent être cachées entre deux blocs de pierre. Notre prototype peut être inséré dans l'espace entre des pierres formant un bâtiment pour inspecter leur surface cachée, mettant en évidence la présence ou non d'inscriptions et permettant ainsi de retracer l'histoire de la structure.

Nous avons publié ces travaux dans le *Journal On Computing and Cultural Heritage* [CTRG19]. Nous avons particulièrement insisté sur les applications en archéologie, mais l'intérêt d'un tel dispositif va au delà. Il peut par exemple être utilisé pour inspecter des machines difficilement déplaçables. Ce projet nous a également permis d'étudier en profondeur le comportement des guides *wedges* qui nous seront particulièrement utiles pour la suite.

# Écran autostereoscopique transparent

Dans le Chapitre 6, nous réunissons les compétences acquises lors des deux projets précédents pour proposer un écran autostéréoscopique transparent. Un tel écran permet de superposer des informations 3D au monde réel, et cela sans avoir à porter de lunettes ni de casque. Cela ouvre donc la porte à de nombreuses applications en réalité augmentée, comme par exemple l'exposition augmentée de produits ou d'objets muséaux à travers une vitre.

Notre solution utilise des projecteurs LBS couplés dans un guide wedge. À la place d'utiliser un film prismatique pour rediriger les rayons comme dans le Chapitre 5, nous utilisons un Élément Optique Holographique (EOH) que nous avons conçu. Son rôle est de diffuser l'image de chaque projecteur dans des zones de vues différentes situées devant le dispositif.

Nous avons créé l'EOH en faisant interférer deux faisceaux lumineux: l'un représentant les rayons issus d'un projecteur et sortant du guide, et l'autre représentant la zone de vue à recréer. L'EOH est un hologramme volumique présentant une sélectivité angulaire qui lui permet de ne diffuser que la lumière des projecteurs tout en étant transparent aux lumières ambientes. Nous l'avons enregistré pour les trois longueurs d'onde des projecteurs et, mis à part quelques problèmes colorimétriques que nous discutons dans le document, il permet bien de générer des zones de vues indépendantes, adjacentes, et en couleur.

Notre prototype utilise 5 projecteurs, générant ainsi 5 zones de vues situées à 50cm de l'écran avec une largeur de 3cm. L'écart interoculaire moyen étant de 6cm, un utilisateur localisé à 50cm peut donc percevoir deux images différentes avec ses yeux: cela reproduit la disparité binoculaire. De plus, l'utilisateur peut bouger sa tête horizontalement pour changer de zone de vue, et cela reproduit la parallaxe de mouvement. L'ensemble permet de recréer une très bonne perception de la 3D, à

condition bien sûr que les images sources des projecteurs soient correctement générées et cela est fait par notre logiciel.

D'abord, nous générons 5 vues différentes d'une même scène 3D, où la disposition des caméras représente la position des zones de vues. Ensuite ces images subissent une transformation en temps réél qui prend en compte la configuration matérielle, comme dans le Chapitre 4. Cette transformation est estimée par une étape de calibration basée sur de la vision par ordinateur. Une caméra définit une grille commune pour tous les projecteurs, et un algorithme trouve, pour chaque projecteur, les points qui s'alignent sur cette grille. Nous extrapolons ensuite les points calibrés pour générer une texture stockant les correspondances entre les points sources et les points cibles permettant l'alignement des 5 images au niveau de l'EOH.

Ces travaux ont été présentés à *Digital Holography And 3D Imaging* [CRGT19], publiés dans le journal *Applied Optics* [CRT+19] et font actuellement l'objet d'une demande de dépôt de brevet. De plus, nous avons présenté le prototype pendant 3 jours de démonstration sur le salon de *Display Week 2019* à San José (Californie). Cela nous a permis d'avoir des retours de nombreuses personnes, dont beaucoup de spécialistes dans le domaine des écrans. Les retours étaient globalement très positifs, à la fois par la qualité de l'expérience 3D et par les possibilités offertes par notre solution.

# Travaux futurs

Les trois projets sur lesquels nous avons travaillé au cours de cette thèse ouvrent la voie à de nouvelles recherches que nous décrivons dans cette section.

L'écran sphérique peut être amélioré en terme de luminosité par un système optique anamorphique, ou alors en remplaçant le projecteur par un module de balayage laser plus contrôlable. Le tracking peut être amélioré en adaptant la technique de *réflexion totale interne frustrée* à une surface sphérique. Enfin, nous pourrons nous servir du prototype pour mener une série d'études utilisateurs concernant la perception et l'interaction avec les données sur un écran sphérique tactile.

Concernant la caméra wedge, nous envisageons le développement d'un prototype de plus petite taille avec un logiciel de reconstruction de panoramas automatique pour explorer des zones restreintes de manière plus pratique et rapide. Nous pourrons également intégrer plusieurs caméras, voire même des projecteurs, pour extraire directement des données 3D des zones imagées. De plus, d'autres moyens de redresser les rayons pourront être étudiés, notament par l'utilisation potentielle d'EOHs.

Nos futurs recherches sur l'écran 3D viseront d'abord à corriger les problémes colorimétriques décrits en travaillant sur l'EOH. Nous pourrons également étudier les alternatives possibles en termes de taille ou position des zones de vues. Notamment,

avoir des zones de vues plus petites et un systême permettant de démultiplier le nombre de projecteurs fournirait une parallaxe horizontale plus dense. En parallèle, nous étudions aussi les bénéfices apportés par des techniques de suivi oculaire. En effet, connaître la position des yeux de l'utilisateur permettrait notamment d'intégrer une parallaxe verticale, d'améliorer la parallaxe horizontale, et d'uniformiser la luminosité. Nous allons également étudier le développement de nouvelles applications en réalité augmentée permises par notre écran, notamment pour la mise en valeur d'objets du patrimoine culturel.

# Conclusion

Dans cette thèse, nous avons exploré différents chemins, menant tous à la réalisation de prototypes ayant chacun leurs applications. Dans chacun de ces projets, nous avons utilisé une approche qui combine l'optique et l'informatique, en utilisant des caméras et des projecteurs avec des systémes optiques et logiciels spécifiques. Notre message est que ces deux approches sont nécessaires pour répondre aux enjeux de demain et qu'elles doivent être abordées en parallèle.

Nous avons proposé une méthode pour fabriquer un écran sphérique tactile à moindre coût, afin de diversifier les possibilités d'afficher et d'interagir avec de l'information. Nous avons également décrit notre prototype de caméra wedge permettant d'imager des zones restreintes et avons présenté des applications en archéologie, mais beaucoup d'autres domaines sont concernés. Enfin, nous avons proposé une solution intégrant un guide wedge, des projecteurs laser, un élément optique holographique et un logiciel pour le dévelopement d'un écran autostéréoscopique transparent.

Les écrans autostéréoscopiques sont de bons candidats dans la réalisation d'écrans 3D car ils sont compatibles, dans une certaine mesure, avec l'état de l'art sur les écrans 2D. Cependant, ils sont selon moi une solution temporaire et l'écran 3D parfait devrait être capable de reproduire un *light-field* complet et de haute résolution. L'approche *light-field* est plus éloignée des technologies actuelles et a donc encore beaucoup de limitations matérielles et logicielles, cependant elle cherche à reproduire la réalité de maniére la plus fidèle. Réinventer les techniques d'affichage, d'acquisition et de génération de scènes 3D en ce sens pourrait finalemement lever les verrous technologiques et ainsi reproduire un *light-field* plus réaliste en un temps interactif.

Nous pensons qu'à terme, des solutions similaires à celles de notre écran 3D pourraient être capable de générer un light-field de haute résolution représentant tous les aspects d'une scène 3D et se mélangeant parfaitement avec une scène réelle. Avec ce type d'écran, la distinction entre le réél et le virtuel sera de plus en plus difficile.

# Contents

# Chapter 1

# Introduction

Nowadays, our daily life is surrounded by displays, and not a single day passes by without being exposed to one. According to a study [Nie18], the average American adult spends more than 9 hours a day in activities involving a display. Displays are not limited to TVs anymore, and a typical day is cadenced by multiple devices: from smartphones, tablets, PCs, to information displays and advertisements in the street. They are today essential in many aspects of our society.

Yet, the vast majority of displays share the same properties: they are rectangular, flat, opaque and they provide 2D images. When we think about it, these common features are clearly restraining compared to the richness and vastness of the world they are supposed to represent. Such displays define a framework within which the content can be created and perceived. Outside this framework, everything is left to interpretation.

Displays are designed for humans: this statement might sound obvious but has many implications. From the choice of the refresh rate or the colors of the pixels, everything is done to trick the spectator into thinking that what he sees is real. You do not need colors to feel terrified when watching Alfred Hitchcock's "Psycho", yet the transition from black and white to color displays has been a huge technical and artistic revolution, giving a tremendous new freedom to content creators and to the feelings a spectator can experience. Our modern displays still have limitations that we do not always notice, but they definitely shape the way we see the world.

In a world where ideas and technologies are constantly evolving, we must question our relationship and experiences with displays. Many companies and institutions are working collectively to push the limits of what a display can offer, with the will to always improve the experience of the end user. Recent advances in optics, electronics, and computer science have led to the development of many alternative ways of creating, perceiving and interacting with digital content.

We notably witnessed the democratization of 3D movies in cinemas, and they fall in the category of stereoscopic displays. Once again, the human perception is at the heart of their development, as they seek to reproduce elements involved in depth perception. Stereoscopic displays are based on *binocular disparity*: the fact that both eyes sense a slightly different image of a 3D scene. They usually rely on wearing glasses to separate the images, and this is an additional restriction compared to regular displays. Moreover, the illusion of depth perception is done through binocular disparity only, and there is so much more involved in natural vision.

Augmented and Virtual Reality (AR/VR) technologies have also recently reached the consumer market. In VR, a user is completely immersed into a virtual world, while in AR, the user can experience a digital augmentation of the real world. There are many ways to achieve this, and a headset is generally required to experience binocular disparity. However, compared to 3D movies, these headsets reproduce a lot more elements of natural vision and among them, *motion parallax*. It is the ability to move the head with respect to the 3D scene in order to appreciate different views of it and to obtain a better 3D perception. Still, the wear of a headset is also a strong limitation and in such a technological world, we might expect to experience 3D perception of digital content with our naked eyes.

The development of glasses-free 3D displays is a hot topic that has rallied many researchers for decades. The general idea is still to deliver different information to each eye, except that this functionality is directly performed by the display. In addition to binocular disparity, they should reproduce other aspects of depth perception, such as motion parallax, in order to provide a complete 3D experience. These challenges can be handled with various approaches, each with its own strengths, limitations and applications, and we will particularly outline them in the present document.

This thesis aims globally at investigating the development of displays that go beyond the traditional 2D plane that we experience every day. We strongly believe that changing our relationship with digital data can change society in many aspects: from entertainment, communication and teaching to transportation, health and science. The revolution of displays is already triggered, and we are willing to contribute to it.

In particular, we noticed a gap between augmented reality and glasses-free 3D displays. In order to experience a 3D augmentation of the real world, the user generally has to wear a heavy headset. Anyone who has tried AR headsets knows that the benefit is worth the effort as the experience is simply unique, however, the headset remains a strong restriction. We came up with the idea of a transparent 3D display that does not require the user to wear any glasses or headsets. The interest of such a display for AR is straightforward: imagine a regular window through which you can see both the real world and a 3D digital world that are superimposed. Our dis-

play shows both binocular disparity and motion parallax, allowing the user to move relatively to the 3D scene and to have an improved depth perception.

Our solution is the result of a long journey that was driven by three main components. First, the analysis of current display technologies: from the mainstream displays that we find everywhere to the more specific ones that have not yet reached the market. Secondly, the use of optics: after all the design of a display is all about modulating and directing light toward a human eye. Lastly, the use of computer science allows the construction of a digital content that is adapted to the real hardware configuration.

Our 3D display is not the only contribution of this thesis. Such a journey requires a lot of curiosity and exploration, and it would not have been possible without two other prior projects, that are as different as complementary. First, we developed a spherical multitouch display that also contributes to change our relationship with virtual data and to think "out of the box". We then developed a specific camera for a use in an archaeological context, which shares a lot of similarities with the 3D display. Both of these projects had a huge impact into our design choices of the proposed 3D display, as we will outline throughout this dissertation.

## 1.1 Organization of the document

This document is organized as follows:

In **Chapter 2**, we introduce the notions that will lead our research. In particular, our approach combines both optics and computer science, and we will review the required knowledge in each of these fields. We also review the basic display technologies, because they are usually the starting point toward more sophisticated displays.

In **Chapter 3**, we introduce the research on 3D displays. We detail how 3D images are processed by the brain, and how a display can produce the illusion of 3D. We briefly describe glasses-based 3D displays before we extensively focus on glasses-free 3D displays. We describe and comment the current state-of-the-art technologies and challenges.

In **Chapter 4**, we present our first prototype of a multitouch spherical display. Our solution offers the possibility to display data directly on a spherical surface and to interact with it by touch. Several solutions already exist, and we propose a low-cost approach in the hope of spreading their use. For this, we combine a focus-free projector in an optical system and an infrared camera for finger tracking.

In **Chapter 5**, we introduce wedge light guides through the development of an imaging device. A wedge guide is able to fold the projection distance of a projector, and similarly, the observing distance of a camera. In this work, we use it for acquisition purposes: we propose a solution to image confines spaces and we study particular applications in an archaeological context.

In **Chapter 6**, we present our proposed solution for a new glasses-free 3D transparent display. Such a display is able to superimpose 3D data onto the real world without glasses and thus may find many augmented reality applications. The idea is to exploit the transparency of the wedge guide and to use multiple laser projectors to project different images. The images are distributed in front of the display by a custom holographic optical element that we precisely describe.

# Chapter 2

# Prior Knowledge

The work presented is this thesis relies on different notions from different scientific domains that we need to briefly introduce. A display is first an optical system that modulates or emits light to create a signal designed to be seen. We thus present some basic concepts on optics (Section 2.1) before introducing more specific vocabulary for display technologies (Section 2.2). One specificity of our approach is that we always consider the design not only from the optical system point of view, but also from a sofware point of view in order to create and adjust the signal. We thus also present some required concepts on computer graphics and computer vision in Section 2.3.

## 2.1 Optics

### 2.1.1 Physical nature of light

Light has the properties of both a particle and a wave. Since centuries, light has been mostly theorized as rays, composed of tiny particles. This gave birth to the theories of geometric optics theorized across centuries by notably Euclid [Bur45] (-4th century), Alhazen [Win53] (2nd century) and Descartes [Des87] (17th century). Newton [New84] developed his corpuscular theory in the 17th century, and meanwhile his contemporary Huygens showed that light has the properties of a wave and demonstrated geometric optics with a wave approach [Huy20]. The famous double-slit experiment of Young validated the wave theory in the 18th century by showing that light passing through two slits constructs interferences that can be constructive or destructive. Based on Huygens theory, Fresnel formulated the Huygens-Fresnel principle [HYFA00] in 1815, stating that every point in space can be considered as point sources, generating spherical waves that continuously sum up and interfere with each other. Maxwell showed later that light was an electromagnetic wave, just like microwaves, X-rays or Gamma rays, and he unified the wave theory [Max65]. Ein-

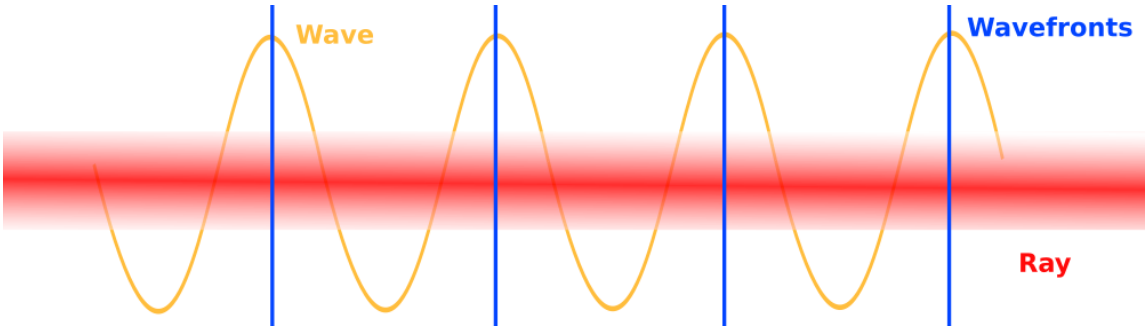



***Figure 2.1:*** *Propagation of light in a medium can be considered in a simplified approach as rays carrying photons is straight lines. In reality, it is also an electromagnetic field oscillating around the direction of propagation and defining a wavefront.*

stein later introduced the concept of photons [Ein05], which are individual quantum particles carrying energy ($E = \frac{hc}{\lambda}$ with h the Planck's constant, $c$ the velocity and $\lambda$ the wavelength). Some effects, such as the photoelectric effect, can be explained only with the quantum particle approach and not the wave theory. Both of these approaches are complementary and necessary to explain the behaviour of light.

### 2.1.2 Wave properties of light

In this thesis, we will independently consider light as a ray or a wave. Alternatively, we can also represent a wave as a wavefront, which is defined as a surface of constant phase. Light is a transverse wave, so the wavefronts are locally perpendicular to the direction of propagation. For planar waves (i.e. directional beams), the wavefronts are parallel planes, and for spherical waves (i.e. point sources), they are concentric spheres.

**Spectrum** Light differs from other electromagnetic waves by its **wavelength**, which is the spatial period of oscillation of a wave, and it is inversely proportional to the energy it carries. As shown in Figure 2.2, visible light has wavelengths comprised between 400nm and 700nm. Ultraviolet and infrared lights are not visible to the human eye and have wavelengths that are respectively below 400nm and above 700nm. Natural light is always **polychromatic**, that means that they are composed of several wavelengths. The **spectrum** of a light source is the intensity of each wavelength composing it. If the spectrum of a light source is so narrow that we can consider that it carries a single wavelength, then the light source is **monochromatic**.

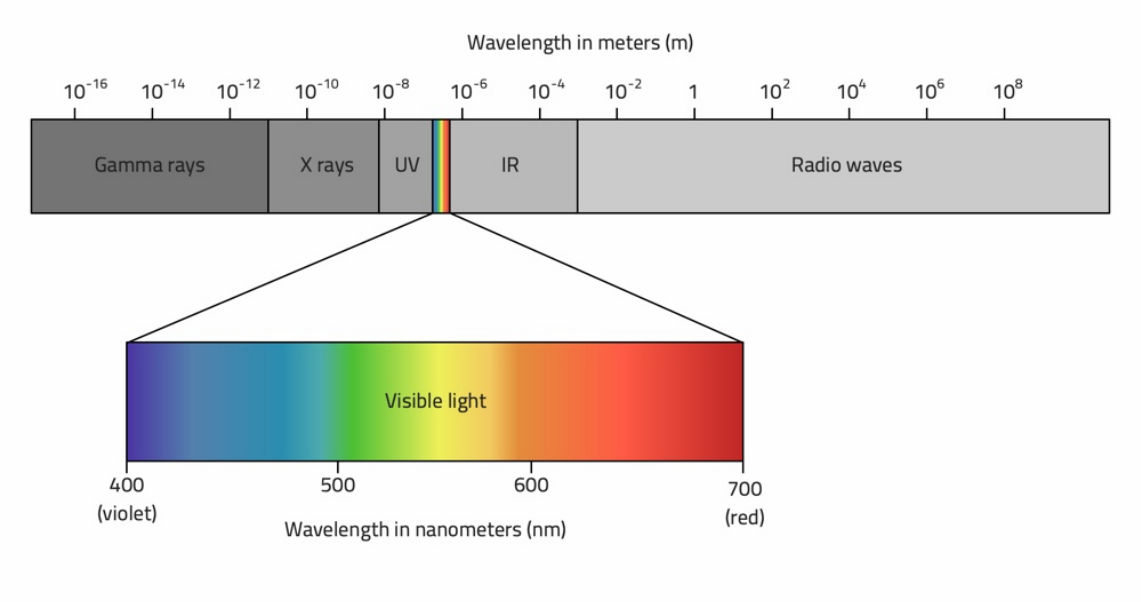



***Figure 2.2:*** *Visible light differs from other electromagnetic waves by its wavelength, comprised between 400 and 700nm (image source: radio2space.com).*

**Mathematical expression** By the principle of superposition, a polychromatic wave can be decomposed as a continuous sum of monochromatic waves of expression:

$$U(\vec{r},t) = U_0 \cos(\omega t - \varphi(\vec{r})) \tag{2.1}$$

where $\vec{r} = (x, y, z)$ represents the coordinates of a point in space, $\omega$ is the temporal pulse of the wave, $t$ the time, $U_0$ is the maximum amplitude of the wave, and $\varphi(\vec{r}, t)$ is the phase of the wave. The phase depends notably on the direction of propagation and length of propagation, expressed as $\varphi(\vec{r}, t) = \vec{k} \cdot \vec{r} + \varphi_0$. In this expression, $\vec{k}$ is the wave vector, which is directed toward the direction of propagation and whose length is $\vec{k} = \frac{2\pi}{\lambda}$ with $\lambda$ the wavelength, i.e. the spatial period of the wave. $\varphi_0$ is the phase at the origin.

To ease the calculations, the complex notation is often used and a light wave is written as:

$$\mathbf{U}(\vec{r},t) = U_0 \exp^{-j(\omega t - \vec{k}\cdot\vec{r} - \varphi_0)} \tag{2.2}$$

We will denote in bold letters the complex numbers.

**Irradiance/Intensity of a wave** The intensity of a wave is defined as $E = |\mathbf{U}|^2$. We draw attention to the reader that the term "intensity" is adapted to describe waves in general. When referring specifically to light, it is better to use the term of

"irradiance". This is to prevent any confusion with the "intensity" that we define later.

**Polarization** As a transverse wave, light oscillates in a direction perpendicular to the direction of propagation. The polarization state of light describes the way the direction of oscillation evolves with time in a perpendicular plane.

In general, natural light is unpolarized and in this case the direction of oscillation of the electric field is random. Light that is polarized has privileged directions of oscillation evolving with time. As represented in Figure 2.3, the wave can be rotated on an elliptic trajectory, with the particular cases of linear and circular polarizations.

**Interferences** In particular, waves are able to interfere with each other and this can be notably exhibited with Young's slits experiment. For simplicity, let us consider two monochromatic waves $U_1$ and $U_2$ incident on a plane located at $z = z_0$, then the total irradiance at a point $\vec{r} = (x, y, z_0)$ is expressed as:

$$E(x,y) = |\mathbf{U_1}(\vec{r}) + \mathbf{U_2}(\vec{r})|^2 \tag{2.3}$$

$$E(x,y) = |\mathbf{U_1}(\vec{r})|^2 + |\mathbf{U_2}(\vec{r})|^2 + \mathbf{U_1}(\vec{r}).\mathbf{U_2}^*(\vec{r}) + \mathbf{U_1}^*(\vec{r}).\mathbf{U_2}(\vec{r}) \tag{2.4}$$

We consider two **coherent** light sources, that means that they oscillate with the same pulse $\omega$ and their phases can be compared. In particular, this can be achieved by separating a laser beam in two. This condition is necessary to exhibit interferences.

Upon these hypotheses, waves can be expressed as

$$\mathbf{U_1}(\vec{r}, t) = U_{1,0} \exp^{-j(\omega t - \varphi_1(\vec{r}))} \tag{2.5}$$

$$\mathbf{U_2}(\vec{r}, t) = U_{2,0} \exp^{-j(\omega t - \varphi_2(\vec{r}))} \tag{2.6}$$

and the total irradiance on a screen can be developed as

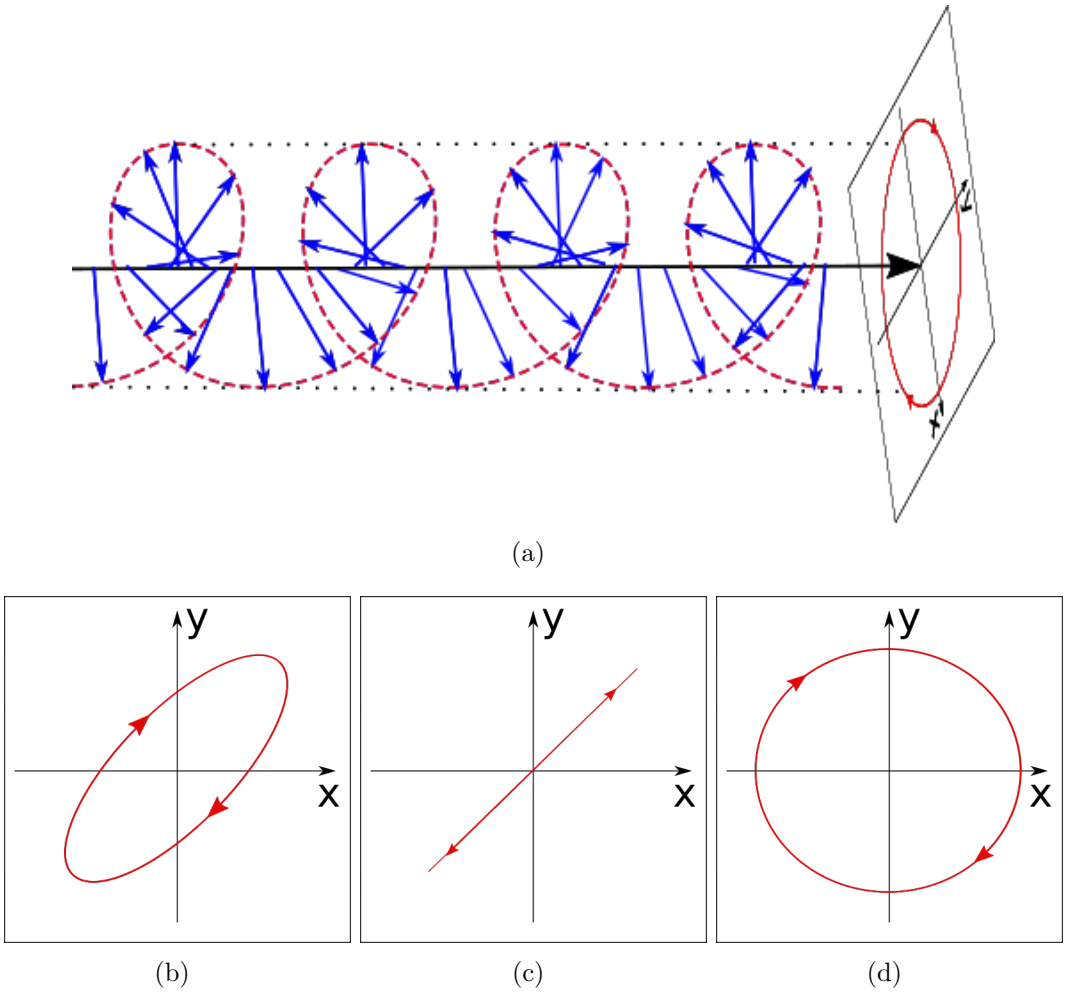
$$E(x,y) = U_{1,0}^2 + U_{2,0}^2 + U_{1,0}U_{2,0} \cos(\varphi_1(x,y,z_0) - \varphi_2(x,y,z_0)) \tag{2.7}$$

The phase difference depends on the difference of traveled distances, which is different for each position on the screen $(x, y)$. The irradiance varies from $U_{1,0}^2 + U_{2,0}^2 - U_{1,0}U_{2,0}$ to $U_{1,0}^2 + U_{2,0}^2 + U_{1,0}U_{2,0}$ with a modulation of $cos(\Delta\varphi)$.

Typical interferences with two spherical waves are illustrated in Figure 2.4.

### 2.1.3 Measuring light

In this section, we introduce the vocabulary and units relative to the measure of light. More specifically, we introduce *radiometric* units, which are used to measure electromagnetic waves in general. Their equivalent *photometric* units are introduced at the end of the section. This is a brief introduction, and more detailed information can be found in [Mey01].




(a)

(b) (c) (d)

***Figure 2.3:*** *(a) Propagation of a polarized electromagnetic wave in space and time; (b) Elliptic polarization; (c) Linear polarization; (d) Circular polarization. Note that spherical and elliptic polarization states also include the direction of rotation.*

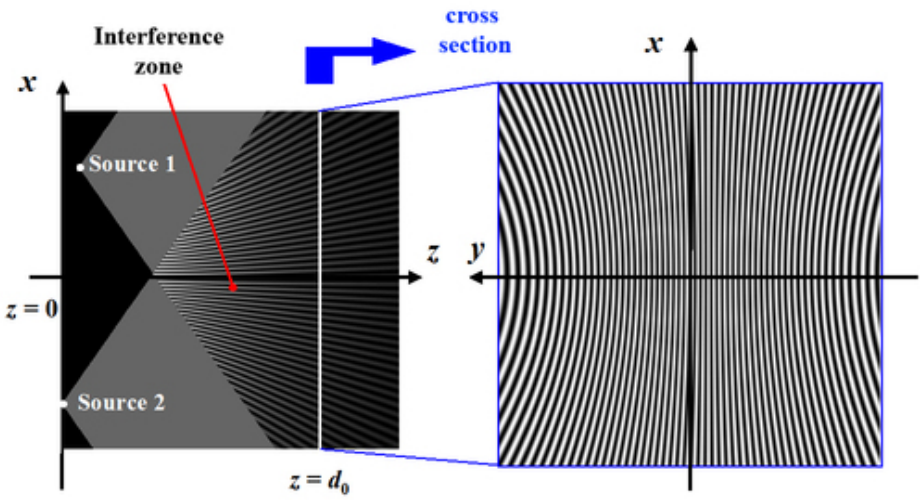



***Figure 2.4:*** *Interferences between two spherical waves in space (left) and on a screen (right) (image source: optique-ingenieur.org)*

**Energy** The energy is the most straightforward notion as it is directly the quantity of light. It is the number of photons, weighted by their individual energy, passing through a surface in a given period of time. The energy is measured in Joules ($J$).

**Flux** The flux is the debit of energy with time, expressed in Watts ($W = J.s^{-1}$). It can be seen, by analogy with other physic fields, as the optical power of a light source. Note that in optics, the term "optical power" rather refers to the "focusing strength" of an optical element.

The luminous flux of a light source represents the total debit of energy distributed along all directions. It should also not be mistaken with the power consumption that is often given for general public products: a 20W light bulb does not deliver 20W of flux but instead consumes 20W of electric power to run.

This quantity is generally not used to characterize broad angle light sources, like a regular TV, as it does not provide much relevant information. Instead, the flux of a laser or a projector is more meaningful because it is confined in a specific "beam".

**Intensity** The intensity of a light source characterizes the flux in a specific direction and is expressed in Watts per steradian ($W.sr^{-1}$). A laser, for example, has a high energy in a specific direction but a small flux. It is more adapted than the flux for characterizing broad angle light sources, as an observer can only see the source from a limited angle.

| Name | Radiometry | Photometry |
|---|---|---|
| Energy | $J$ | $lm.s$ |
| Flux | $W$ | $lm$ |
| Intensity | $W.sr^{-1}$ | $cd$ $(= lm.sr^{-1})$ |
| Irradiance | $W.m^{-2}$ | $lx$ $(= lm.m^{-2})$ |
| Luminance | $W.m^{-2}.sr^{-1}$ | $nt$ $(= cd.m^{-2})$ |

***Table 2.1:*** *Equivalence of radiometric and photometric units*

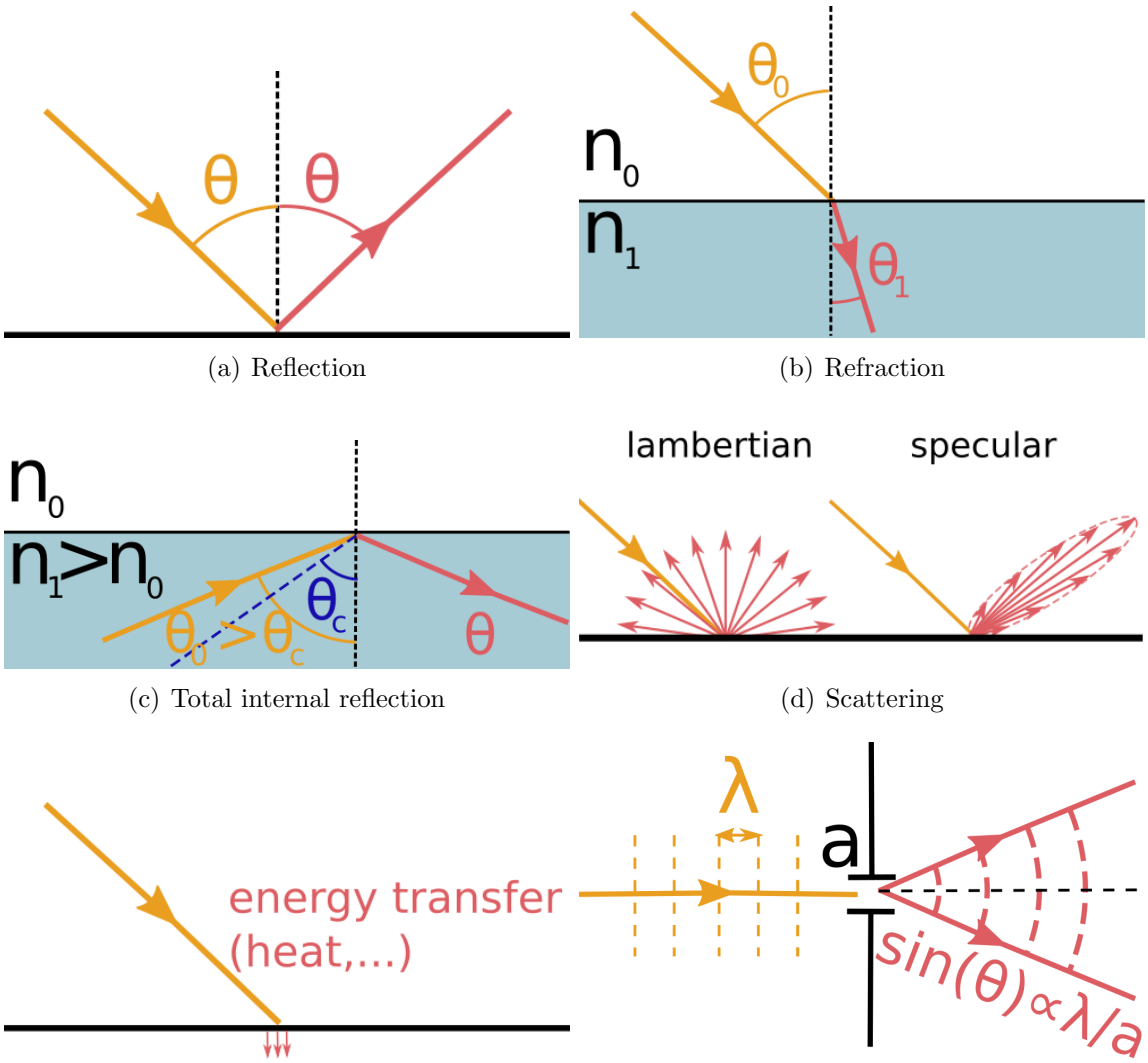
**Irradiance** The irradiance is the amount of flux incident on a surface. It is denoted $E$ and is expressed in Watt per surface unit ($W.m^2$). From a wave approach, it is the squared modulus of the complex amplitude $E = |\mathbf{U}|^2$. The notion of irradiance usually does not characterize point light sources, as it requires a receptor located at a certain distance. Instead, the notion of *illuminance* can be used for area sources and characterizes similarly the total flux emitted by unit surface.

**Luminance** Luminance characterizes the flux ($W$) in a specific direction ($sr^{-1}$) per unit surface ($m^{-2}$) and is thus expressed in $W.sr^{-1}.m^{-2}$. It can characterize either light received by a surface, emitted by a surface, or travelling in space.

**Photometric units** *Photometry* is the science of measuring light from a perceptual approach. All the above notions still exist but they are weighted by the spectral response of the eye at the considered wavelengths. Lumen ($lm$), candela ($cd$), lux ($lx$) and Nits ($nt$) are introduced with their radiometric equivalent in Table 2.1.

### 2.1.4 Interaction of light with matter

When incident on matter, light can behave in a variety of ways depending on the matter's properties. In this section, we describe the most common types of elementary interaction of light with matter. Note that the complete interaction is generally a combination of several of these interactions. The angle of incidence of a ray or wave on a surface is defined with respect to the surface's normal, which is a unit vector locally perpendicular to the surface.




(a) Reflection

(b) Refraction

(c) Total internal reflection

(d) Scattering

(e) Absorption

(f) Diffraction

***Figure 2.5:*** *Light interaction with matter*

### Reflection

Light can be reflected, and an element performing mostly reflection is a mirror. In this case, Descartes' laws specify that the reflection happens in a plane that contains the input light ray and the normal of the surface, and the angle with the normal is reversed.

### Refraction

Dielectric media have the ability to change the direction of light. The index of refraction $n$ characterizes the refractive power of a material.

When passing through a medium of a different index of refraction, a light ray is refracted. Its direction of propagation is changed depending of the ratio of the indexes of refraction of both media. Snell-Descartes' laws specify notably that the refracted ray lies in a plane containing the normal of the surface and the input ray. For a ray incident on the interface with an angle $\theta_0$ from a medium of index of refraction $n_0$, to a medium of index of refraction $n_1$, then the refracted angle $\theta_1$ obeys to

$$n_0 \sin\theta_0 = n_1 \sin\theta_1 \tag{2.8}$$

If $n_1 > n_0$, then the ray if refracted toward the normal, otherwise it goes away from the normal. Typical examples of refractive material are water, glass and transparent plastic such as acrylic.

### Total internal reflection

In the case where $n_1 < n_0$, for example when leaving an acrylic surface ($n_0 = 1.49$) toward the air ($n_1 = 1$) the refracted angle is higher than the incident angle. From this property emerges the notion of a limiting angle of incidence from which $\theta_1$ reaches $90°$ and cannot go further. This angle is called the critical angle, denoted $\theta_c$ and can be derived from Descartes' law with $\theta_1 = 90°$.

$$\theta_c = \sin^{-1}\left(\frac{n_1}{n_0}\right) \tag{2.9}$$

Beyond the critical angle, most of the energy is reflected instead of refracted.

### Scattering

When meeting a diffusing material, the energy of a ray can be distributed along a variety of directions. This is the most common interaction and is the reason why things are visible from different angles. The scattering can happen in several ways:

in reflection, transmission, or through more complex light paths (e.g. subsurface scattering). In case of a uniform distribution of light, then the surface is called lambertian. When the diffusing angle is close to the reflection direction, then the surface is specular. An element for which scattering is the major interaction with light is called a diffuser. For projection devices for example, a lambertian diffuser is generally used to see the projected image from a wide angle. A diffuser with a narrow scattering angle is a directional diffuser.

#### Absorption

Depending on the material, a small or large amount of the energy carried by the light can be absorbed by it. The luminous energy is transformed into other types of energy. In general the generated energy is thermal (as for a black cloth) but it can also be electric energy for semiconductors, chemical energy for the photosynthesis, and so on. In a photosensitive material, like an unexposed camera photofilm, the absorption of a photon causes chemical reactions that result in a modulation of the local transmittance.

#### Diffraction

When light meets an obstacle or a slit, it is bended. This property is inherent to the wave aspect of light and can be explained with the Huygens-Fresnel principle by considering interferences between a continuous distribution of point sources.

Without going too much into details, diffraction is visible when light passes through a small slit or obstacle of size $a$ of the same order of magnitude of the wavelength $\lambda$. Then light is diffracted, in a simplified approach, in a cone of angle $\theta$ such that $sin\theta \propto \lambda/a$. For small angles, the approximation $sin\theta \approx \theta$ is generally used. In reality, higher orders of diffraction exist at higher angles, but we do not consider them in the scope of this document.

A diffracting element composed of tiny slits or showing a transmittance that is modulated at high frequency is a grating. Holograms are particular cases of gratings and we specifically describe them in the following section.

### 2.1.5 Holography

Holography has been invented by Gabor in the 1960's [GSB+65]. It is an imaging technique based on interferences to record and reproduce the relief of an object. This process recreates the full depth of an object, and an example hologram can be seen in Figure 2.6.

(a) (b) (c)

***Figure 2.6:*** *Three different views of a reflection color hologram made by Yves Gentet. Note that no physical objects are present, instead the hologram is a thin plate shaping light as if the object was there. (image source: ultimate-holography.com)*

In this section, we focus on analogical holography to exhibit the general principle and properties of holograms. Digital holography and holographic displays will be detailed later in Chapter 3. The general principle of holography can be explained in two phases: the recording and the reconstruction.

### Recording

The recording of a hologram is done on a photosensitive material that reacts to incident light, as in classical photography. The transmittance $T$ of a photosensitive film under normal exposure (neither overexposed nor underexposed with regards to its absorbance) is proportional to the incident irradiance $E$.

When an object is lit, it re-emits a wavefront that can be roughly modelled as

$$\mathbf{O}(\vec{r}, t) = O_0 e^{-j(\omega t - \varphi_O(\vec{r}))} \tag{2.10}$$

If the recording film is directly exposed to $O$ then, after development, $T \propto E = |\mathbf{O}|^2 = O_0^2$. It can be observed that the phase information, containing the information of relief, is not recorded. This is what happens in classical photography, except that the object is first imaged by a lens, but only the amplitude of light is recorded.

In holography, the key idea is to record interferences between the object beam and a reference beam that is coherent with it. In this case, the phase difference is exhibited in the recorded irradiance. Denoting $\mathbf{R}$ the reference beam, the irradiance

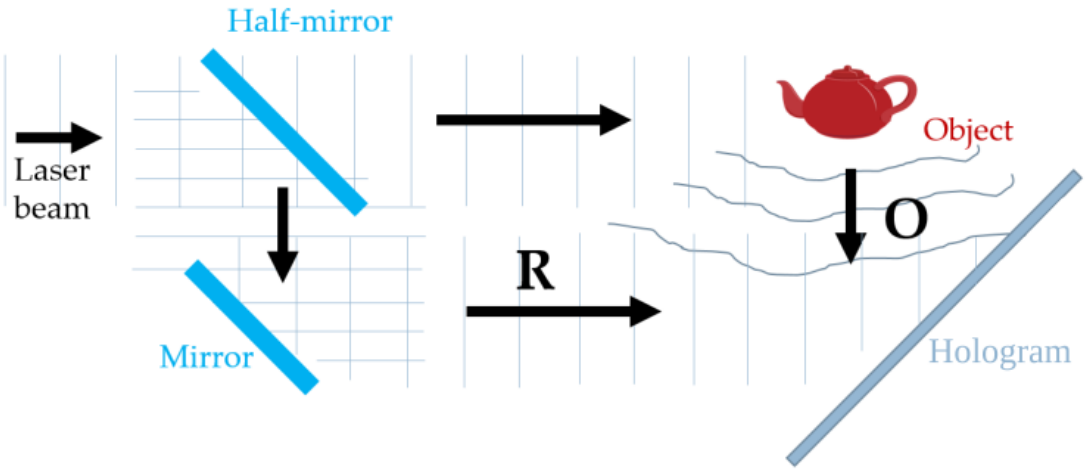



***Figure 2.7:*** *Typical recording of a transmission hologram: a laser beam is divided in two beams. The reference beam R is a simple beam incident on the hologram. The object beam O is created by the scattering of the object. Note that for reflection-type holograms, the reference and object beams are incident on both sides.*

on the plane of the film is (see Section 2.1.2):

$$\begin{aligned} E &= |\mathbf{R} + \mathbf{O}|^2 \\ &= |\mathbf{R}|^2 + |\mathbf{O}|^2 + \mathbf{R}\mathbf{O}^* + \mathbf{R}^*\mathbf{O} \\ &= R_0^2 + O_0^2 + 2R_0O_0cos(\varphi_R - \varphi_O) \end{aligned}$$

After development, the photosensitive film exposed to the interferences is called a hologram of an object. Such recorded holograms then behave as diffraction gratings because the frequency of modulation is of the order of the wavelength.

Holograms can be recorded either in transmission or reflection, and a typical setup to record a transmission hologram is represented in Figure 2.7.

### Reconstruction

A hologram recorded as described in the previous section and developed has a transmittance of

$$T \propto |\mathbf{R}|^2 + |\mathbf{O}|^2 + \mathbf{R}\mathbf{O}^* + \mathbf{R}^*\mathbf{O}$$

In a general way, a complex wave $\mathbf{U_i}$ incident on an element of transmittance $T$, transmits $\mathbf{U_t} = T \times \mathbf{U_i}$.

**With the reference wave** If the reference wave $\mathbf{R}$ is used to reconstruct the hologram, then the transmitted wave is

$$\begin{aligned}\mathbf{A}_t &= T \times \mathbf{R} \\ &= (|\mathbf{R}|^2 + |\mathbf{O}|^2)\mathbf{R} + \mathbf{R}^2\mathbf{O}^* + |\mathbf{R}|^2\mathbf{O}\end{aligned}$$

In this equation:

- $(|\mathbf{R}|^2 + |\mathbf{O}|^2)\mathbf{R}$ is a complex wave directly proportional to $\mathbf{R}$, because the modulation factor $(|\mathbf{R}|^2 + |\mathbf{O}|^2)$ is real. It is the order of diffraction 0.
- $\mathbf{R}^2\mathbf{O}^*$ is the order of diffraction $-1$, a complex wave that is harder to interpret at first sight because it is a complex multiplication between three waves. It can be shown that this reconstructs the *pseudoscopic* image of the object, i.e. the object with inverted relief. It is generally an undesired order of diffraction.
- $|\mathbf{R}|^2\mathbf{O}$ is the most interesting term because it is directly proportional to the object beam, with a real modulation factor. It is the order of diffraction $+1$ and reconstructs the object wave as if the object was at its original position with regards to the hologram. The light rays diverge from the object, making it a virtual image.

In "in-line" holography, these three orders are aligned and overlap each other, while in "off-axis" holography, they have different positions, and so one can generally exploit one of the three orders independently.

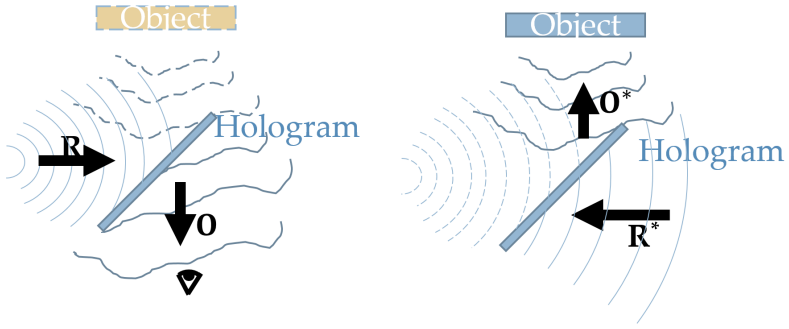
**With the conjugated reference wave** Instead of using directly the reference wave, a hologram can also be reconstructed using the *complex conjugated* beam $\mathbf{R}^*$ of the reference wave. The conjugated of a complex wave can be interpreted as the same wave propagating with the $-\vec{k}$ vector, i.e. the *counterpropagative* wave: the conjugated of a wave that converges to a point is the same wave but diverging from the same point. A hologram reconstructed with $\mathbf{R}^*$ transmits:

$$\begin{aligned}\mathbf{A_t} &= T \times \mathbf{R}^* \\ &\propto (|\mathbf{R}|^2 + |\mathbf{O}|^2)\mathbf{R}^* + |\mathbf{R}|^2\mathbf{O}^* + \mathbf{R}^{*2}\mathbf{O}\end{aligned}$$

In this equation, we find again the zero-th order $(|\mathbf{R}|^2+|\mathbf{O}|^2)\mathbf{R}^*$ that has no impact and an undesired term $\mathbf{R}^{*2}\mathbf{O}$, but the term that is the most meaningful is $|\mathbf{R}|^2\mathbf{O}^*$ ($-1$ diffraction order) because it is directly proportional to $\mathbf{O}^*$. The conjugated of

the object wave is reconstructed: it is a light wave that converges toward the object, thus it is a real image.

The differences between the virtual image of the object wave and the real image of the conjugated object wave are highlighted in Figure 2.8. The first one is viewable when looking through the hologram, the second one is visible when a diffuser is located at the intersection of rays.




(a) (b)

***Figure 2.8:*** *a) The reference beam is used for the reconstruction: a virtual image is visible when looking through the hologram. b) The conjugated reference beam is used for the reconstruction: a real image is visible on a screen*

### Properties of volume holograms

A volume hologram, or Bragg hologram, is a hologram whose thickness is larger than the wavelength used for the recording. They have interesting properties that have been theorized by Herwig Kogelnik in 1969 [Kog69]. The most important characteristic that we use in this thesis is the Bragg selectivity, and we explain it in the following. If the reconstruction beam has the exact wavelength and angle, called the Bragg angle/wavelength, of the reference beam used in the recording, then the image is reconstructed with a maximal efficiency $\eta_0$. If the angle and/or wavelength is slightly different, then the efficiency of diffraction lowers, typically with a $sinc^2$ behavior as shown in Figure 2.9. The range between the first extinctions is the bandwidth. An incident wave whose properties are outside the bandwidth is reconstructed with a nearly-0 efficiency, so they are not modified by the hologram that thus appears transparent. The tighter the bandwidth is, the more selective the hologram is.

These properties allow volume holograms to be multiplexed. Several recordings, at different angles of incidence or wavelengths, can be superimposed onto the same material. If the bandwidths are narrow enough not to overlap, they can reconstruct

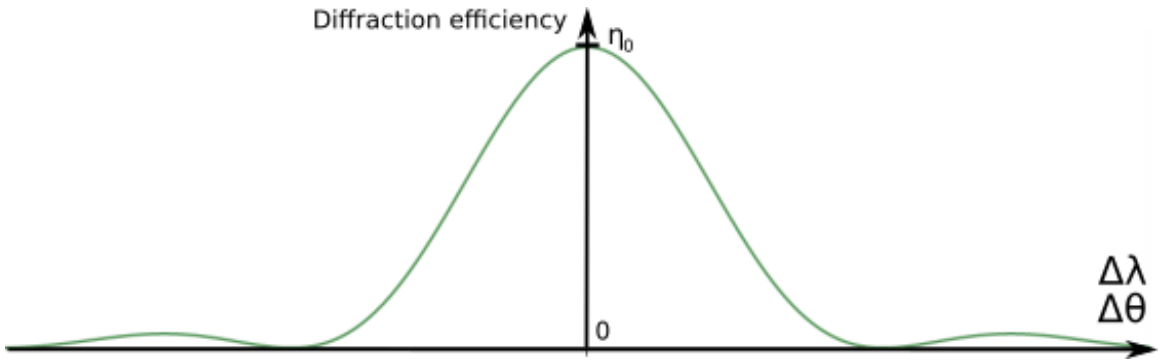


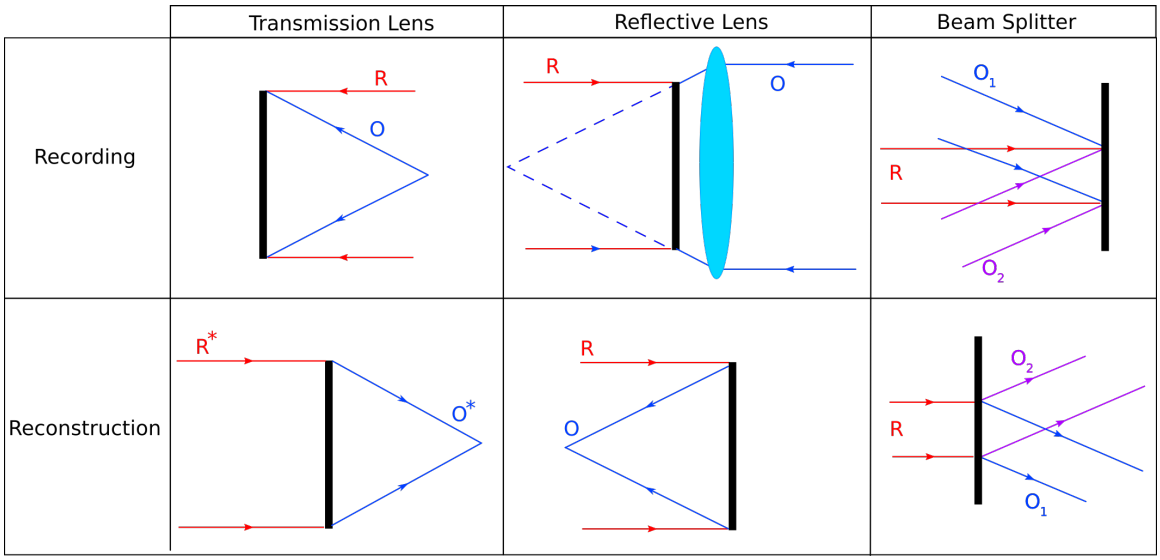

***Figure 2.9:*** *Typical efficiency curve of volume holograms with respect to angle or wavelength variations from the Bragg angle or wavelength.*

different images for several incident waves without interfering between each other. This is notably used for digital storage, where data can be superimposed on the same surface and read with different beams [CPS+00]. Wavelength-multiplexing is exploited notably to build color holograms. A full-color hologram is typically obtained by superimposing three gratings of different wavelengths, corresponding to red, green and blue colors.

### Holographic Optical Elements

Holographic optical elements are particular holograms that do not intend to reproduce 3D objects but an optical function. In particular, they reproduce the behaviour of regular optical elements such as lenses or mirrors. Figure 2.10 shows an example of recording and reconstruction of several optical elements. HOEs have interesting properties compared to conventional optics. Notably, they can have geometric properties that are difficult to achieve with their conventional equivalents. For example, a lens with a high optical power requires a large curvature that limits its size, whereas a HOE can perform the focusing function from a large surface, and the same goes for spherical mirrors. On the contrary, the larger the aperture of a lens, the thicker it gets, and a HOE can be a thin element performing the same function. Moreover, regular optical elements may exhibit optical aberrations that get worse when the angle of incidence goes away from the optical axis, and a HOE can record an optical function free from optical aberrations.

The wavelength and angular selectivity is also an interesting property for HOEs as several optical functions can be superimposed on the same surface. For augmented reality, the angular selectivity is used so that HOEs are transparent to ambient light while performing an optical function for the digital image to superimpose. In particular, we also use this property to design our 3D transparent display as we explain in




(a)

***Figure 2.10:*** *Examples of simple holographic optical elements: recording and reconstruction.*

Chapter 6.

## 2.2 Conventional 2D display technologies

In this section, we first review the vocabulary related to displays in Section 2.2.1. We review the existing technologies for flat panel displays in Section 2.2.2 that we define as a device that generates an updatable image from a 2D surface to the eye. Typical examples of flat panel displays are televisions (TVs) or smartphones. Finally, we review projector technologies in Section 2.2.3, which project the image in space, and require a scattering surface to see the image.

### 2.2.1 Generic terms

Digital displays divide the image in picture elements, called **pixels**, and we review in this section the different methods of controlling a pixel color.

The total number of pixels is called the **resolution** and is generally given as the number of pixels columns times the number of pixel rows (e.g. 1920x1080p for the HD resolution). **Resolution** is an ambiguous term as it also often refers to the pixel density. This comes directly from the optical meaning of resolution, which is the

minimum distance that can be resolved with an optical system. In this thesis, we use this term for both of the definitions, but the context will clearly indicate the used meaning.

Displayed images have an **aspect ratio**, which is the ratio between its width and its height. The aspect ratio of a single pixel can also be an interesting property, as they are not always square.

The image is visible from a certain solid angle that is called the **Field Of View** (FOV). Similarly, in this thesis we use the term FOV for a projector as the solid angle of projection, although it is sometimes referred to as the size of the projected image at a certain distance in existing literature. We can also specifically refer to the horizontal or vertical FOV, which are simple angles in radians or degree.

Projection displays have specific vocabulary. The **projection distance**, also called throw, is the distance between the projector's output and the projection screen. The **throw ratio** $t$ is defined as the ratio between the projection distance $D$ on the horizontal image size $W$: $t = D/W$. It is directly linked to the horizontal FOV.

The **refresh rate** of a display is the temporal frequency of the displayed images. A typical refresh rate for most of displays is 60Hz, which means 60 images per second.

### 2.2.2 Flat panel displays

#### Cathode Ray Tube (CRT) displays

CRT displays might not be considered as "flat" displays, because their apparatus is quite large, but it is the pioneer technology that allowed TVs to enter the home entertainment market, so it is interesting to review briefly their principle.

The term cathode ray tube refers to an electron gun. A CRT display uses a phosphorescent screen that contains red, green and blue phosphors. When the phosphors are struck by an electron beam, it glows with a controllable intensity. CRT displays uses temporal multiplexing, as the electron beam is rapidly scanned on the image area. CRT displays are a deprecated technology that is bulky, consumes a lot of energy and suffer from flickering due to the time multiplexing.

#### Plasma

Plasma displays generate pixels thanks to gas cells surrounded by electrodes. When subject to an electric field, the gas is ionized and emits ultraviolet light. The screen consists of a layer of colored phosphors, corresponding to the pixels, that emit red, green or blue light when illuminated by the ultraviolet light. They have been gradually abandoned as well, because other display technologies have outperformed them.

### Liquid Crystal Displays (LCD)

LCDs are the dominant technology in the current display industry. Their principle relies on attenuating a white backlight through several layers by means of polarization. A color filter array can be added to this arrangement for color displays.

The backlight is generally a high power light source, which is the case for TVs or smart phones. The light source can be of various types, fluorescent lamps or LEDs. For device running on battery, such as watches or calculators, the backlight can be a simple mirror reflecting ambient light, to reduce power consumption. For transparent displays, it can also be directly the ambient light or a projected light.

The attenuating layers consist of a LCD module enclosed in two polarizing films (see Figure 2.11). The LCD module is composed of a liquid crystals layer enclosed in two electrode layers. Liquid crystals are elongated molecules having birefringent properties. In particular, they align when exposed to an electric field and rotate the polarization of light. Several arrangements of layers are possible. The TN (Twisted Nematic) type of LCDs is the original technology and is represented in Figure 2.11. In the TN technology, the polarisers have crossed directions so that without the LCD layer no light passes through them, but the liquid crystals arrange in a shape of helix and rotate the polarisation. In this configuration, the light gets linearly polarized by the first polariser, is rotated by 90° by the LCD module, and exits the second polariser with maximum intensity. A controlled voltage applied on the electrodes can partially or totally untwist the molecules and rotate the polarisation to a controlled angle. The angle of polarisation defines how much light passes through the second polariser, giving the final pixel intensity. More recent technologies are IPS (In Plane Switching), where polarisers are parallel to each other and electrodes are in the same plane, but the principle remains the same.

The electrodes can be a passive matrix or an active matrix. A passive matrix addresses the current by applying a voltage from the edges, and the addressed pixel is determined at the intersection. Pixels are thus controlled sequentially at a high frequency. In an active matrix, each electrode is addressed individually by a TFT (thin film transistor). This method is generally preferred over the passive matrix because of the better image quality and refresh rate, however, passive matrices can still be found in low-cost displays.

### LED displays

LED displays directly use a 2D grid of LEDs that are directly pixels. They should not be mistaken with LCD displays that use LEDs as a backlight. LED displays are rather used in public displays such as public transports or store signs. They do not provide a high density of pixels, but are cheap and have a bright illumination. LED-

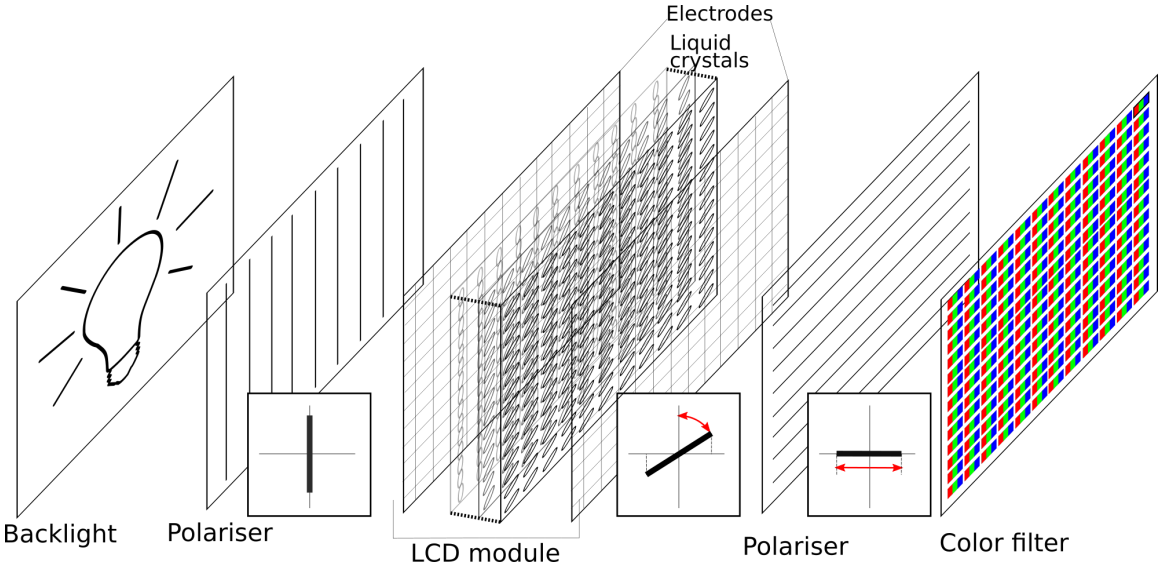



***Figure 2.11:*** *Decomposition of the layers composing a LCD display and states of polarization during the propagation of the back light. The light is generated by a backlight source, polarized with a first polariser, then the polarization is rotated with the electrically addressed liquid crystals, then the polarisation state is projected on a second polariser to give the final intensity value. The color filter array gives the final $R,G$ or $B$ color component.*

based displays intended for personal use are currently under development, notably with the MicroLED technology that consists in microscopic LED pixels.

#### OLED displays

OLED displays are an increasing market in the display industry since the end of the 2000's, and they are gradually replacing LCD displays in high quality display applications.

An OLED is a particular LED based on an organic layer situated between two electrodes, where at least one of them is transparent. When subject to an electric current, the organic layer emits light. Contrary to the LED displays where LEDs are individual electronic components, the organic layer is continuous and only requires an array of electrodes (as for LCDs), thus they can be made with a higher density. As for LCDs, the electrodes can either be a passive matrix (PMOLED) or an active matrix (AMOLED). As for LED displays, OLED displays do not require a backlight as the light is emitted directly by the organic layer, thus they can provide deeper blacks than LCDs.

OLEDs notably allow new kinds of displays. First, they have better transparency than other technologies because they do not attenuate light (like LCDs) and do not require large components blocking light (like LEDs). So transparent displays can be made out of OLED by designing transparent TFT [GSM+06]. Moreover, the organic layers can be curved surfaces, resulting in curved displays [Cok08] that can be even stretchable [HSK+17].

### 2.2.3 Projection displays

#### General considerations

The general principle of digital projection displays can first be explained with earliest non-digital projection devices as for example slide projectors. In a slide projector, a transparent slide containing an image is lit by a powerful light source. The slide is then magnified and imaged at a further distance by a projection lens. The slide is imaged on a plane called the imaging plane, where the screen should be located to have a sharp image. The distance between the projection lens and the object is adjustable, so that the position of the imaging plane can be tuned. The basic idea of a digital projector is to replace the slide with a digital display module.

If the screen is not at the position of the imaging plane, tilted or not planar, then the image is blurred. This comes from the fact that the lens has a large aperture, and then many rays from the same point have different output directions after the lens. They have to be focused on the screen so that a point is imaged on a point.

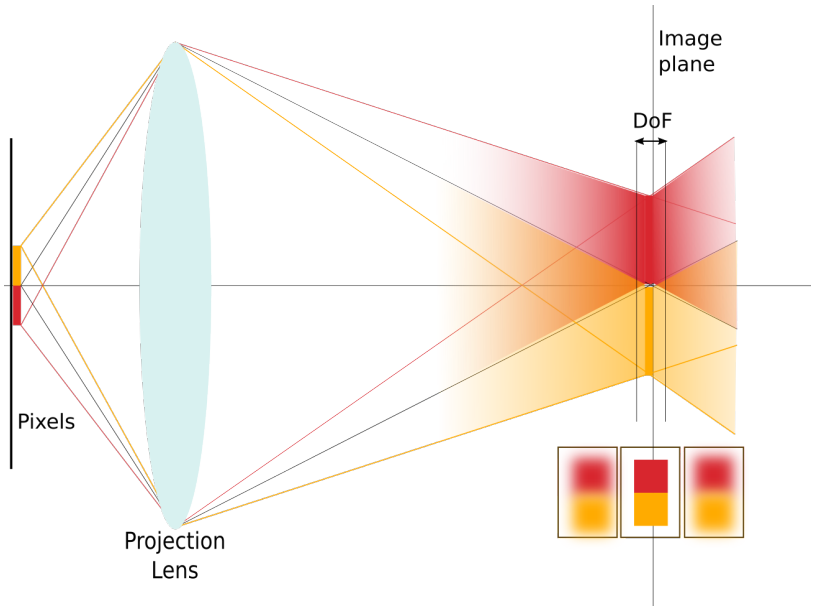



***Figure 2.12:*** *A large imaging lens introduces depth of field (DoF), because the rays focus from many directions. The image of a projector gets blurred from a certain distance from the imaging plane.*

This defines the concept of **depth of field**, which is the distance range around the imaging plane position where an image is considered to be sharp. The depth of field depends on the apparent size of the lens aperture at some distance. It is larger with small lenses and large projection distances.

There are basically four main projector technologies that use a projection lens: LCD, DLP, LED and LCoS. Compared to a slide projector, the principle is basically to replace the slide by a digital image. LBS projectors involve a particular technology that does not rely on this principle.

### LCD projectors

Following the analogy between a slide projector, replacing the slide directly with a LCD display can result in a digital projection device. However, in practice the apparatus is optimized for projection. A high-power collimated white light source is divided into red, green and blue beams that are incident on 3 different LCD chips. They are combined with dichroic mirrors into a single virtual full color LCD that is imaged on the screen as illustrated in Figure 2.13(a).

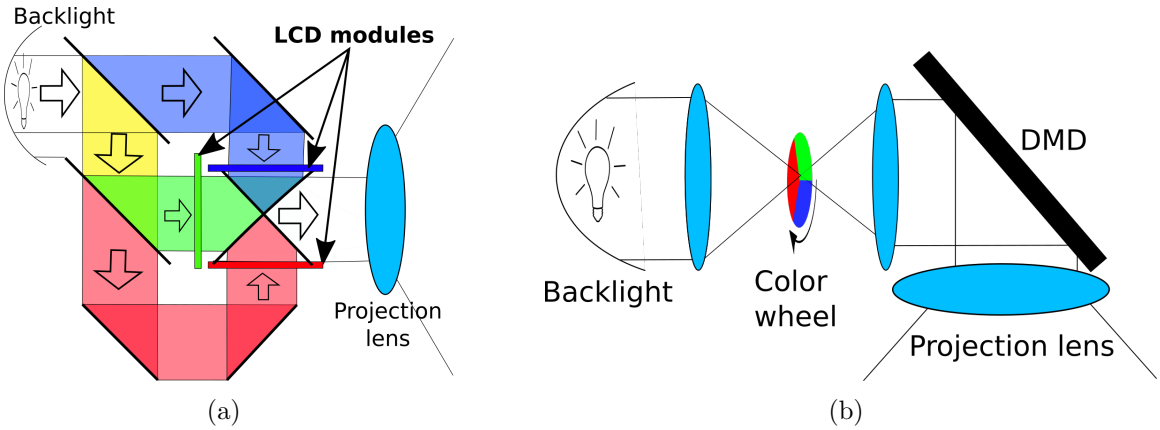



***Figure 2.13:*** *(a) Principle of a LCD projector; (b) Principle of a DLP projector.*

### Digital Light Processing (DLP) projectors

DLP projectors are time-mutiplexed devices, based on an electronic component called a Digital Micromirror Device (DMD). A DMD consists of a matrix of micro-mirrors that has two different orientation states: one that directs light toward the projection lens, and one that directs light toward an absorbing material. These two positions are basically "ON" and "OFF" states that are rapidly switched during the construction of a frame. The intensity of a pixel is defined by the amount of time the pixel is on the "ON" state. A DLP projector builds a color image by displaying the red, green and blue images sequentially, and this is done thanks to a rotating wheel located in the light path (see Figure 2.13(b)). This results in the well-known "rainbow effect" of these projectors that makes the sequential colors visible when a user quickly moves his eyes.

### LCoS

A Liquid Crystal on Silicon (LCoS) panel can be seen as a LCD that is based on reflection instead of transmission. LCoS projector features a back light that is reflected by the LCoS chip. This is quite similar as for DLP projectors, except that the direction of reflection is fixed and light is instead directly modulated by the LCoS. A color LCoS projector is generally achieved with a combination of three single-color LCoS modules, just like a LCD projector. It is also possible to build a color image sequentially on a single chip with a color wheel like the DLP technology.

LCoS projectors show better image quality than LCD or DMD but they are more expensive, so they still have not really reached the consumer market.

### LED and Laser projectors

LED and laser projectors do not refer to a projection technique and they can be either DLP, LCD or LCoS. They simply refer to the nature of the light source.

Generally, the backlight is generated by a bulb lamp and a spherical mirror to collimate all the energy in the same direction. Using a LED backlight consumes less energy and increases the lifespan of the lamp.

Lasers can be also used as light sources. They present many advantages but are still very expensive. The main advantage is the color fidelity: they have a wider color gamut than other technologies. Other light sources base color reconstruction on the combination of three wide spectra considered as "red", "green" and "blue", laser projectors modulates "pure" colors with very narrow spectra.

Moreover, they have a higher contrast, as the intensity of the backlight can be modulated faster than a bulb lamp, for darker frames for example. This property makes them also more energy efficient. Finally, the lifespan of the laser source is longer than the traditional bulb lamp.

The term "laser projector" is often used to qualify these projectors and they are famous notably in movie theaters. However, they should be referred to as "DLP Laser projector", for example, to be more precise. Another type of laser projectors, LBS projectors, relies on a totally different approach.

### Laser Beam Steering (LBS)

A LBS projector is a completely different type of projector that does not rely on a projection lens. They were first developped and commercialized by Microvision [FCM09] and current LBS projectors still rely on their technology.

The principle of work is depicted in Figure 2.14(a). In a LBS projector, the image is constructed pixel by pixel during a frame like in a CRT display. Each pixel is generated by a single laser beam that is rapidly scanned through the image. The beam is created by the combination of three laser diodes (R,G,B) (see Figure 2.14(a)) and its color can be controlled directly by modulating at high frequency the power of the input diodes. This beam is incident on a scanning micromirror device that deflects the ray on a trajectory represented in Figure 2.14(b).

The divergence of the laser beam is set up such that, beyond a minimal distance, its size evolves linearly with the image size (Figure 2.14(c)). This gives LBS projectors a focus-free property because the pixel size is consistent with the image size at any projection distance, contrary to conventional projectors.

This technology results in very compact designs and low power consumption, making them suitable for portable devices. The image has very high contrast ratios (which is the ratio of luminance of white pixels over black pixels) as the laser is

switched off during black pixels. They also have a very high color fidelity, for the same reasons as laser backlight projectors. However, eye safety considerations limit the brightness of such projectors because of the employed lasers.

## 2.3 Software considerations for displays

In this thesis, the solutions that we present rely on both hardware and software. In general, digital displays are based on pixels that must be computed and addressed by a computer. Displays feature a large number of pixels that have to be processed typically 60 times a second, so Graphics Processing Units (GPUs) are used. GPUs are dedicated hardware units for displays that allow parallel computations, and they are mandatory for any display application. In particular, they are very effective to perform 3D rendering and to create computer generated images as we describe in Section 2.3.1. They are also essential in many applications requiring massive computations and/or interactive rates.

To extend the limits of a display, computer vision can be used and this is described in Section 2.3.2. This field consists of analyzing data from one or several cameras in real-time, and this can also be performed on GPUs. For display applications notably, it can be combined with computer graphics to integrate real world data for the rendering of the final image, as we will see in Section 2.3.3.

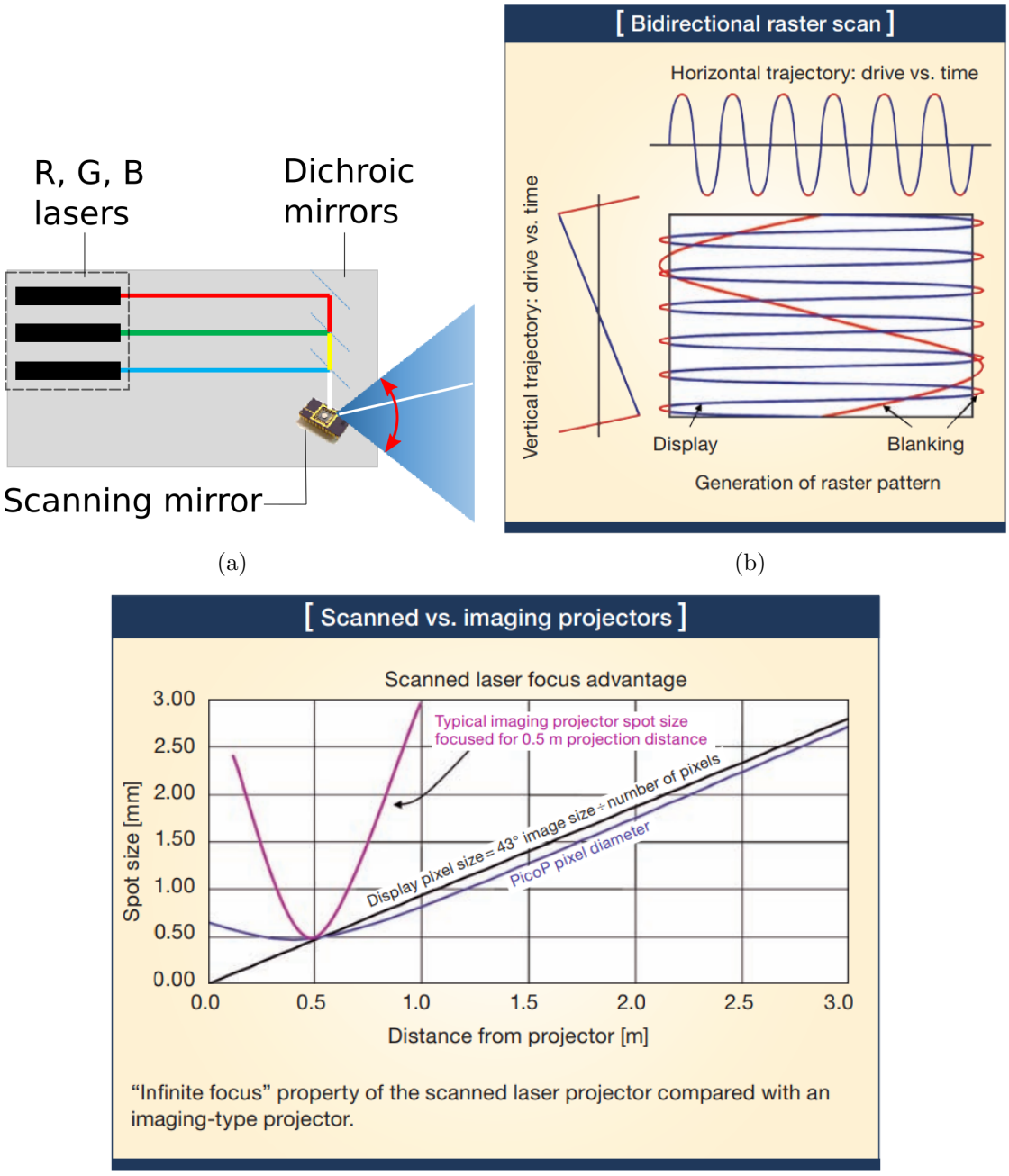
### 2.3.1 Computer graphics: from the digital world to the real world

Computer graphics is the process of transforming a 3D scene from an abstract digital representation to a 2D image. This is a crucial step as the generated 2D images are directly the input of a regular display. For more complex displays, and as we do in this thesis, the images can also be modified based on the hardware configuration.

#### Representation of a virtual scene

As for a real scene, a 3D virtual scene is composed basically of 3D objects and light sources, and 2D images are created with virtual cameras. A 3D scene is mathematically represented on a computer, and in this section we present the most widely used representation.

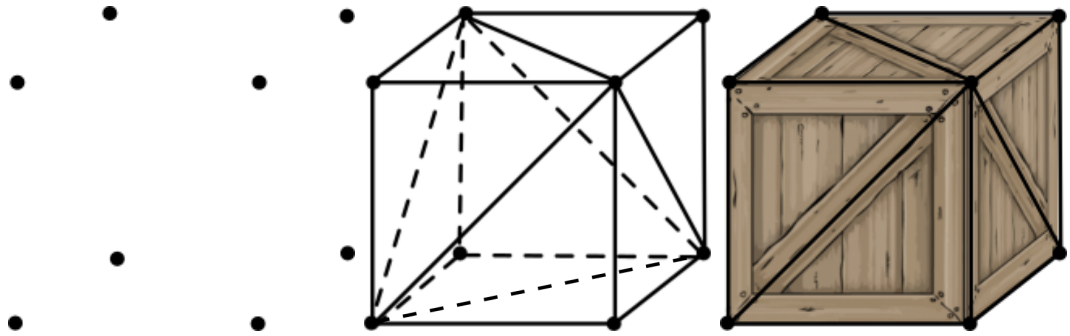
A 3D object is discretized and represented by a list of 3D points called vertices. The vertices may be linked to each other with primitives to define a surface. The primitives are generally triangles (a set of three vertices), but can also be quads (four vertices). A set of all connected vertices defines a **mesh** and represents, more or less




(a) (b)

(c)

***Figure 2.14:*** *a) Evolution of the pixel size with the distance with a conventional projector and a LBS projector. (image source: [FCM09]); b) Scan trajectory of the MEMS micromirror on the resulting image. (image source: [FCM09])*



(a) Vertices (b) Mesh (c) Textured mesh

***Figure 2.15:*** *Digital representation of a 3D object*

accurately, the geometry of the 3D object. Its color can be defined by vertex, but for more complex appearance this is generally done through textures. In this context, a texture is a 2D image that is intended to be mapped on a mesh. A virtual object can also have a material attached to it, which models interaction of light with the object, in particular its reflection properties.

Light sources are located in space by their coordinates, and they can be of different types: point (e.g. a candle), directional (e.g. the sun), surface (e.g. a neon tube) and so on. They interact with the meshes according to its material properties.

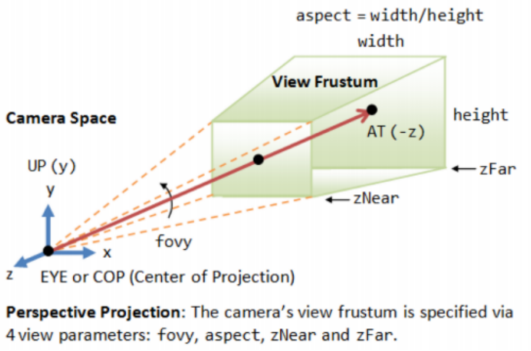
### Rendering

In order for the mesh to be rendered, at least one virtual camera is required. They can be modeled in various ways, the most common being the perspective camera model represented in Figure 2.16. A camera model defines a projection matrix that projects the 3D model on a 2D plane, along with its final color values.

Computing the final color from the point of view is called the rendering and is the most challenging and time-consuming step. The resulting image is computed from the intrinsic color of the object and light interaction from the light sources. In a 3D scene, the path of rays from light sources to the camera can be complex, with multiple bounces on several surfaces. A basic rendering technique is ray tracing, where light interaction is studied reversely: from the camera, toward the objects and then toward the light sources.

Figure 2.17 shows a basic configuration of a 3D scene in Blender[1], an open source 3D creation software.

[1]https://www.blender.org/ (accessed on 10/02/2019)




$$f = \cot(fovy/2)$$

$$M_{proj} = \begin{pmatrix} \frac{f}{aspect} & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & -\frac{zFar+zNear}{zFar-zNear} & -\frac{2\cdot zFar\cdot zNear}{zFar-zNear} \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

***Figure 2.16:*** *Perspective camera model and projection matrix (source: [Wet19])*

(a) (b)

***Figure 2.17:*** *(a) 3D scene in Blender, where the camera is highlighted; (b) A rendering of the scene from the camera (source: blender.org)*

#### GPU programming

A display is directly connected to the GPU of a computer for the pixels to be addressed and interaction with the GPU is possible thanks to dedicated libraries. In this section we introduce specific notions of GPU programming.

Small programs, called shaders, are intended to run in parallel on the GPU. In particular, two kinds of shaders are widely used and executed at different steps of the rendering pipeline: the **vertex shader** and the **fragment shader**. The vertex shader is executed for all (or a subset of) vertices of the scene. The fragment shader is executed later in the pipeline and acts on fragments, which are somehow comparable to the final pixel values in image space.

Shaders can modify the attributes of either vertices or fragments and can also use variables that are sent from the CPU. In particular, **textures** have specific locations on the GPU memory because they play a major role in computer graphics and often require rapid access from the fragment shader. Here, and for the rest of the thesis, the term "texture" simply stands for "2D images" that are specifically intended to be sent over the GPU. Note that a texture can also be 1D or 3D as it is simply an array of values.

**Framebuffers** are special textures that are filled during the execution and are usually used to store a rendering. The final image to display is stored in a framebuffer that is directly sent to the display. Framebuffers can also store off-screen renderings that can be subject to other transformations before being sent to the display. For example, the presence of a screen in a virtual scene can be rendered in two steps: the screen image is rendered in an off-screen framebuffer before being used for the final image.

In this thesis, we use OpenGL[2] which is a powerful library specifically designed for graphic applications in a C++ environment, but other libraries works quite the same. The shaders are written in GLSL programming language.

### 2.3.2 Computer vision: from the real world to the digital world

Computer vision is, in a way, the inverse of computer graphics: 2D images of the real world are taken and analysed to retrieve real world properties. It is clearly an anthropomorphic term as it designates a field where computers are provided the ability to "see", and most importantly "understand" what they see in order to take decisions.

It can be used to automatize task that can be achieved with the human visual

[2]https://www.opengl.org/ (accessed on 10/02/2019)

system (e.g. for waste sorting [RHS+11] or video monitoring [CLK00]), but not only, as computers are able to go beyond human capabilities for certain tasks. For example, sensors can see beyond the visible spectrum and analyse notably infrared [BBC+07] or ultraviolet [BAGM07]. Computer vision is often associated with artificial intelligence, as it allows vision-based decisions for robots, self-driving cars, and so on.

This field is largely multidisciplinary: it may require multiple scientific domains such as optics, graphics, image processing and deep learning, and has applications in various fields such as human-computer interaction, robotics, artificial intelligence or healthcare. In this section, we introduce the basic concepts of computer vision.

### Basic pipeline

**Acquisition** The first step of computer vision is to sample the illuminance of the environment, so that the computer can "see". Several sensors can be used depending on the applications, and for most applications, regular cameras are used. In this case, 2D images are at the root level of the application.

A real camera exhibits optical distortion due to the lens, and this might be undesired for some computer vision algorithm. In particular, radial distortion distorts straight lines at the edges of the images (see Figure 2.18). Optical distortion can be corrected by applying a transformation to the image. Denoting $(x_u, y_u)$ and $(x_c, y_c)$ the coordinates of the uncorrected and corrected images, respectively, the corrected image can be expressed as [Zha00]:

$$x_c = x_u(1 + k_1(x_u - x_0)^2 + k_2(x_u - x_0)^2 + k_3(x_u - x_0)^4 + ...) \tag{2.11}$$

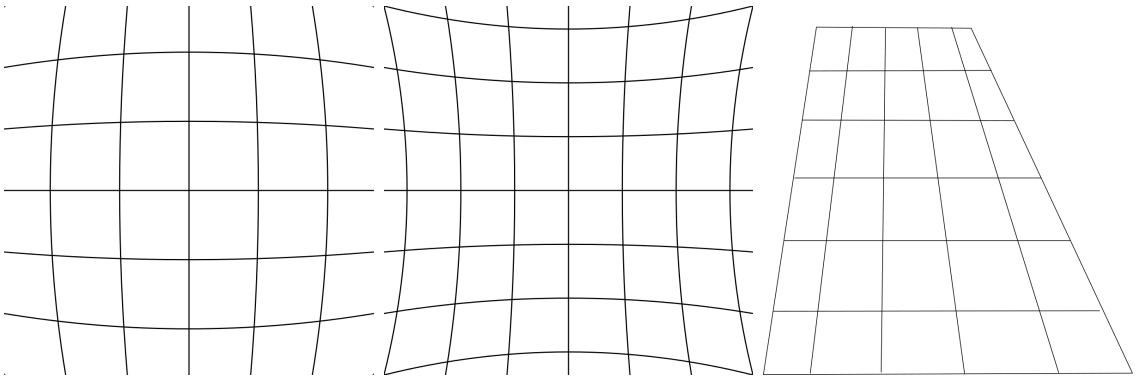
$$y_c = y_u(1 + k_1(y_u - y_0)^2 + k_2(y_u - y_0)^2 + k_3(y_u - y_0)^4 + ...) \tag{2.12}$$

where $(x_0, y_0)$ are the coordinates of the center of distortion and $k_i$ are the coefficients of distortion. Generally, 2 or 3 coefficients are enough to correct the images. These coefficients can be estimated with Zhang's algorithm [Zha00] by taking several images of checkerboards, consisting of a grid made of straight lines.

Note that in addition to radial distortion, a tangential distortion can also happen when the lens and the sensor are not parallel. It can be corrected with a different model relying on a similar approach.

Other types of distortions are purely geometric, like the keystone distortion. It happens when an object is taken from an angle, then the parts that are closer to the sensor will appear bigger. Figure 2.18(c) illustrates this with a tilted plane. A correction of the keystone distortion can be done directly on the image by a homographic transformation. The parameters of this transformation can be evaluated by specifying the correspondence between 4 uncorrected and 4 corrected points.

Before being analyzed, some other operations can be performed on the input images, such as noise removal, normalization, intensity correction, and so on.



(a) Barrel (b) Pincushion (c) Keystone

***Figure 2.18:*** *Different types of distortion. Barrel and pincushion distortions are intrinsic to the camera whereas the keystone distortion is extrinsic.*

**Image analysis** The analysis of images is the core of computer vision algorithms and allow the computer to "understand" the image. Our brain is very good at identifying quickly regions of interest in an image, such as lines, faces or objects. On the other hand, computers are intrinsically bad at it and image analysis algorithms aim at "reverse-engineering" the human perception in some extent. This step depends on the application, and we review some of the common tools for computer vision algorithms.

- Identifying objects of interest: This is generally done by binarizing the image through a series of filters. For example, Sobel filters (edge detection) or Hough transforms (line detection) can be used to identify objects of known geometries, and the same goes for the color. The result is a black and white image exhibiting the objects of interest, from there the coordinates are easily retrieved. This is generally a tough task that depends on the target application and the objects of interest. The identification can also be done using features that are known to be present in the objects of interest. For example, for eye tracking applications, visual characteristics, or features, of the pupil and the iris can be used to retrieve the eye position and/or orientation [MM05].

  Nowadays, computer vision is often based on neural networks and deep learning to identify the objects of interest [VDDP18]. Instead of trying to explicitly describe them through colors or geometric features, algorithms are trained on huge image databases to recognize specific objects like cars, animals, and persons.

- Tracking: It is sometimes required to track objects and this is done by analysing the coherence between a sequence of frames [TP06]. In particular, optical flow

techniques [SRB10] or Kalman filters [**?**] can be used to study the trajectory of a tracked object.

- Image matching: for some applications, for example image stitching (e.g. panoramas), it is sometimes necessary to identify the same features in two different images. This can be done through SIFT descriptors [Lin12], which are elemental features of images independent of scaling or orientation. In this way, the same features can be identified in different images and put into correspondence in order to extract notably the relative position of the camera. To build a panorama, a homography is applied on each image before stitching them together [Taa+16].

- 3D reconstruction: Computer vision allows to build a 3D model of real object by analyzing regular camera images. Notably, photogrammetry [MBM01] is a technique that consists of taking numerous pictures of the same object from different viewpoints. The images are matched together by finding common features, and the location of each viewpoint with regard to the object are estimated. From there, a 3D points cloud is constructed, and mesh reconstruction techniques are applied to obtain the final 3D model.

**Trigger events** Following the parallel with human behavior, after "seeing", and "understanding", a decision based on the understanding can be taken. This step is particularly useful for Artificial Intelligence systems and robotics. For example, in a self-driving car context, the presence of a pedestrian or a car may trigger brakes or a change of direction. Robots can also perform tasks (e.g. catching a ball [CLRS15]) with a disturbingly human-like behavior.

We do not insist more on this step because it highly depends on the application. The specific examples of interactions with displays are detailed in Section 2.3.3.

### Computer vision programming

OpenCV[3] is an open source library for different programming languages, including C++ and Python. It provides tools for computer vision-based applications, including camera calibrations, features extractions, tracking and so on. In this thesis, we use OpenCV in each of the prototypes as they all include cameras and computer visions algorithms.

[3] opencv.org

### 2.3.3 Examples of display applications

Computer graphics and computer vision can be used in combination to extend the possibilities offered by conventional displays, which is also the goal of this thesis. In this section, we review some examples that illustrate this.

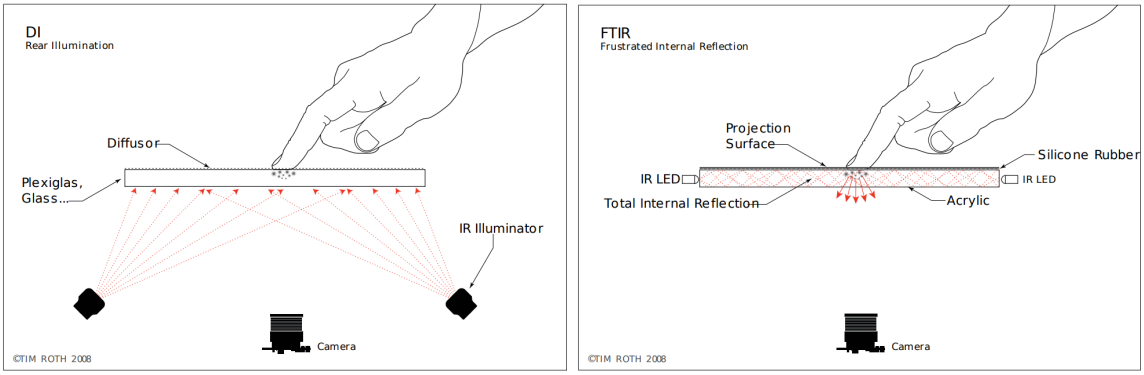
**Optical finger tracking** Cameras allow a projection surface to be multitouch with optical finger tracking [SBD+08], and we particularly do this in Chapter 4. Generally, an infrared light source lights the surface, either from the same side as the user (front illumination) or from the other (rear illumination) as represented in Figure 2.19(a). This requires a surface that is partially transmissive in the infrared in order for the IR light to reach the fingers and travel back to the camera. Another technique, illustrated in Figure 2.19(b) is the Frustrated Total Internal Reflection. It is based on total internal reflection, as explained in Section 2.1.4 and thus requires a clear surface, which is generally covered by a diffusing film so that the projected image remains visible. In the FTIR approach, infrared LEDs are located at the edges of the surface and rays are propagated along the thickness. When a finger is in contact with the surface, the index of refraction is locally changed so that the total reflection condition (see Section 2.1.4) is no longer met and the light passes the interface and gets scattered by the finger. This allows a better detection by ensuring that *only* light from the finger reaches the camera.

**Eye and gaze tracking** Eye or head tracking can be used in combination with a display. This can allow notably the virtual camera for the rendering to move with respect to the location of the user in front of the display [CZF09]. This can be used to move around a displayed 3D object and exhibit motion parallax (see Section 3.1).

Gaze tracking is a subset of eye tracking and aims at analysing the eyes orientation to estimate the coordinate in screen space of the observed point. Commercially available eye tracker (e.g. Tobii™[4]) are able to determine the gaze location after a calibration process that consist in estimating the position of the sensor with respect to the display. This information can be used in various ways, notably to improve human-computer interactions [PB06, ZJ04], perform foveated rendering [BTV+17] or to improve the interactivity of disabled people [BL11].

**Body interaction** It is also possible to interact with a display with remote gestures that are analyzed with a camera, in particular it had been showed that this improves greatly the user experience in video games [SCP11]. Personally, a memorizing experience as a child with such a system was with the EyeToyPlay™ video game in 2003.

[4] www.tobii.com/




(a) (b)

***Figure 2.19:*** *Optical finger tracking techniques (image source: [SBD+08]). (a) Rear Diffused Illumination setup. The camera tracks the changes of brightness coming from scattering on fingers. The front illumination is the same with the light sources on the same side of the finger and by tracking shadows. (b) Frustrated Total Internal Reflection (FTIR) is based on total internal reflection. A finger on the surface locally changes the index of refraction and scatters light toward the sensor.*

The player was immersed in a virtual environment and interacted with the virtual data through gestures. Figure 2.20(a) show two children remotely "cleaning" the display. This was an early algorithm that did not recognize body pose but consisted of comparing directly the pixel values between two frames.

The technology has evolved in video game applications to give notably the Kinect™ by Microsoft. The technology was based on an infrared projector projecting a structured pattern that is captured with an infrared camera. The deformation of the pattern gives information about the depth of the real scene, and from there the body pose could be retrieved (Figure 2.20(b)). This computer vision vision technique is called structured lightning [Ger12] and allows to generate depth maps and 3D models. This had allowed notably dancing video games, but was also a great tool for computer vision research. We refer the reader to reference [HSXS13] for a review of computer vision algorithms and applications based on this device.

Microsoft has stopped selling the Kinect™ in 2017[5] but is in the process of commercializing the Azure Kinect[6] specifically designed for Artificial Intelligence. Another example of such infrared tracking based device is the LeapMotion™[7] which is

[5] `https://www.fastcompany.com/90147868/exclusive-microsoft-has-stopped-manufacturing-the-kinect`

[6] `https://www.microsoft.com/en-us/p/azure-kinect-dk/8pp5vxmd9nhq`

[7] `https://www.leapmotion.com/`

more specifically designed for gesture interactions.

**Augmented and virtual reality** Computer vision allows the integration of real world data for the rendering of a virtual scene, and thus has obviously a lot of applications in the augmented reality (AR) and virtual reality (VR) fields. We see notably the arrival of AR/VR headsets, as we detail later in Section 3.2.2, and computer vision allows natural interaction with the virtual data. This also stands for video-based AR applications.

First, the body pose tracking and gesture recognition that we discussed earlier can be used notably to grab virtual objects with the hand or avoid them with the full body. For AR specifically, the act of placing a virtual object on a real one (see Figure 2.20(c)) requires an estimation of the 3D configuration of the user environment. For that, AR markers can be used [PSB13]: they are printed features located in the real environment. They have known sizes and shape to they can be easily detected by software, and the homography estimation gives the relative position of the camera. This can also be done in unprepared scenes like feature-based detection [LH08] or Simultaneous Localization And Mapping (SLAM) [BDW06]. The latter allows an on-the-fly reconstruction of a coarse 3D mapping of the room and is particularly used in robotics application.

AR and VR rely also on non camera-based sensors that help extracting information from the real world (e.g. inertial sensors), but these do not belong to the computer vision field.

**Spatial Augmented Reality (SAR)** Camera-projector systems are also used for augmented reality, and they belong to the Spatial Augmented Reality (SAR) field. SAR brings digital augmentations directly on the real world by the mean of projectors. This changes our perception regarding displays, because every object is then perceived as a display.

This generally requires a camera, which is calibrated with regard to the projector's position and orientation, allowing direct interaction between the real and virtual worlds. A typical example is the PapARt [LH12] (see Figure 2.20(d)), where a camera tracks the position of a paper for the projector to project an image on it, whatever its position. This allows notably the real drawings to be augmented. The camera also allows natural touch and gesture interactions thanks to computer vision algorithms.

## 2.4 Conclusion

In this section, we reviewed the essential background notions that lead our research and will help the reader understanding the works presented in this manuscript.

(a) (b) (c) (d)

***Figure 2.20:*** *a) The EyeToy Play$^{TM}$ video game integrates the player in a virtual world that can be interacted with (image source: playstation.com); (b) The Kinect$^{TM}$ sensor can retrieve the skeleton of the user by generating a depth map and analyzing it (image source: [GCR15]); (c) Computer vision allows a virtual object to be placed on a physical table (image source: [LH08]) (d) PapARt is a projector-camera system relying on computer vision that turns a paper into an interactive display (image source: [LH12])*

Light will be described independently as rays or waves to explain its behavior. In particular, the concept of interferences, holography and HOE will be useful for Chapter 6 where we describe our 3D transparent display.

We reviewed the basic display technologies and saw that software considerations are necessary for a display to work, notably to address pixel values from a GPU and generate the content. Computer vision can extend a display's capabilities by taking into account the real environment, for example to give a display multitouch capabilities as we do in Chapter 4. Augmented Reality requires both computer vision and computer graphics, as 3D virtual objects are rendered depending on real world parameters.

This combination of hardware and software provides the possibility of creating more and more sophisticated displays. Different elements detailed in this section are used to create 3D displays as we describe in the following chapter.

# Chapter 3

# State of the art on 3D displays

In Section 2.2, we reviewed existing technologies for the displays of our daily life, that is to say conventional displays like a regular TVs or smart phones. They now have achieved a very high quality with impressive resolution and contrast ratio, reproducing reality with high fidelity. However, they fall upon the category of 2D displays: they provide a single view of the scene and most of depth information is discarded.

In this chapter, we describe technologies related to 3D displays. We use the term "3D display" as a generic term encapsulating displays that are able to increase depth perception compared to 2D displays. Depth perception can be simulated by recreating some of the *depth cues* involved in our visual system, as we will review in Section 3.1. We then briefly describe glasses-based solutions and their limitations in Section 3.2 before going deeply into the subject of glasses-free 3D display in the remainder. We will clearly classify them into different categories: autostereoscopic displays (Section 3.3), volumetric displays (Section 3.4), light-field displays (Section 3.5) and holographic displays (Section 3.6).

## 3.1 Depth cues

An eye is a complex sensor, and the brain processes the visual information to acquire a 3D representation of a real world scene. This representation is based on visual cues, called *depth cues*, that have been reviewed in both psychological [HR08] and 3D displays [MPWL13] communities. We introduce the depth cues in this section and classify them into two categories: monocular and binocular cues.

### Monocular cues

Monocular cues provide depth information when viewing a scene with a single eye. Most of the monocular cues can be used to interpret 3D information of both natural real-world scenes and 2D representations of 3D scenes, like on a regular flat 2D display or a photograph. The main 2D monocular cues are:

**Perspective**: This is the property of parallel lines to appear converging at an infinite distance. This has been theorized in arts in the 15th century [Pri70] to give depth effects in paintings.

**Occlusion**: When an object occludes another object, the brain interprets it to be closer.

**Familiar size**: The *apriori* knowledge that we have on the size of familiar objects makes it possible to locate objects in space.

**Relative size**: The different perceived sizes of objects that are supposed to be the same size, for example, provides information on their relative depth.

**Aerial perspective**: When an outdoor scene provides a large depth, the atmosphere tends to make objects hazier with growing distance.

**Texture gradient**: Repeated texture on objects appears to have greater frequencies with the distance.

**Blur**: Due to the imaging optics of the eye or a camera, images may exhibit depth of field. Objects that are too far away from the focal plane are blurred.

**Depth from motion** is another effect due to the relative motion between the sensor and the scene. When an object approaches the observer, its projected size increases. This can be only perceived in a moving scene, not in a still picture.

Figure 3.1 illustrates some of these effects.

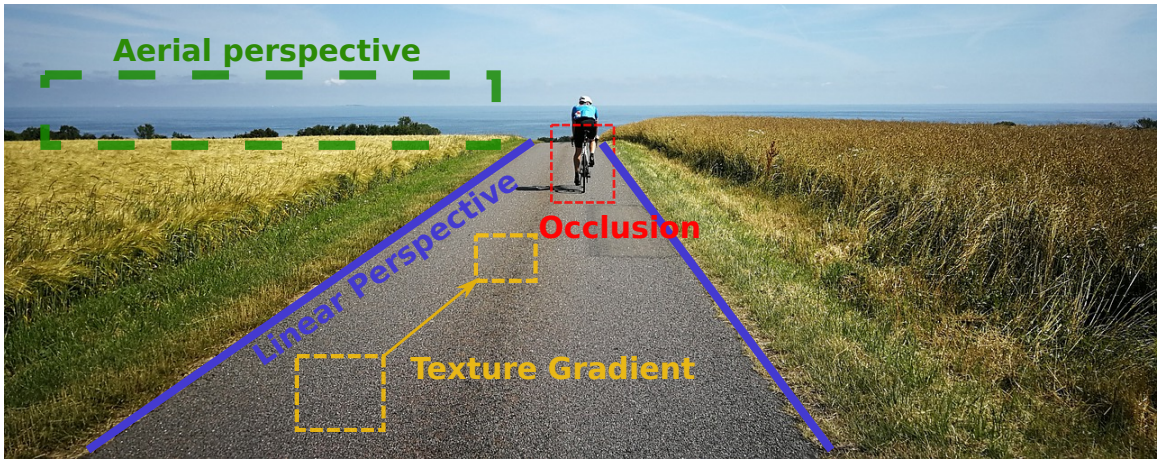



***Figure 3.1:*** *Illustration of some of the monocular depth cues*

Other monocular cues occur only when perceiving a real scene and not on 2D displays:

**Accomodation** is an *oculomotor* cue. It is the ability of the eye to adjust its focal length. This is the same principle as for cameras: due to the pupil size, several rays from a same 3D point enter the eye. The eye has to adjust its focal length in order to image a point sharply on the retina. This focal length is sensed by the brain through the involved muscles, and it is used to estimate the distance of the object that the eye focuses on.

**Motion Parallax** happens when the head moves relatively to the 3D scene, and the relative motion between close and far objects provides information regarding their relative distance. A 2D display can reproduce only a limited motion parallax: the motion of objects from a static viewpoint, while a moving viewpoint can be reproduced through interaction (e.g. for a video game). In real life, motion parallax is also associated with head movements and this is mostly what 3D displays seek to reproduce and the way we use this term in this document.

### Binocular cues

Depth perception is improved by the fact that we have two eyes.

**Binocular disparity** is the fact that the left and right eyes perceive an image with a slightly different perspective. The brain interprets these differences to build a mental 3D representation of both 2D images.

**Convergence**, or simply vergence, is another oculomotor cue that is a direct consequence of binocular disparity. Eyes physically converge toward close objects to merge their images, and the brain receives signals about the involved muscles, thus providing depth information.

### Fooling the brain

Monocular 2D cues are widely used when looking at photographs or movies because they are directly captured by the sensors. Computer generated images, for animation movies or video games for example, include camera models and pipelines that mimic the real world in order to reproduce these depth cues. And indeed, if a 3D scene is correctly generated, depth can be easily interpreted even on a 2D display. This is done through several steps in the rendering pipeline: perspective camera model, depth test, mip-mapping, simulated depth of field, and so on [PJH16]. Hence, 2D displays are able to exhibit monocular depth cues and an artificial motion parallax.

3D displays seek to reproduce additional depth cues, and in particular binocular cues. By showing left and right eyes images with a slight parallax, binocular disparity and convergence improve greatly the depth perception and user experience

in general [YJXP10]. This is on what glasses-based displays, which are the most democratized type of 3D displays, are based as we will to see in Section 3.2.

However, binocular cues should not be considered as the only requirements for a 3D display. Indeed, 3% to 15% of the population are subject to stereoblindness [Fre12], yet they are able to have a better depth perception of the real world than with a 2D display. A good 3D display should be able to generate additional monocular cues. Notably, the importance of motion parallax in depth perception for example should not be minimized [RG79]. Accomodation is also a strong monocular cue, but it is harder to implement on a display, and 3D displays often suffer from what is called the *vergence-accomodation conflict* (VAC) [YIMT02]. As shown in Figure 3.2, the VAC happens when eyes are forced to converge to a point that is outside the display plane, due to the binocular disparity, while each eye is also forced to accomodate on the display plane, where the light really comes from, to obtain a sharp image on the retina. This gives contradictory information to the brain and creates a visual discomfort.

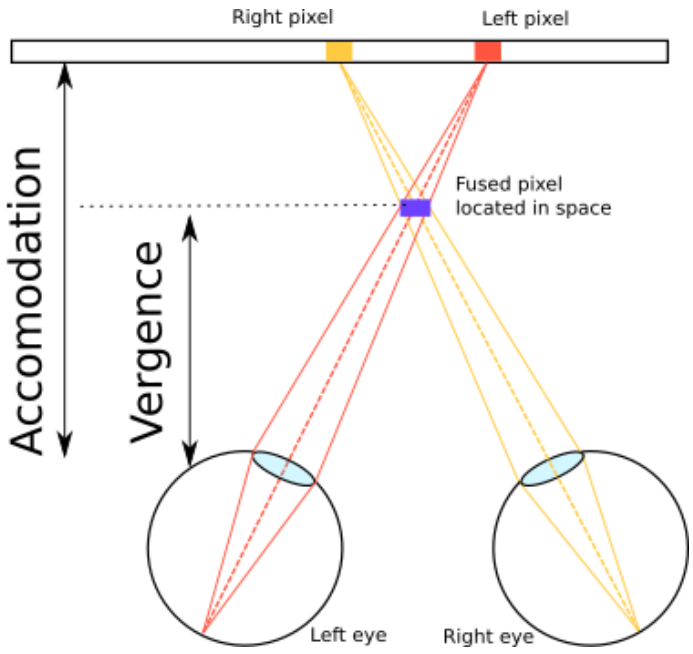



***Figure 3.2:*** *Vergence-Accomodation Conflict (VAC)*

## 3.2 Glasses-based 3D displays

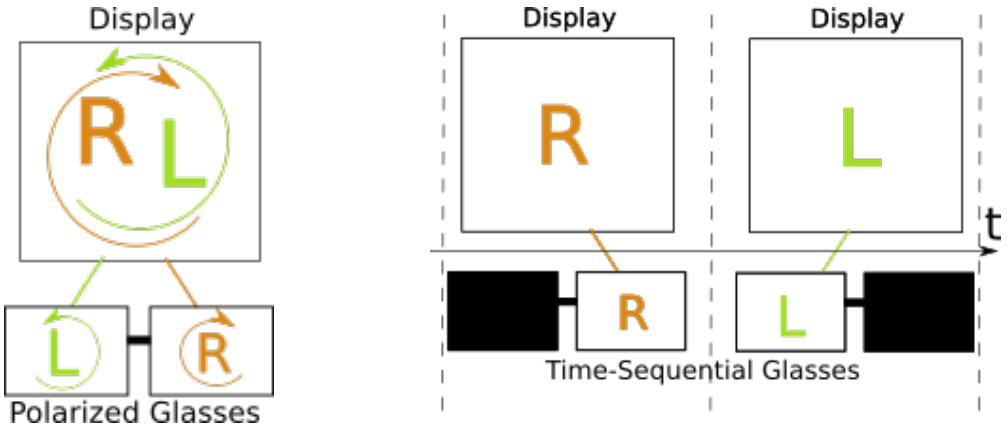
### 3.2.1 Stereoscopic glasses

The most common way of sensing a three dimensional image on a display is with passive or active glasses. These approaches aim at creating binocular cues by sending different images to the right and left eyes. The image separation is performed by the glasses. This has been democratized notably with cinema or home entertainment.

**Passive glasses** were the first type of 3D glasses with the pioneering work reported back to 1853 [SHS+08]. Among them, the anaglyph technique requires the left and right image to be encoded into different spectra, and the image separation is done with color filters [Bei81]. For years, it was the only way to perceive a 3D illustration, and printed anaglyphs were used since the 1920's notably in magazines and comics. Anaglyphs do not require any specific display, so we do not consider them as 3D displays. Nowadays, this technique is only marginally used, and movie theaters have opted for polarization-based glasses. This technology was introduced in the 1930's and became popular in cinema in 2005 [Wal13]. The principle of work of today's polarization stereoscopic glasses is depicted in Figure 3.3(a). The projector projects simultaneously images for the left and right eye that are circularly polarized but each in the opposite direction. They are reflected with a maintaining polarization screen toward the users' eyes. The left and right glasses are circular polarizers with opposite directions that only allow light with the associated polarization to pass through. Generating these polarization states with a flat TV screen is a little bit trickier but polarization films can do so by interlacing the polarization states of the pixel lines [LTL08].

**Active glasses**, also called field-sequential, are based on time-multiplexing. Both eyes have a high speed shutter in front of it that is synchronized with the display. The display alternatively shows the image intended for the left eye while the left shutter is opened, then rapidly switches on the other image for the right eye, and so on, as illustrated in Figure 3.3(b). They were first developed in 1922 with mechanical rotating shutters [Lip91], today the shutter is done with LCDs [MC11]. This approach works similar for projectors and televisions, as the only requirements are a high frame rate and synchronization between the glasses and the display. They thus are preferred for home entertainment compared to polarization-based solution, partly because they do not have to split the resolution in two like polarized TV. However, field-sequential approaches are not adapted to a large audience because the shutter glasses are more expensive than polarized ones, and so movie theaters still use polarization-based glasses.

Glasses-based 3D displays are **stereoscopic** displays, that is to say they provide binocular vision and convergence.




(a) (b)

***Figure 3.3:*** *a) In polarization passive glasses, two images are superimposed with two different polarization states and selected by the spectacle; b) Temporal active glasses use a high speed shutter synchronized with the display to alternatively hide each eye.*

### 3.2.2 Head-Mounted Displays (HMD)

More recent approaches allow to experience a three-dimensional scene through headsets, such as dedicated Virtual Reality (VR) or Augmented Reality (AR) headsets that have the ability to display different views of a scene to the left and right eyes separately. VR and AR headsets recently became commercially available to the general public.

Probably today's most popular examples of VR headsets are Oculus Rift [Ocu16] and HTC Vive [VIV16]. As shown in Figure 3.4(a), they use a regular 2D display in front of each eye, associated with a positive lens that magnifies the image at a further distance. It allows total immersion in a virtual world perceived in 3D. As for stereoscopic glasses, they provide binocular vision and convergence, but above all they have the ability to create motion parallax in any direction thanks to external and/or internal tracking systems.

In AR headsets, a user is able to see both the real world and computer-generated images. Their design is generally based on the imaging of a 2D display by a transparent combiner located in front of each eye. The Meta$^{TM}$ headset [Pul17] for example images the reflection of a LCD screen on a concave half-mirror (Figure 3.4(b)), while Hololens$^{TM}$ [KC17] and Digilens [WGP18] use Holographic Optical Elements (HOE) to couple-in, expand the pupil and scatter out the light in a guide (Figure 3.4(c)). The HOE are not holographically recorded but printed so they are rather called Diffractive Optical Elements (DOE) or diffraction gratings, and more detail about the use

of this technology for near-eye displays can be found in [Lev06].

AR headsets can be monoscopic, like Google Glass™ [OMHC14] or Optinvent [SDB12], so of course they do not have any of the binocular cues but can still provide motion parallax. Even though they might be able to reproduce the illusion of 3D presence thanks to head tracking or computer vision techniques, such as SLAM (Simulateous Localization And Mapping) [ESC14], we do not classify them as 3D displays, as they do not provide binocular vision of the digital world. They can also be stereoscopic, like Meta™, Magic Leap™ [BMA16] or Hololens™, and in this case they provide the same depth cues as virtual reality headsets and we consider them as 3D displays.

Current technological challenges in the VR/AR headset aim at increasing resolution and field of view (FOV), solve the VAC and balance a good image quality for a low latency. To solve the VAC, eye tracking can be used with a tunable lens [LCH08] that adjusts the accomodation distance. Eye tracking allows also to perform foveated rendering [BTV+17] that lowers rendering times. We draw attention to the reader that when referring to a HMD, the FOV does not describe the scattering angle of a pixel but the apparent angle of the display area from the eye (Figure 3.4), so it is basically the image size. This distinction is necessary because in a HMD, the eye does not move (or only slightly) with regards to the display unit.

A hybrid approach consists of rendering a stereoscopic view of the real world in an opaque HMD. For example, the Lenovo Explorer ™headset [Len17] shows a hardware that is quite similar to a VR headset with two cameras at the eyes' location. In this way, the user perceives a binocular view of the real world along with augmented content, simulating a see-through effect as in a AR headset. Unlike AR headsets, the user does not have a direct view of its surroundings but a stereoscopic rendering of it: this lowers the perception of the real world but allows notably true occlusion and fewer misalignment issues between virtual augmentation and true content.

### 3.2.3 Discussion

Stereoscopic glasses are a cheap and effective way of bringing 3D life to movies or video games. Indeed, they only require two (virtual or real) cameras for the content creation, and existing displays can be easily adapted to display stereoscopic images, when the image separation is done by the glasses. However, the user experience in this kind of solution is not yet fully satisfactory as 3D movies are reported to provide some sickness [Sol13]. Visual sickness because of the VAC, but also motion sickness due to the contradictory signals between a moving 3D scene while the spectator is personally still, just like transport sickness. The commercialization of personal glasses-based 3D displays has failed [3DT19], and the movie industry does no longer focus on 3D production [3Di17] due to several factors, including the aforementioned

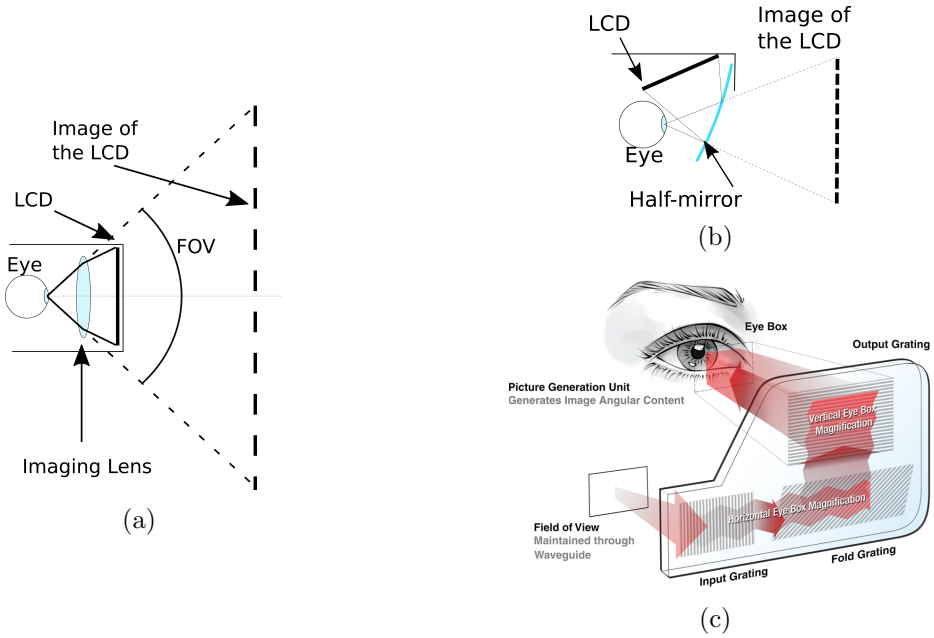



***Figure 3.4:*** *a) Optics of a VR headset: a LCD display is imaged by a lens; b) Optics of a half-mirror-based AR headset: a LCD display is imaged by a flat or convex half-mirror; c) Optics of waveguide-based AR headsets using diffraction gratings (source: [WGP18])*

sickness, the cost of the glasses (for the active type), and the annoyance to wear them.

The same goes for HMDs. Motion sickness is stronger for the VR technology as the user is totally immersed in a moving virtual world. Moreover, all types of headsets are generally heavy and this causes a discomfort [YCX+18]. Another form of discomfort comes from the generated heat [WCH18].

As a conclusion, even though they can achieve high quality and immersion, glasses-based 3D displays causes a discomfort that partially comes from the eyewear themselves. They have reached the general consumer market, but the enthusiasm is limited. They are not perfectly adapted neither for public demonstrations, as they require preparation time and assistance. For these reasons, glasses-free solutions have been largely investigated for decades and they improve some of these aspects.

## 3.3 Autostereoscopic displays

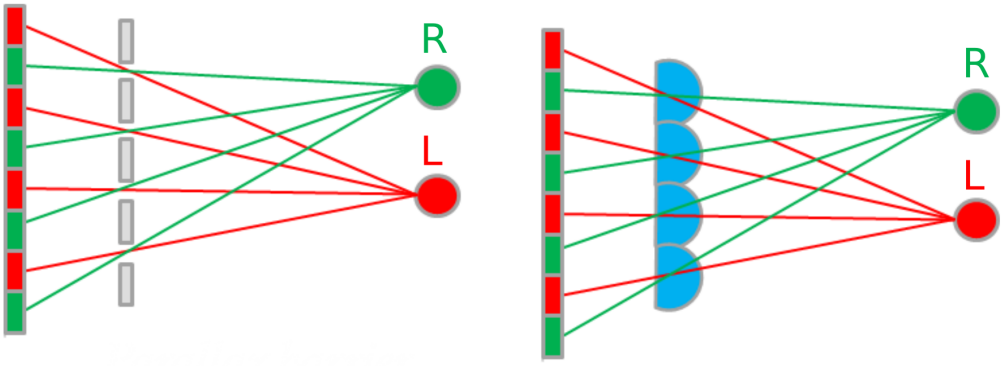
In an autostereoscopic display, the left and right eye image separation is performed by the display itself rather than with portative optics. Discrete views of a 3D scene are generated and distributed toward different viewing zones. If the viewing zones are correctly located, then each eye of the observer sees a different image, providing binocular disparity and eye convergence. If more than two views are generated and distributed in front of the display, then the observer can move his head to select the view to see and thus perceive motion parallax. Most of autostereoscopic displays rely on the fact that our eyes are separated in the horizontal direction, and they focus on providing horizontal-only parallax, opposed to full parallax. To achieve this, several methods are possible.

### 3.3.1 Parallax and lenticular displays

The most popular methods are **parallax barriers** and **lenticular lenses**.

Figure 3.5(a) shows the general principle of parallax barriers-based autostereoscopic displays: a vertical grid is located in front of a LCD display, so that from a single point of view, only one vertical pixel line over two is visible and the others are hidden behind the barrier. If the barrier pitch and distance from pixels are correctly chosen and the user is located at the correct distance from the display, then only half the pixel grid is visible from one eye and the other half visible from the other eye. This allows each eye to be addressed with different images.

Following the same idea, replacing the parallax barrier with lenticular lenses, i.e. a dense array of small vertical cylinder lenses as shown in Figure 3.5(b), achieves the same goal but allow to collect more light from the pixels [vBPF96].




(a) (b)

***Figure 3.5:*** *a) In the parallax barrier approach, a grid separates the image in vertical lines alternatively visible from one eye or the other.; b) The lenticular lenslet approach works the same except that the pixels are imaged by tiny lenses.*

In both approaches, the horizontal resolution of the 3D display is half the horizontal resolution of the LCD display underneath. A popular technique, used notably in the Nintendo 3DS$^{TM}$ display, is to slant the lenticular sheet or parallax barrier, so that the loss of resolution is distributed in both vertical and horizontal directions [VB99]. This also reduces Moiré effects due to superimposition of both the barrier grid and the pixel grid.

Limitations of these autostereoscopic displays are that they show a stereoscopic image from only a single point of view. Eye tracking can be implemented to update the image according to the user's position [BDM$^{+}$00] but this is generally limited to one user. Autostereoscopic displays can also generate more than two viewing zones but they sacrifice even more resolution. In this case, they are sometimes called automultiscopic [EDF$^{+}$16, DDD$^{+}$14] or multiscopic [Kar12, KA03], but in this thesis, we will still call them autostereoscopic for consistency with most of the existing literature.

The visual quality of parallax or lenticular-based autostereoscopic displays do not reach the quality of 2D displays [BBdH05].

Projectors can be used instead of a LCD display. Matusik et al [MP04] associated an array of projectors, a screen and lenticular sheets, either in front or rear projection configuration. In the rear projection case, a lenticular sheet focuses the light of each projector on a vertical line, the screen transmits it and another lenticular sheet casts the pixel line in a specific direction as in a regular lenticular display. In the front

projection case, a retroreflective screen is used, then only one sheet of lenticular lenses is required.

Theses displays were inspired by the integral photography technique introduced by Lipman [Lip08]. A display can use a 2D array of elementary lenses instead of a 1D array of cylinder lenses, in this case it is called an integral display [AJ01] and can exhibit vertical motion parallax in addition to horizontal. These displays, however, have to sacrifice even more spatial resolution.

### 3.3.2 Dynamic parallax barrier

A barrier or lenticular sheet is manufactured according to the desired properties, such as the observation distance range or the number of views. These properties are fixed once the barrier is manufactured and positioned. Barriers can also be dynamic to release these constraints. A dynamic parallax barrier can be implemented by physically moving a parallax barrier [SH98] or by using a LCD module as a barrier [IYS93]. In the latter, pixels in "ON" state let the light pass trough it while pixels in "OFF" state block the light. The barrier parameters (position, orientation and pitch) can then be changed according to the desired observation distance range and number of views. Moreover, the display can be turned in 2D mode when no barrier is displayed. Time-multiplexing the position of the parallax barrier, and thus the generated image behind it, improves the perceived horizontal resolution [PPK00].

Head tracking can be used to adapt the parallax barrier parameters according to the user's head position [PPK00, PKG+07]. Eye tracking is implemented in the latest version of New Nintendo 3DS™ with a dynamic parallax barrier that allows the display to be switched between 2D and 3D.

### 3.3.3 Random Hole Displays (RHD)

A parallax barrier does not have to be a regular grid. Nashel et al. [NF09] show that a mask consisting of random holes can also be used as a parallax barrier. The user needs to be tracked, so that the visible pixels can be determined at each frame in order to generate the correct image. In a regular parallax display, if the user is at a wrong location, then several entire lines of pixels are perceived wrong, whereas in a random hole display, the error pixels are randomly distributed along the whole resolution, making them less perceivable. Time-multiplexing is also used in order to have more degrees of freedom for conflict correction, especially when there is more than one viewer. This work has been extended to a tabletop display [YSF10] and a see-through display [KMCS12b].

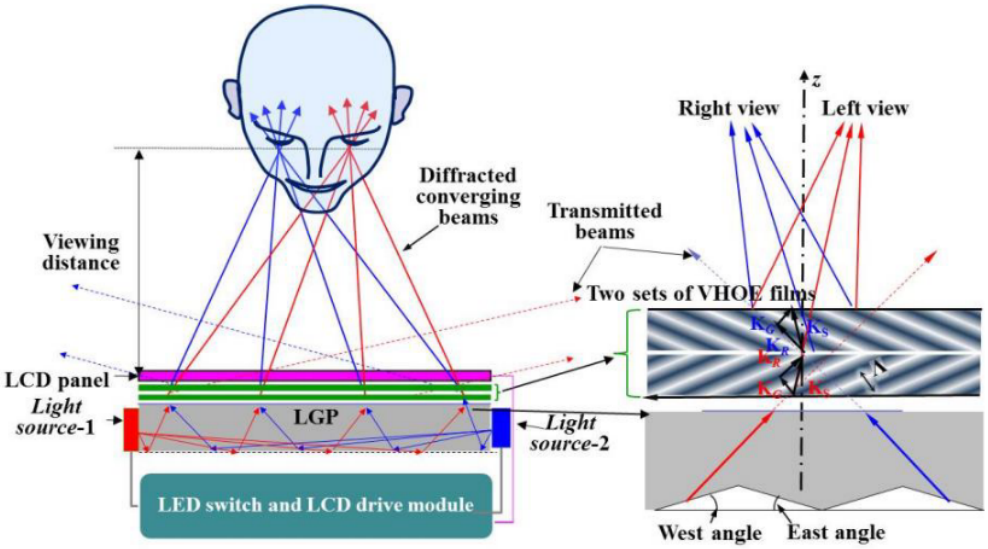



***Figure 3.6:*** *A light guide and HOE solution to provide a directional backlight: LED sources emit light in a guide from different sides and a HOE directs the rays toward different viewing zones (source: [HBF+14])*

### 3.3.4 Directional backlight

Different views can also be spatially separated with a directional backlight. In a regular LCD display, the backlight is a strong uniform light source located behind the liquid crystal panel, which casts rays in the whole hemisphere. A directional backlight, on the other hand, emits light in a specific direction. Autostereoscopic displays based on directional backlight are usually time-multiplexed, so the images are generated at a high frame rate and the user sees the left and right views at the same time thanks to the persistence of vision. This way, the full resolution of the display can be used, but time-multiplexing introduces flicker artifacts and lowers the perceived brightness. A directional backlight can be built using various methods. One of the earliest displays of this type is the Cambridge display [DML+00] where a CRT is associated with lenses in order to directionally light a LCD panel. A light guide can also be used to light a LCD with a directional backlight: as depicted in Figure 3.6, two light sources can be located at opposite edges of a light guide with a holographic optical element that couples light out of the guide toward different directions for the left and right eyes [HBF+14]. A wedge light guide can also be used as a directional backlight by adding LED sources at different positions of the entrance side [MT05]. A more atypical backlight can also be built using a hemispherical mirror and LED assembly, in order to create a 360° autostereosopic tabletop display [SS16].

### 3.3.5 Super Multi-View displays (SMV)

Autostereoscopic displays are generally subject to the VAC that we explained in Section 3.1. The eye accomodates to make different rays from a same point converge on the retina, so if an autostereoscopic display is able to send slightly different perspectives of the same scene to the same eye, then it is forced to adjust its focal length until the focused object is sharp as illustrated in Figure 3.7. In this case, the displays are called *Super Multi-View* (SMV).

Tataki et al. [TTN11] developed a SMV display using a regular 16-view lenticular display that they adapted to display 8 views per eye. In such displays, eye tracking can be implemented to move both groups of views along with eyes position.

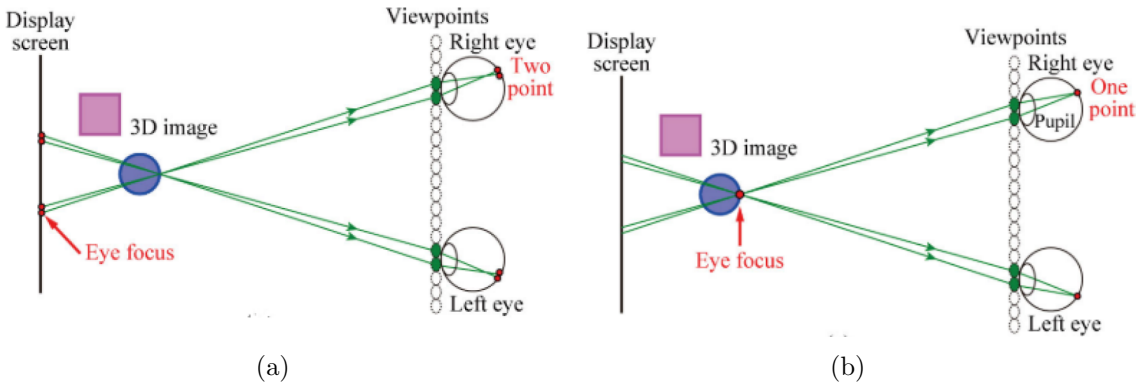



***Figure 3.7:*** *a) If the eye focuses on the screen, the image on the retina is blurred. b) If the eye focuses on the 3D object, the image is sharp (source: [Tak14])*

Multiprojection-based SMV displays are also reported [Tak14]. They also use a projector array and a vertical diffuser to provide horizontal parallax. Their prototype uses 128 small projectors arranged on a slanted 2D grid as shown in Figure 3.8. The vertical shift has no effect because of the vertical diffuser, so this simply increases the horizontal density of viewpoints.

## 3.4 Volumetric displays

A volumetric display is designed to show information in a volume. This is usually done by discretizing the volume into several stacks of planes, for example with multiple screens or a single high-speed moving screen. The displayed images are 2D slices of a 3D model. In this kind of display, the pixels are distributed in a volume, and so we rather call them voxels. Due to their design, volumetric displays generally cannot provide occlusion capabilities because all the generated voxels are always visible.

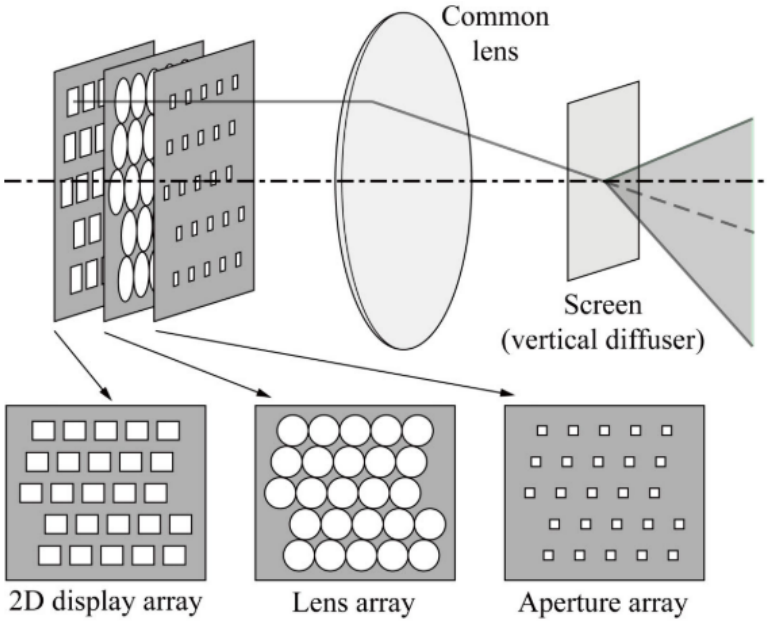
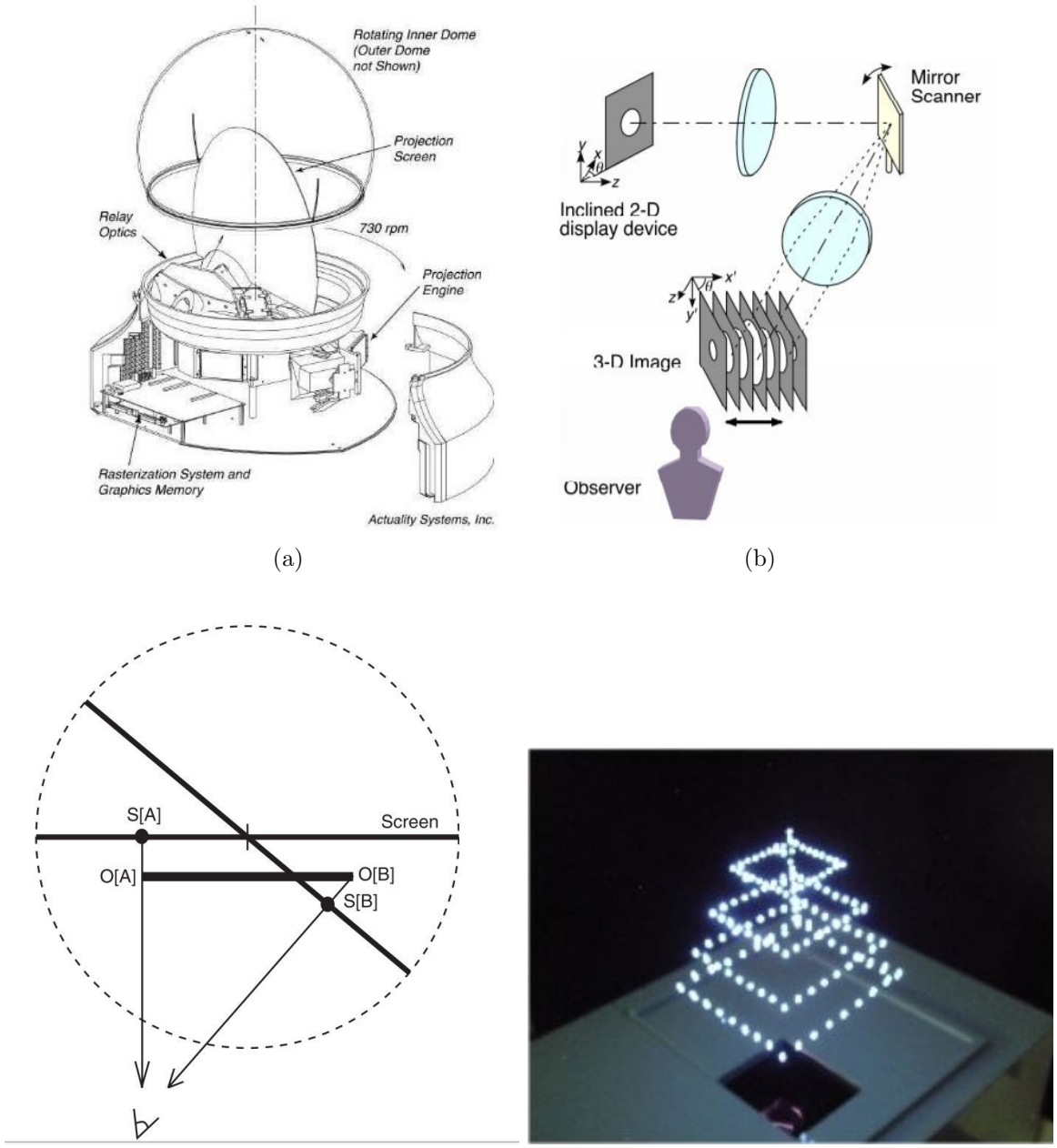



***Figure 3.8:*** *2D slanted array of projectors to produce a denser horizontal parallax (source: [Tak14]).*

Perhaps the most popular and representative research in this field is the Perspecta$^{TM}$ display [FNH+02] by Actuality. Their system is represented in Figure 3.9(a): it consists of a high-speed rotating projector and a projection screen. The projector is synchronized with the rotation and thus can display the correct slice of the 3D model.

Sadovnik et al. [SR98, lig05] created the DepthCube: a high-speed projector associated with a stack of several LCD layers that act as projection screens. The LCDs rapidly switch from a transparent state to an opaque state. In the opaque state, the projected light is scattered in all directions. Only one layer at a time is visible, and the projected image is synchronized with the layer position. Sweeping all LCD planes rapidly creates the volumetric rendering. Quite similarly, other methods use stacked LCD to display, this time simultaneously, images of different depths [SOT+04, STU+00, BCP+08].

Another method is to image a single LCD with an optical system, and sweeping this image at various positions to create the volume. In this case, a virtual image is created and thus it appears floating in mid-air. Figure 3.9(b) demonstrates the concept of such a display as presented by Miyazaki et al. [MSSM06]: 2D images are created with a regular LCD display, and a lens assembly creates a virtual image of it. A scanning mirror moves and the 2D image is modified accordingly to display the




(a) (b)

(c) (d)

***Figure 3.9:*** *a) The Perspecta$^{TM}$ display design: a rotating screen with a synchronized projector (source: [FNH+02]).; b) Volumetric display system based on three dimensional scanning of inclined optical image (source: [MSSM06]); c) Modification of the Pespecta display to make it occlusion-capable: the drawing shows how the ends of a line are created on two screen positions (source: [CNH+07]); d) Mid-air volumetric image made by plasma generation by a scanning femtosecond laser (source: [KUY06]).*

correct slice of the 2D model. The authors have later developed a prototype replacing lenses by a dihedral corner reflector array [MHM+13] or a roof mirror array [MMM15] to reduce optical aberrations.

Alternatively, it is also possible to use micro bubbles [KHH17] or smoke [TSN+10] as volume scattering media and to project light on it.

A more complex approach is to create volumetric images by making a femtosecond laser focus on a mid-air point so that a plasma is created. The point becomes a point light source that is visible in all directions, as shown in Figure 3.9(d). Fast scanning a volume to create every point of a 3D model creates impressive volumetric images in mid-air [KUY06, OKH+16].

Another complex approach to design a mid-air volumetric display are optical trap displays [SNS+18]. A cellulose micro particle is optically trapped in air and illuminated in red, green and blue depending on the voxel value. The particle scatters light in all directions and is scanned through a volume.

Finally, the work of Cossairt et al. [CNH+07] is an example on how to adapt volumetric displays for providing occlusion capabilities. They replaced the reflective screen of the Perspecta$^{TM}$display [FNH+02] with a vertical diffuser. If the screen is static, then the observer perceives only a small portion of it. When the screen is spinning, the full image as perceived from a single viewpoint is generated by several screen positions. The question whether this display can still be called volumetric arises, because it does not show slices of a volume. Instead, full images from multiple viewpoints are generated and distributed over all screen positions. Moreover, contrary to a volumetric display, the point source position in the displayed volume does not represent the position of the point in the virtual scene, resulting in depth inconsistencies. This display may be more related to light-field displays.

## 3.5 Light-field displays

### 3.5.1 Introduction to light-field

The concept of light-field has been introduced in 1996 by Levoy and Harahan [LH96]. It derives from the concept of the plenoptic function [AB91] that we need to present first. It is a 7D function $P(x, y, z, \theta, \phi, \lambda, t)$ that models the light radiance that is incident:

- in every point in space $(x, y, z)$
- from every angular direction $(\theta, \phi)$
- with every wavelength $(\lambda)$
- at any time $(t)$

It is a theoretical function that represents every light ray in time and space, and it cannot be fully evaluated. Levoy [LH96] defined in 1991 light-field as a restriction of the plenoptic function at a particular instant of time, for a single wavelength, and assuming no changes along the direction of propagation. It thus reduces to a 4D function, that carries the luminance along a ray ($[W.m^{-2}.sr^{-1}]$) and is usually represented with the two-plane parametrization: $LF(u, v, s, t)$ where $(u, v)$ and $(s, t)$ are coordinates in two different planes, defining the ray direction.

### 3.5.2 Examples of light-field displays

Light-field displays control the radiance of a ray along a particular direction to approximate the light-field emitted by the virtual scene. In a sense, one can consider autotereoscopic displays as a light-field display because they are a coarse approximation of a light-field. Let us consider a schematic two-view horizontal-parallax-only autostereoscopic display, as show in Figure 3.10. We abstract the hardware implementation and consider that a pixel is able to generate different intensity values along two different horizontal directions so that two different viewing zones are created. Then, for each pixel $(u, v)$, the generated light-field is $LF(u, v, s, t) = LF(u, v, s_k(u), t_0)$ with $s_k$ taking discrete values for each $u$ ($k = 1$ or 2 in this case) and $t_0$ being the coordinate of the plane of the drawing, it has no influence since no vertical parallax is visible.

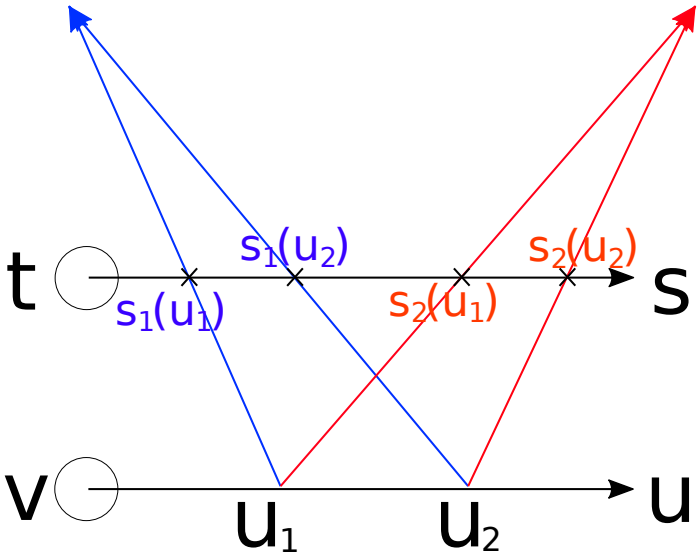



***Figure 3.10:*** *light-field parametrization of a schematic autostereoscopic display (seen from above)*

Even if the term of "light-field display" may be somehow applicable to all other sections of this chapter, we will focus in this section on displays that base both the hardware and image calculation on light-field approximation. The approach of light-field displays is completely different from autostereoscopic displays. Instead of addressing different views toward different locations, the question is reversed: given a bunch of available rays $(u, v, s, t)$, what intensity value should we assign so that they approximate the rays that would have been casted by a physical 3D scene?

A well-known research in this field is the Holovizio$^{TM}$ display [BFB+05] by Holografika. The principle, as depicted in Figure 3.11(a), is to use a 2D array of projectors as light sources and a "holographic" screen. This screen has the ability to scatter light into a small lobe maintaining the incident direction. With proper hardware-related calibrations and a custom OpenGL-based library, the authors are able to generate the correct light-field of a virtual scene. The general principle is the following: instead of putting virtual cameras at viewing zone positions as in autostereoscopic displays, the calculation aims at determining, at each position of the screen, the radiance created by the virtual scene from the direction of each projector.

Other important state-of-the-art light-field displays are layered displays [WLHR11, LA02]. They are also called multiplicative, additive, compressive or attenuation displays. The principle of the hardware is quite simple: a stack of LCD layers separated by a spacer is located in front of a backlight and is used to control ray intensity along controlled directions. A standalone LCD module, labeled $i$, controls the amount of transmitted light $t_i(x, y)$. The resulting intensity of a ray from a backlight position $(x, y)$ along the direction $(u, v)$ is obtained by multiplying transmittances of all layered pixels along the path as shown in Figure 3.11(b). The software aims at computing the transmittance value of each pixel in order to reconstruct the target light-field. The appropriate pixel values are computed using neural networks for the earliest work [LA02] and later as a constrained least-squares problem [WLHR11]. One of the difficulties is that one pixel is used for several ray directions and so the light-field cannot be exactly sampled, only approximated, and artifacts are visible. These displays are also called *compressive* displays for that reason: the hardware does not have enough degrees of freedom to provide an exact solution. Time-multiplexing is introduced as an additional degree of freedom in tensor displays [WLHR12b] to lower the artifacts of the light-field reconstruction. The values of pixels in each LCD layer and in each frame in time are computed using a tensor model and non-negative tensor factorization. Directional backlight, instead of uniform, is also introduced to extend the FOV. However, these approaches have some limitations. First, a strong Moiré pattern is visible due to the overlapping pixel grids, and each layer has a low transmittance itself so the result has low brightness. Finally, the light-field computation is expensive and cannot be done in real-time with a satisfying angular resolution.

In the same way as SMV displays, if the angular resolution of a light-field display

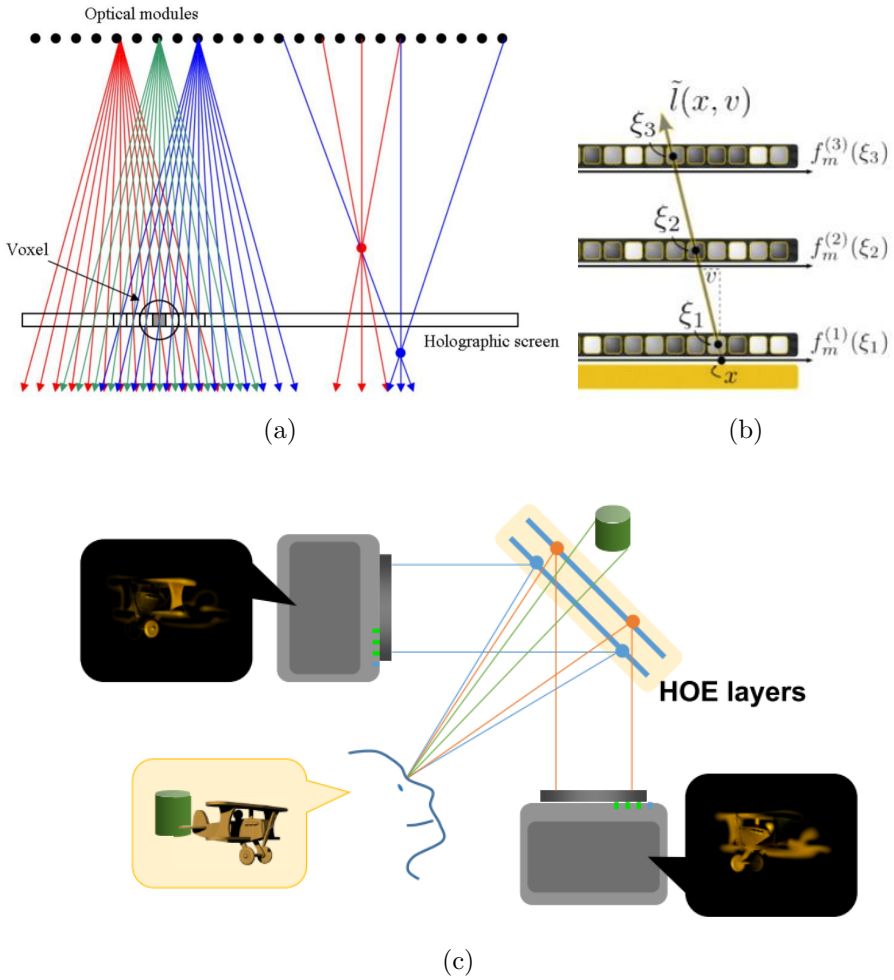



***Figure 3.11:*** *a) Schematic view of the Holografika display(source: [BFB+05]); b) Principle of work of layered light-field displays (source: [WLHR12b]); c) A see-through layered display using holographic optical elements (source: [LJM+16])*

is high enough to have multiple rays from the same point entering the eye, then the eye is forced to accomodate to see it sharp, solving the VAC. This has been done by Maimone et al. [MWL+13] by using a tensor display with a high angular resolution backlight.

Additive light-field displays were also adapted to a see-though display by using holographic optical elements (HOE) and projectors [LJM+16]. Multiple projectors modulate light rays that are incident with different angles on several layered HOEs, as illustrated in Figure 3.11(c). Due to angular selectivity of the HOEs, each layer reflects selectively the light coming from a specific projector toward the eye. This display has several advantages compared to previous work: first, it is transparent and thus can be adapted to AR, it also free from Moiré effects, diffraction and the HOEs have better transmittance than LCD panels. The target light-field can then be calculated like in the other previous methods, so it is also computationally expensive.

Another method is based on a projector associated with a rotating directive horizontal screen to build a 360° light-field display tabletop [XLL+13].

As a conclusion, light-field displays are all about trying to reproduce rays as if they were issued from a real scene. They are theoretically more precise than other approaches because they reproduce better a real life situation. However, hardware limits the quality of light-field reproduction and expensive computations make them impracticable for real-time interaction. We also saw that HOEs are a good candidate to overcome some limitations of conventional optics, and for adapting technologies to AR thanks to their angular selectivity.

## 3.6 Holographic displays

In Section 2.1.5, we explained the optical principle of holography and clearly defined a hologram. As a reminder, a hologram is a recording material that has been exposed to interferences and is able to display a 3D image when lit with particular beams. Similarly to a photograph, the term "hologram" can refer at the same time to the physical object (the recording plate) or to the image it contains.

This describes analogical holography, and in this case the image cannot be updated once it has been recorded. The idea of holographic displays has been around since the beginning of holography, and misinterpretation of this concept have been widely spread. We clarify this in Section 3.6.1. Digital holography is a technique that allows notably the generation of computer-generated holograms and we review its general principle in Section 3.6.2. We review the current state of the art of real holographic displays in Section 3.6.3.

### 3.6.1 "Fake" holographic displays

Science fiction, media or companies often distort the meaning of the word "hologram" and this generates a false idea to the general public of what a hologram is and what a holographic display would be.

The first popular use of this word might come from the Star Wars movie (Figure 3.12) that suggested that a hologram is a volumetric image floating in mid-air and allowing telepresence. The closest existing glasses-free displays are volumetric displays, except that they cannot reproduce occlusion and so could not have a true telepresence application. A 3D floating object or a person can also be created in a AR headset, and companies call them holograms although they are generally stereoscopic 2D images interpreted as 3D. This keeps on nourishing the misconception of a hologram, and it sometimes simply refers to a "floating image".

Most of so-called holographic displays are simply based on the Pepper's ghost effect [Pep90], which is an on-stage trick developed in 1890 for magic shows. It consists basically in reflecting an image on a transparent surface and makes the illusion of presence. This idea was recently applied for on-stage "holographic" presence of singers[1] or politicians[2] by using a semi-transparent mirror reflecting a 2D LCD display (Figure 3.13(a)), and shares nothing with a hologram. Commercial products based on the same effect also exists, such as the DreamHoc$^{TM}$ product by RealFiction [Rea09] that is advertised as being a holographic display although it simply shows floating 2D images. Even some research articles misuse the term of "hologram", such as Holografika [ABF+06] that we reviewed in Section 3.5.2, so one should be careful when reviewing the existing literature.

We do not intend to blame these misuses of the term. On the contrary, we observe that its signification has shifted. In this thesis, to prevent any misinterpretation, we will use the word hologram only in its true optical meaning as introduced in Section 2.1.5: the physical media on which interferences are recorded (or computed) and by extension, the 3D image that is visible when looking at the reconstructed beam.

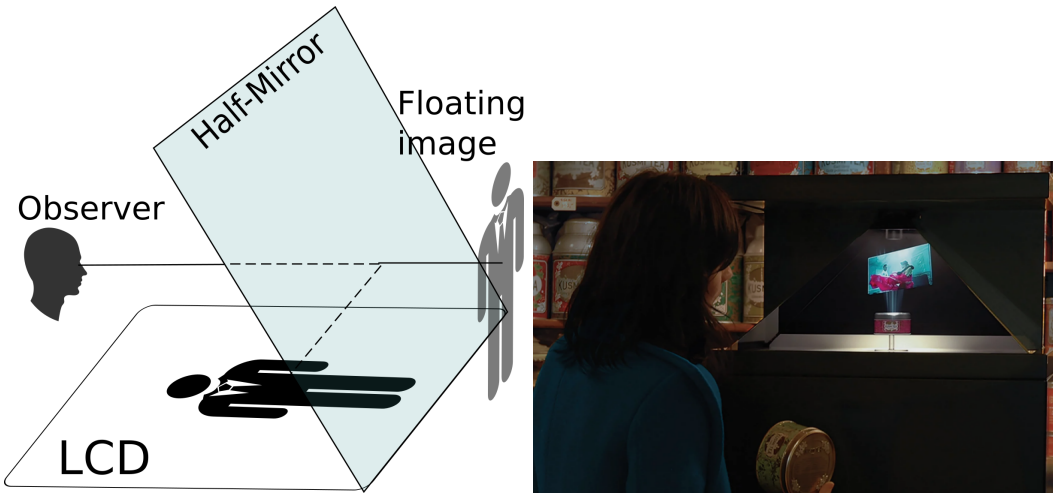
### 3.6.2 Introduction to digital holography

The step from analogical holograms to a holographic display is the same than from traditional photography to television. Before introducing concrete display application, we need to introduce briefly the concept of Digital Holography and Computer-

[1] `www.telegraph.co.uk/culture/music/9207812/Hologram-resurrects-Tupac-Shakur-at-Coachella-2012-music-festival.html` (Accessed on 08/15/19)

[2] `www.telegraph.co.uk/news/2017/02/05/watch-frances-far-left-candidate-appears-via-hologram-paris/` (Accessed on 08/15/19)

***Figure 3.12:*** *A "hologram", as seen in "Star Wars: A New Hope", or the origin of fifty years of misunderstandings (source: imdb.com, ©LucasFilm Ltd)*




(a) (b)

***Figure 3.13:*** *(a) Pepper's ghost effect to reproduce "holographic" presence; (b) "Holographic" pyramid by RealFiction based on the Pepper's ghost effect (`www.realfiction.com/holographic-displays`)*

Generated Holograms (CGH). This section is a quick overview and for more detailed explanations, we refer the reader to [LP12].

The idea of digital holography was proposed in 1967 by Goodman and Lawrence [GWL67] and the theory was developped in 1980 by Yaroslavskii and Merzlyakov [IM80]. Digital Holography can be used either to record or reconstruct wavefronts, and is thus has a very wide range of applications: interference-based imaging systems [SHPW07], 3D microscopy [DSG+08], profilometry [MB97], particle tracking [MY00], adaptative optics [LYK13], and so on. In this section, we will focus on applications of digital holography for CGH and 3D displays.

To create updatable holograms, one needs to be able to modulate the transmittance of a medium in real-time. To that extent, Spatial Light Modulators (SLM) (such as LCD, LCoS or DMD modules) replace the photosensitive material and are used to display a diffraction pattern. Similarly to analogical holograms, the diffraction pattern is constructed by computing the interferences between the virtual object and a virtual reference beam. A real reference beam, similar to the one used for the computation, lights the SLM to reconstruct the object wave.

Generally, the reference beam is quite simple: monochromatic and collimated, so the difficulty of the computation is the object wave propagation from the 3D model to the virtual hologram location. The computation of the complex wavefront on the plane of the hologram can be achieved in several ways.

A 3D model can be decomposed as a sum of punctual light sources emitting spherical waves, as in the Huygens-Fresnel principle. Each point propagates its spherical wave to the hologram plane, and the sum of all point contributions gives the total wavefront. These approaches need to take into account occlusion and shading separately, otherwise all emitting points would be visible [Und97]. Other approaches sum up waves generated by primitives, such as triangles [KHL08], instead of points to accelerate rendering time. The propagation of sources is done through theoretical diffraction models such as:

- Fresnel approximation for diffraction at finite distance. The diffraction can be computed using Fresnel transforms.
- Fraunhofer approximation for diffraction at infinity, or at the focal plane of a positive lens. The diffraction is modeled with Fourier transforms.

Ray-casting techniques are also used to compute CGH [IYS13], inspired from computer graphics techniques. The principle is to cast a bunch of rays from sampled areas of the hologram, called "elemental holograms", and compute the intersection with the closest polygon of the scene, as illustrated on Figure 3.14. The distances of intersection of each ray is stored and used to calculate the resulting diffraction pattern. Occlusion, shading and reflectance models are also adapted from computer graphics to the hologram computations.

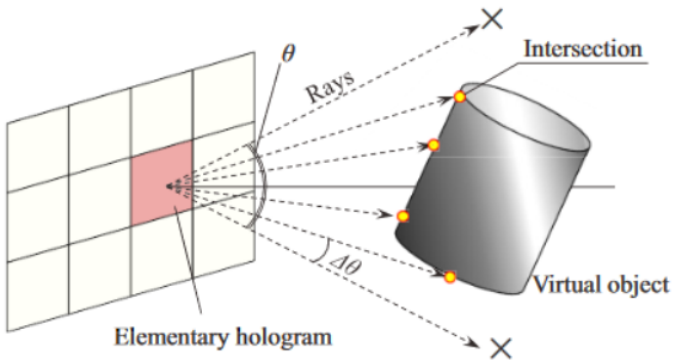



***Figure 3.14:*** *The ray tracing technique for CGH computation. Rays are sent from the hologram and the distances of intersection with the 3D objects are used to calculate the phase difference (source: [IYS13])*

The complexity of hologram computation compared to regular computer graphics rendering is that every point of the scene is not imaged through a lens and then has an influence on the whole sensor. The contribution of all points, or primitives, is computed then summed up, making a complexity that is linear with the number of points/primitives and pixels. Beyond computing parallelization, some acceleration techniques are used. Instead of computing the Fresnel/Fourier transforms of each points, lookup tables can be used [SNMI10, WLF+18] to pre-compute the light wave generated by point light sources on the hologram. The computation can also be done in an intermediate plane close to the 3D model, so that the field extension is smaller and fewer pixels are required. The plane can then be propagated as a whole to the CGH location with wave propagation models [WSO+12].

### 3.6.3 "True" holographic displays

As opposed to the "false" holographic displays reviewed in Section 3.6.1, true holographic displays exist in some extent, but they are not yet mature enough to enter the consumer market. Holographic displays are considered to be the ultimate 3D displays, providing all depth cues. Indeed, the role of holography is to reconstruct a complete wave front as if it was originated from a real object. As we already introduced in Section 3.6.2: CGHs use SLMs to create a diffraction pattern intended to reconstruct a 3D object by diffraction.

There are some issues that prevent CGH to be directly used for a holographic display. First, the computation resources are very high as we already highlighted and this is a brake to real-time holographic displays. Moreover, CGH are known

to have a very low FOV. Indeed, diffraction theory dictates that for an element of size $a$ diffracting a light of wavelength $\lambda$, the angle of diffraction $\theta$ is governed by the relation $sin(\theta) \approx \lambda/a$ (see Section 2.1.4). For current SLM pixel sizes of about $a \approx 10\mu m$ and a visible wavelength of $\lambda = 0.6\mu m$, the total FOV is limited to about 7°. Even if sufficiently small pixels would be available to provide a satisfying FOV, then a very high number of them would be required to have a satisfying image size, resulting in higher computation and storage needs. According to St Hilaire et al. in 1990 [SHBL+90], a hologram with a size of 100mm by 100mm, with a 30° FOV and a framerate of 60Hz would require a data-rate of 12 Terabits per second.

### Holographic projectors

As explained in Section 2.1.5, a hologram is able to reconstruct the object either as a virtual image or a real image. In case of a real image, meaning visible on a screen, the display is a holographic projector [Buc08]. Unlike suggested in Figure 3.12, holographic projectors usually do not provide 3D images. Instead, they only provide projection of a 2D image on a screen, because light generally cannot be scattered in mid-air. The main advantage of holographic projectors over conventional projectors is efficiency and contrast. In a conventional projector, such as a LCD or DLP, light is modulated by attenuating the backlight as we saw in Section 2.2.3. A black pixel is created by attenuating as much as possible the energy coming from the backlight, so a huge loss of energy is inevitable. In a holographic projector, even though some energy is lost due to the SLM, the principle is to redirect light by diffraction toward desired locations. Black areas, in theory, receive no energy so the contrast ratio is far better than in conventional projection devices, as for the LBS projectors.

### Holographic displays

Holographic displays that aim at creating virtual images are able to generate 3D content visible when looking through the display. We will focus on these displays in the remainder of this section.

One of the earliest functional holographic displays that overcame known limitations was designed by St-Hilaire et al. [SHBL+90, SHBL+93]. They use acousto-optic modulators SLM to modulate light phase and a supercomputer to calculate the CGH. To overcome the CGH limitations, they discarded vertical parallax, chose a small image size and achieved a 15° FOV using a horizontal scanning mirror. Other work by the same authors [SHBL+93] improved this display with additional scanning mirrors, extending the horizontal FOV to 38°, thus improving resolution and image size.

More recent work on holographic displays show exciting perspectives. The SeeReal company has been able to achieve high resolution holographic display, that we

can almost call holographic television [SLH07, SLH+08]. The key of their technology is to compute *sub-holograms* instead of the whole hologram, requiring a lot less computational resources. Eye tracking the user allows to address two sub-holograms at each eye position: unseen information is not generated. The user perceives binocular parallax with the two perspective corrected holograms, and all other depth cues, including convergence.

A super-wide FOV CGH was achieved using a standard CGH imaging setup coupled with a convex parabolic mirror [SSK+18]. In this case, the hologram calculation must take into account the reflection by the mirror. However, it results in a very small image size because of the curved mirror.

Holographic near-eye displays prototype for augmented reality applications are also introduced. For near-eye displays, the limiting diffraction angle is not as critical than for distant displays because the eye is at a relatively fixed position with regard to the display. Notably, Microsoft Research designed a prototype using an off-axis holographic projector and a holographic optical element [MGK17] and proposed an optimized algorithm that takes into account optical aberrations correction. VividQ [Viv17] is a company that also developed a true holographic HMD. Their optical system is quite standard with a SLM imaged on a half mirror, and their technology relies mostly on proprietary hologram computation and storage algorithms. Their software also corrects optical aberrations.

## 3.7 Conclusion

In this chapter, we introduced the basics of 3D vision, and particularly the depth cues that will lead our research on 3D displays. The most important one to reproduce are the binocular disparity and convergence, but motion parallax is also a strong depth cue that is often neglected. Binocular cues are reproduced by glasses-based 3D displays, and motion parallax can be introduced by software through the wear of headsets. However, it is sometimes undesirable that the user has to wear a device to experience 3D digital images, as it may cause discomfort and somehow breaks the natural viewing conditions. Many methods to design glasses-free 3D displays exist and we reviewed the main ones.

Autostereoscopic displays base their technology on generating multiple viewpoints and assigning them to different viewing zones. They generally are adaptations of existing displays such as LCD displays or projectors, combined with an optical element that performs the view separation. Computing a signal for such a display consists of rendering multiple views of the same scene through conventional rendering pipelines. In general, they have the vergence-accomodation conflict, but it can be solved with the SMV techniques that consist in generating multiple views per eye.

Volumetric displays generate voxels directly in a volume, generally by sweeping a 2D image. They display voxels directly at the correct spatial location so no conflict occurs, but they generally cannot reproduce occlusion and are thus limited to volumetric data. Moreover, the depth of the 3D scene is limited to the volume of the display.

Light-field displays generate directional rays with controlled intensities so that they reproduce light rays coming from a virtual scene as closely as possible. For now, the hardware is not able to offer as much control as required so only approximations are provided, resulting in visual aberrations. In addition, the computational complexity of such approaches makes them impracticable for real-time rendering.

Holographic displays provide a way to directly reconstruct the wavefront of a virtual scene, but they still suffer from hardware and software issues that prevent the perfect updatable holographic display to exist.

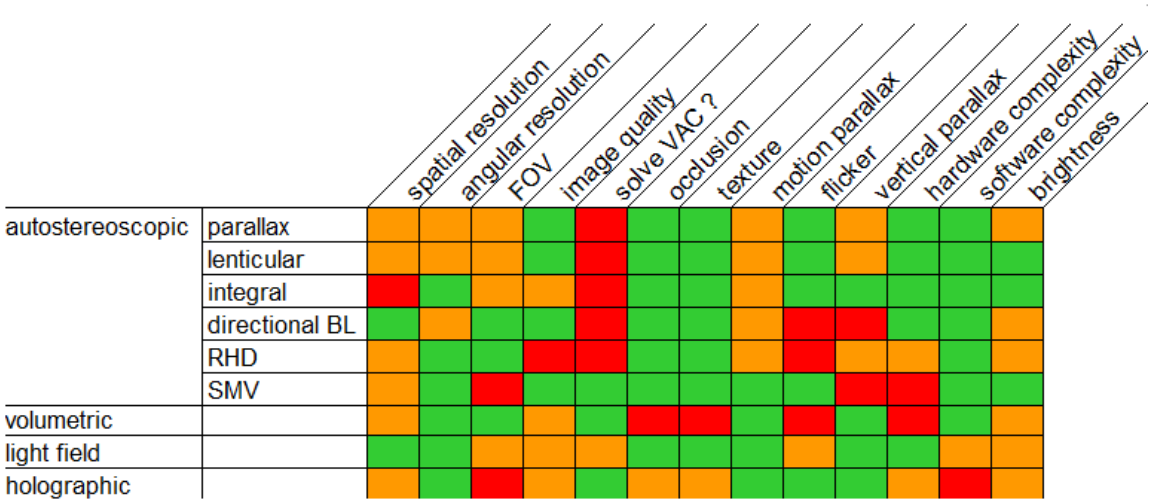



***Figure 3.15:*** *Summary of advantages and drawbacks of the different 3D display technologies.*

# Chapter 4

# A low-cost multitouch spherical display

Despite constantly improving image quality, conventional flat 2D displays have inherent limitations. As we reviewed in the previous chapter, 3D displays extend the possibilities of 2D displays, by providing additional depth information through the depth cues.

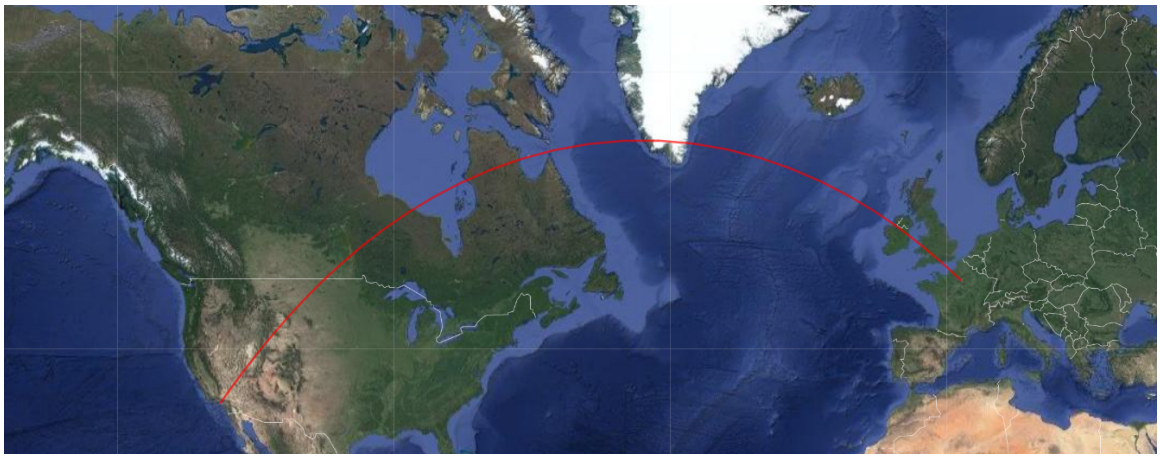
However, the lack of depth cues in conventional displays is not the only limitation. In addition, their shapes and properties are often the same, and a virtual scene is often perceived through a rectangular flat screen. Most of the time this is sufficient but they are not adapted to every kind of data. Even though curved displays emerge thanks to notably the OLED technology (see Section 2.2.2), they cannot reach non developable shapes such as spheres.

In this chapter, we propose an innovative system that provides a way to display data on a spherical surface and interact with it directly by touch. We follow a low-cost and "Do It Yourself" approach so that it is easily reproducible. This work has been validated by a conference publication [CRG17] at SID Display Week 2017.

We will also discuss optical and software tools that we use later in the development of a 3D display. The main similarities of both projects are LBS projectors' properties, software calibration, computer vision and fast prototyping using a 3D printer.

## 4.1 Introduction

A spherical multitouch display is an opaque globe on which information is displayed and can be interacted with directly by touch. It offers many interesting potentials that are not possible with regular 2D displays. First, it has a 360° viewing angle, allowing users to be located all around the display. Secondly, it is more suitable



(a)

***Figure 4.1:*** *A plane trajectory between Paris and Los Angeles on a Mercator projection looks surprisingly curved, but this is the shortest path. (source: scribblemaps.com)*

to display spherical data, in particular geographical data: who has not been fooled once by a plane trip representation on a planisphere that looks exaggeratedly curved, as shown in Figure 4.1? Last but not least, such a display arises curiosity and interest. It has numerous applications in education, communication, collaborative work or data visualization.

Commercial spherical displays already exist, and we review them in Section 4.2, but they have not yet entered every classroom, most probably due to the high cost of existing products. They thus remain non-conventional displays: users and developers are generally not familiar with.

In this chapter, we propose a new design of a low-cost, yet effective, spherical multitouch display. The principal components are a laser beam steering projector (see Section 2.2.3), a fisheye lens, a webcam for optical tracking, off-the-shelf optical components, and custom 3D printed parts, as well as software for tracking, calibration, and color correction. Our work makes multitouch spherical displays accessible to a larger number of people, so that any researcher or enthusiast interested in this field can contribute with hardware or software improvements, as well as user interaction studies.

## 4.2 Related work

Global Imagination was the first company to patent and commercialize spherical displays that are not touchable [URF08]. Their product, the Magic Planet$^{TM}$, is

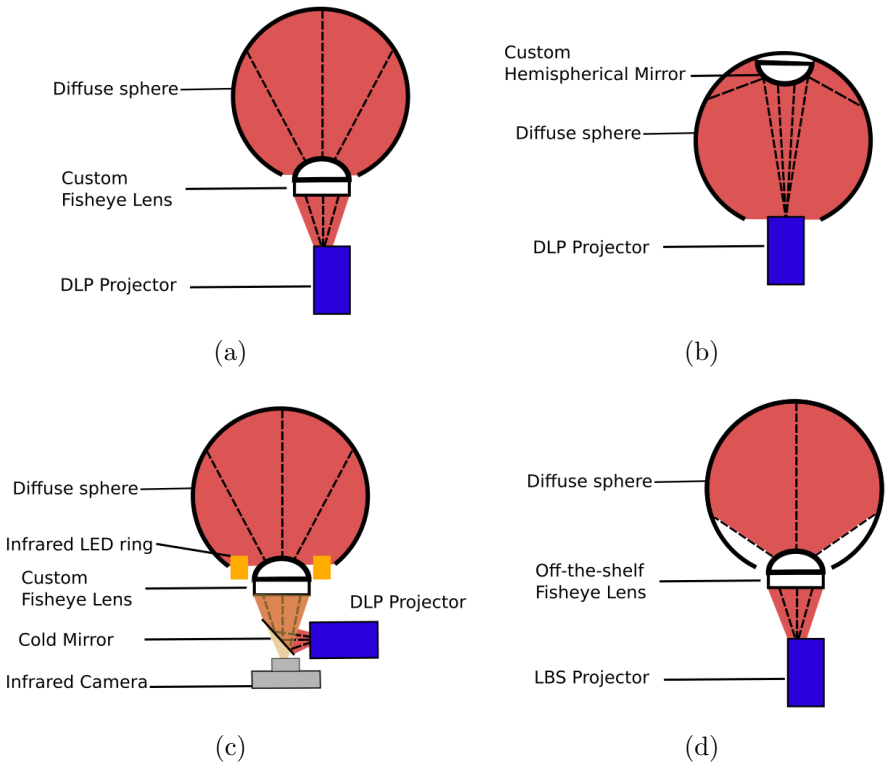



***Figure 4.2:*** *Different rear projection techniques for spherical displays: (a) Custom fisheye lens [URF08]; (b) Custom hemispherical mirror [LFL03]; (c) Custom fisheye lens and infrared camera for finger tracking [BWB08]; (d) LBS projector and off-the-shelf fisheye lens [Nir11]*

based on an optical coupling between a DLP projector and a tailored fisheye lens. The fisheye lens is specifically designed to conjugate the image plane of the projector onto the surface of the sphere.

ArcScience also sells non-touchable spherical displays following a different approach: they use a patented hemispherical convex mirror [LFL03] to project the beam of the projector onto the inside of the sphere. The projection is done from the south pole of the sphere, and the hemispherical mirror is located at the north pole as illustrated in Figure 4.2(b). The sphere is not fully covered because of casted shadows at the north pole.

To our knowledge, the first work on developing touchable spherical displays was presented by Microsoft Research in 2008 [BWB08]. In particular, they added multi-touch capabilities to the spherical display from Global Imagination. For that, they placed an infrared LED ring at the bottom of the sphere that lights the sphere from

the inside, and an infrared camera is added to image the inside surface of the sphere and to make it touchable with finger tracking. The projector and camera light paths are separated thanks to a cold mirror that reflects visible light and that is transparent to infrared. This research prototype was used to perform a series of user experiments on interaction with spherical displays. However, not many details are given regarding the hardware implementation, making it hardly reproducible.

A commercially available multitouch spherical display is the PufferSphere$^{TM}$, sold by Pufferfish [Ltd02]. The projection part relies, like previous ones, on the association of a DLP projector and a specially designed fisheye lens, called the Super Umami Lens. The method to make it touchable is not detailed by the company, but is also based on finger tracking.

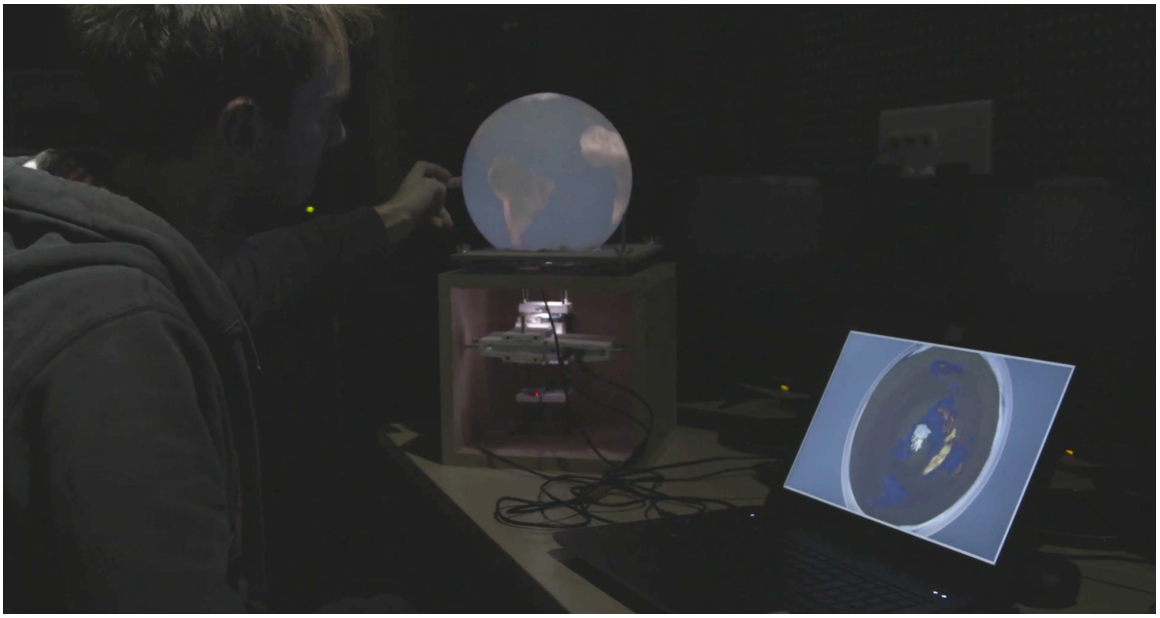
All of the examples above rely on tailored optics that are difficult to design and expensive to produce.

Another approach is to use several projectors and no optics [ZWM+17] in order to cover the whole spherical surface. However, this solution is not cost-friendly neither due to the involved multiple projectors.

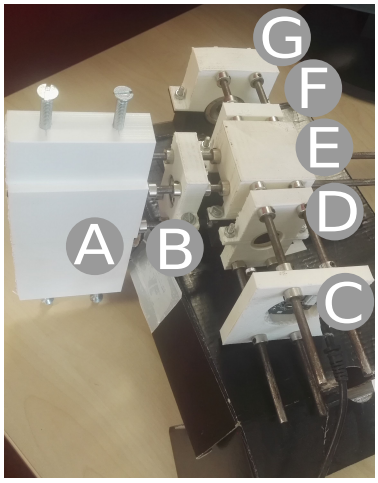
Alternatively, Patel [Nir11] proposed a low-cost and "Do it yourself" spherical display, however, it is not touchable. It exploits the technology of a *focus-free* laser beam steering projector, which allows the projection of a sharp image on curved surfaces. The projector is associated with a commercially available fisheye lens allowing to cover almost the entire surface of the sphere. Contrary to other solutions, there is no need to specially design the fisheye lens because of the focus-free property of laser projectors.

## 4.3 Overview

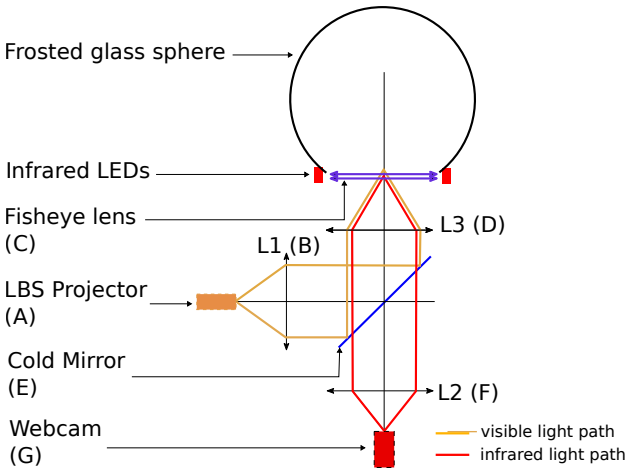
The main contribution of our work is the design of a spherical display with multitouch capabilities that is low-cost and reproducible by using only off-the shelf optical components and a single projector. Our hardware design is illustrated in Figure 4.3 and detailed in Section 4.4. For the display, we use a LBS projector combined with a standard fisheye lens to project the image on the interior surface of a frosted-glass sphere with an opening neck, as inspired by Patel [Nir11]. For multitouch capabilities, we use optical tracking on a second light path with an infrared camera that senses the reflection of rear-projected infrared LED light, inspired by [BWB08], but without requiring a custom shaped lens. We combine both light paths with a cold mirror, and we change the FOV of the projector and the camera with a lens system in order to reach almost 180° at the output of the fisheye lens. Our lens system is designed to maintain the "focus-free" property of the projector through the optical system. The finger tracking (Section 4.5.3) is done by analyzing the image of the infrared camera.



(a)




(b)




(c)

***Figure 4.3:*** *(a) Picture of the prototype in use, showing the projected image on both the laptop and the sphere; (b) Picture of the optical system, assembled with 3D printed parts; (c) Annotated scheme of the whole system and correspondences with the picture.*

Since we use standard optical components, we rely on calibrations for the projection (Section 4.5.2), the tracking (Section 4.5.4) and the correction of chromatic aberrations (Section 4.5.5). Other considerations on the prototype, together with its cost, are detailed in Section 4.6.

## 4.4 Hardware

### 4.4.1 Sphere

The sphere that we use is actually a 3/4 portion of a 25cm diameter sphere, with an opening neck of 12cm diameter where the light comes from. It is a part of a globe lamp used for interior lightning, bought from the local hardware store. It is made of frosted glass, which is a translucent material adapted for the use as a rear projection screen.

### 4.4.2 Projector

For a spherical display, we need to have a projection system that is in focus everywhere on a spherical surface. Classical projectors do not allow such a property due to their limited depth of field, unless custom optics are used as in [BWB08, Ltd02, URF08, LFL03]. Consequently, we use a LBS projector. As explained in Section 2.2.3, they are based on a combination of three laser diodes and a fast-scanning mirror. The linear evolution of pixel size along the projection distance establishes a very interesting focus-free property. Consequently, projections on non-planar surfaces are possible without custom optics.

We have chosen the Celluon PicoPro projector. It has an equivalent resolution of 1920x720, and its FOV is 43° horizontally and 22° vertically. At the time of the project, it was the brightest LBS projectors available with 30 lumens. Nowadays, it is not sold anymore but it can be replaced with a Sony MPCL1A, with a 32 lumens brightness. We expect that the brightness of LBS projectors continues to grow in the future.

### 4.4.3 Fisheye lens

The field of view of the projector is not wide enough to cover the entire inside surface of a sphere, so we use a fisheye lens to increase the projection angle up to nearly ±90°. We use an afocal fisheye lens, which means that in contrast to fisheye objectives, it does not conjugate the object on an image plane but multiplies the angles by a certain magnification. We decided to use a ×0.2 magnification Opteka

Fisheye lens, similar to Patel [Nir11]. It is a low-cost fisheye lens whose properties are suitable for our prototype. The intended use of the lens is to be mounted on a camera so that the light coming from wide angles narrows towards the camera's lens. In our work, we use the lens not only for this purpose, with the infrared camera imaging the whole surface of the sphere, but also for the exact opposite: we want to expand the beam of the projector from small angles to reach nearly $\pm 90°$ in the output. For the projection, the relevant magnification of this lens is $1/0.2 = \times 5$, and so it requires an input angle of $\pm 18°$ to achieve the $\pm 90°$ output.

### 4.4.4 Infrared camera and illumination

In order to make the spherical display touchable, we use optical tracking of the user's fingers illuminated in the infrared, similar to Benko et al. [BWB08]. Note that the camera for the tracking must work on the same optical axis as the projector to be able to use the fisheye lens, but in a different spectral range so that the displayed image does not interfere with the tracking. For the illumination, we have built a circular infrared LED ring operating at 880nm, which is a suitable wavelength since it is far enough from the maximum wavelength of the projector, but close enough to the visible spectrum in order to still be sensed by affordable cameras. Technically, we use 16 OSRAM SFH485 LEDs distributed circularly around the opening neck of the sphere, in order to achieve an as-uniform-as-possible illumination (see Figure 4.9). For convenience and mobility, the electric circuit is USB powered.

We use a PlayStation Eye camera, since it is a cheap and effective webcam, and its sensor can work in the near infrared range. This camera is commonly used for infrared tracking and can be modified with existing adapted material (e.g. PeauProduction[1]).

Originally, the webcam works only in the visible spectrum because of an IR-cut filter located between the lens and the sensor, although the camera's sensor has a wider spectral range. We have replaced this IR-cut filter by a bandpass filter centered on 880nm.

We also modified the FOV by changing the main lens of the camera. Originally, the main lens of the camera can be switched between two focal lengths, resulting in either a 56° or 75° FOV. This is too wide to work in our optical system as too much of the resolution would be lost. The new lens has a focal length of 6mm to obtain a field of view of 30° horizontally and 23° vertically. This makes the camera's FOV closer to the projector's FOV, and so they can share common lenses. Finally, we removed the case of the webcam to reduce useless volumes.

[1] `www.peauproductions.com` (Accessed on 09/02/2019)

### 4.4.5 Lenses

In our optical system illustrated in Figure 4.3(c), the micro-mirror of the projector is located at the focal point of the lens L1 in order to collimate the beam of the projector. The beam is then reflected on the cold mirror towards the lens L3 that makes it converge with the desired input angle $\alpha$ on the fisheye lens. The angle $\alpha$ depends on the height $H_3$ of the beam on the lens L3 and its focal length $f'_3$, such that $\tan(\alpha) = H_3/f'_3$. Assuming that the diameter of the lens is not limiting, $H_3$ is defined on the first lens as $H_3 = H_1 = f'_1 \tan(FOV/2)$. According to the previous equations, the following setup results in the full cover of the sphere for projection and imaging, where $D_i$ is the diameter of lens $i$:

- L1: $f'_1 = 43mm$, $D_1 = 23.5mm$
- L2: $f'_2 = 43mm$, $D_2 = 23.5mm$
- L3: $f'_3 = 22mm$, $D_3 = 18mm$

Theoretically, the focal length of L2 could be $f'_2 = 27.3mm$, so that the input angle at the fisheye would be $\pm 18°$ resulting in $\pm 90°$ at the output. However, as a single degree of error in the prototype results in a 5 degree loss on the sphere, we have chosen the effective L2 lens to have a focal length of 22mm, so that up to $\pm 3°$ of error are acceptable while keeping the full coverage.

Another important issue to take into consideration when doing optics with a LBS projector is the evolution of the scanning laser beam diameter. As we explained in Section 2.2.3, the image is sharp because the beam has low divergence, and adding a lens makes the beam more diverging, resulting in a blurry image. The second lens must put the waist near its focal point in order to restore back the low-divergence of the beam (see Figure 4.4) and to make the image sharp again.

### 4.4.6 Cold mirror

We use a cold mirror to separate the infrared illumination used for tracking from the visible light of the projector. More precisely, we use a 45° cold mirror from KnightOptical[2]. A cold mirror is reflective in the visible spectrum ($400 - 700nm$) and transparent in the infrared spectrum ($> 700nm$). If L1 is completely lit, the size of the collimated beam is approximately 17.5mm. The mirror we have chosen has a size of 40mm x 25mm corresponding to an apparent surface at 45° of 28.3mm x 25mm, so that the mirror is not a limiting optical element in the system.

[2] www.knightoptical.com

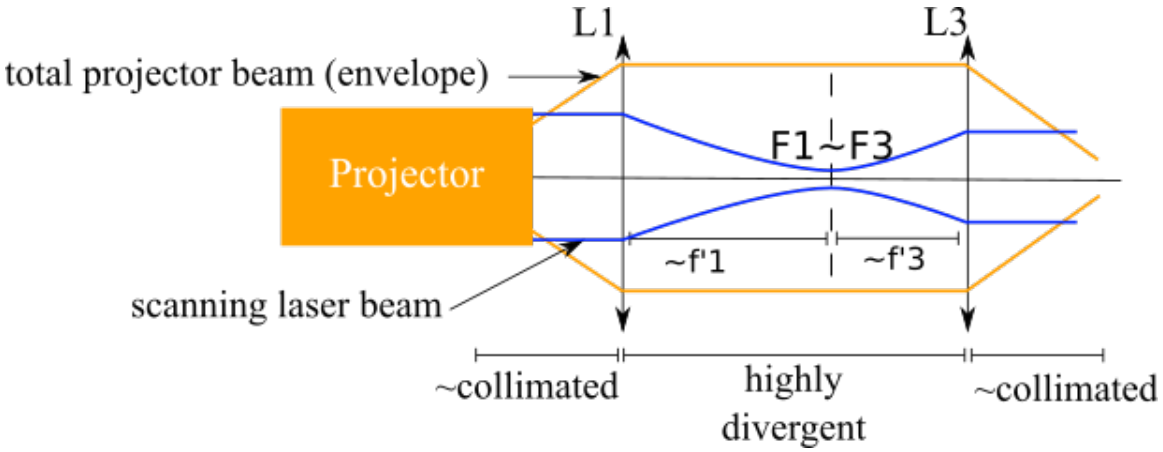



***Figure 4.4:*** *Variation of the scanning laser beam diameter (not in scale) with the distance. The quasi-collimated laser of the projector is focused at the focal point F1 of the lens L1 before diverging a lot. L3 makes the beam collimated again if the waist is also on its focal point.*

## 4.5 Software

In this section, we detail the software contributions. First, the input image needs to undergo a transformation to be suitable for the fisheye projection on a sphere that is described in Section 4.5.1. Multitouch interaction is done with a third party software communicating with ours on a network protocol. Calibrations are required to balance the hardware imperfections due to our low-cost approach, and we calibrate our display for projection (Section 4.5.2), tracking (Section 4.5.4) and the correction of chromatic aberrations (Section 4.5.5)

The software is implemented in C++ and OpenGL, using the TUIO library [KBB+05] and GLSL shaders [RLKG+09]. An overview of the software pipeline can be seen in Figure 4.5.

### 4.5.1 Rendering

A regular 2D window adapted for a regular flat display cannot be directly projected as an image source in the spherical display. It has to undergo an equidistant azimuthal projection: let us consider an Earth map in cylindrical projection that corresponds to a 2D application development environment. As shown in Figure 4.6, the North Pole point is elongated along the top pixels line and the South Pole in the most bottom pixel lines. In our system, the North Pole needs to be the center of projection and the South Pole elongated in the outer circle. The required image transformation is an

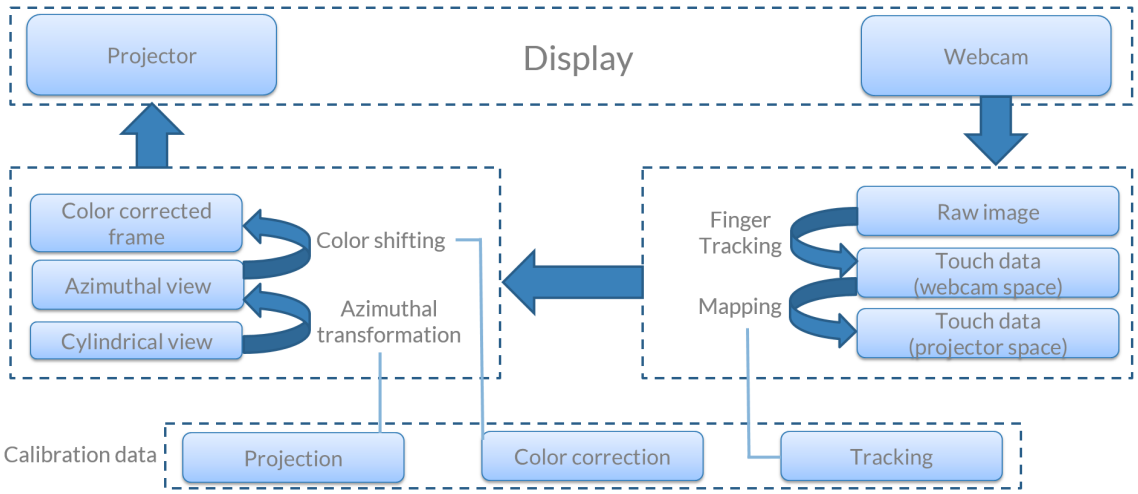



***Figure 4.5:*** *Overview of the software pipeline*

equidistant azimuthal projection, used for centuries for map representation [Sny87]. It is a projection of a globe on a tangent plane.

The projection is done directly on the GPU by using vertex and fragment shaders, and so we need to use the inverse projection formula: given a pixel in the final image, whose OpenGL coordinates are $(x, y) \in [-1, 1]$, we look for the corresponding point on the globe parametrized by its longitude $\lambda \in [-\pi, \pi]$ and latitude $\varphi \in [-\pi/2, \pi/2]$. The globe coordinates are then transformed to texture coordinates $[u, v] \in [0, 1]$ to indicate which pixel has to be read. This is convenient since in a cylindrical view of the globe, $\lambda$ and $\varphi$ are aligned on the pixel grid. Figure 4.6 illustrates the different used frames.

The inverse equidistant azimuthal projection formula [Sny87] gives the longitude $\lambda$ and latitude $\phi$ according to the coordinate of the plane of projection $(x_{plane}, y_{plane}) \in [-\pi, \pi]$ , which are simply obtained by multiplying OpenGL coordinates by $\pi$. Let $\varphi_1$ and $\lambda_0$ be the center of projection (common point on the tangent plane) and $c = \sqrt{x_{plane}^2 + y_{plane}^2}$ the distance of the current point from the center, we have:

$$\begin{aligned} \varphi &= \sin^{-1}\left(\cos(c)\sin(\varphi_1) + \frac{y\sin(c)\cos(\varphi_1)}{c}\right) \\ \lambda &= \lambda_0 + \tan^{-1}\left(\frac{x\sin(c)}{c\cos(\varphi_1)\cos(c) - y\sin(\varphi_1)\sin(c)}\right) \end{aligned} \tag{4.1}$$

And the remapping of globe coordinates to texture coordinates is straightforward:

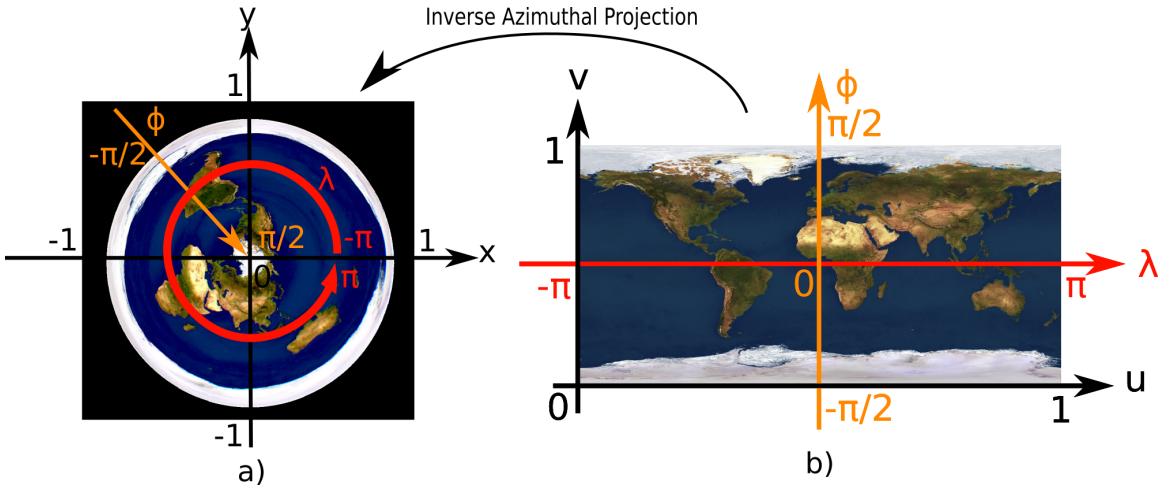



***Figure 4.6:*** *a) Spherical coordinates and OpenGL coordinates in the transformed image; b) Spherical coordinates and texture coordinates in the cylindrical projection (input texture)*

$$\begin{aligned} u &= \frac{\lambda}{2\pi} + 0.5 \\ v &= \frac{\varphi}{\pi} + 0.5 \end{aligned} \tag{4.2}$$

### 4.5.2 Projection calibration

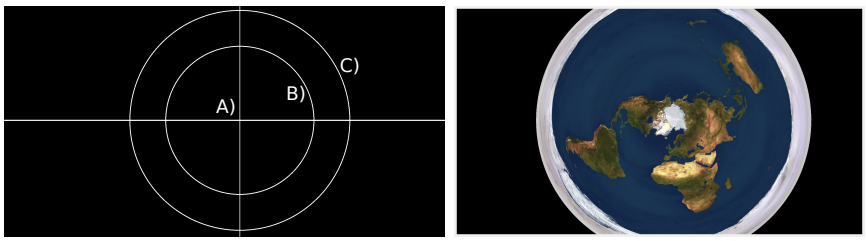
The projection needs to take into account hardware considerations. Due to slight misalignments of the projector with the optical system, the center of the projected window does not coincide necessarily with the north pole of the globe, neither does the equator. To prepare the image to be rendered on our hardware, we need to go through a calibration step. This calibration has to be done once as long as hardware elements do not move relatively to each other.

We use an interactive calibration process illustrated in Figure 4.7(b) that requires the user to:

A) set the center of the cross $(x_0, y_0)$ to the north pole of the sphere, so that the OpenGL position of the center of the images to display is known.

B) set the middle circle on the equator of the sphere $r_{eq}$, so that a radial scaling can be applied to the image.

C) set the outside circle as the maximum visible circle on the sphere $r_{lim}$, so that a working area is defined where the tracking needs to be calibrated.




(a) (b)

***Figure 4.7:*** *a) Calibration steps: setting up the pole, radius of the equator and maximum visible radius; b) Result of the azimuthal projection shader.*

With OpenGL coordinates and the calibrated parameters $x_0, y_0, r_{eq}$, we use the same projection formula defined in Equation 4.1 in a viewport centered on $(x_0, y_0)$ and with the radial scaling applied on coordinates $(x_{plane}, y_{plane}) = \frac{-\pi \times (x,y)}{2r_{eq}}$. The minus sign comes from the presence of the mirror in the projection light path.

The circle defined in C) contains 55% of the total projected image, which corresponds approximately to 760k pixels over the 1.4M pixels of the projector. In Section 4.7, we propose a method to improve the final resolution.

Note that the distribution of pixels over the sphere is not regular, as pixels close to the north pole are bigger than pixels of low latitudes because of the high distortion of the fisheye lens (see Figure 4.8). Contrary to DLP projectors, there is no visible pixel grid and the image is continuous.

### 4.5.3 Finger tracking

We use optical finger tracking for the multitouch capability: the entire inside surface of the sphere is imaged onto the sensor of the infrared webcam thanks to the lens system, and then the image is processed in order to detect the position of the fingers. We use *rear diffused illumination* [SBD+08]: the infrared light coming from the LED ring is partly reflected by the frosted glass sphere, and partly transmitted. When the user's finger points on the sphere, the transmitted infrared light is reflected by the fingers towards the webcam as shown in Figure 4.9.

To perform the finger tracking, we use the third-party open source software Community Core Vision (CCV), developed by the NuiGroup community [Moo06]. It is an image processing software dedicated to optical finger tracking, with a graphical interface enabling to adjust different parameters in order to increase the robustness of the tracking.

The computer vision algorithm requires the user to first set up a background image where no fingers are visible (Figure 4.9(a)). Then, in each frame, the cur-

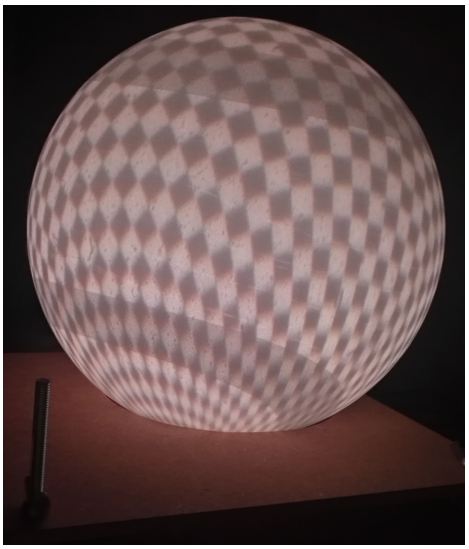


***Figure 4.8:*** *A regular checkerboard without deformations is projected, and the non-uniformity of the output resolution can be observed.*

rent image (Figure 4.9(b)) is subtracted from the background before being binarized (Figure 4.9(c)) to retrieve the position of the fingers in so-called "blobs". A series of adjustable filters offered by CCV are used to improve the tracking robustness. The blobs are then sent over a specific network protocol called TUIO [KBB+05] and retrieved in our custom software as a C++ object, whose attributes are notably position, size, speed, a tracking ID, and so on. As shown in Figure 4.9, despite the low difference of intensity with the quite intense background image, the fingers are correctly detected. In Section 4.7, we discuss future improvements, notably to reduce the intensity of the background image.

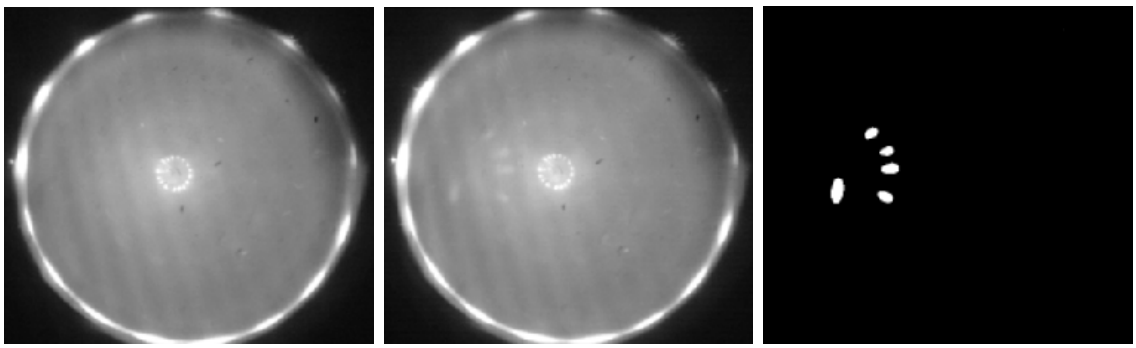
This information can then be used to identify touch gestures, such as tapping, double tapping, dragging, flicking, pinching and others [VWW10].

### 4.5.4 Finger tracking calibration

A tracking calibration is necessary in order to make a one-to-one correspondence between the projector's and camera's pixels.

To do so, our calibration preprocess displays a series of points $(x_{proj,i}, y_{proj,i})$ lying inside the maximum visible circle defined in the first calibration step. Then, the user touches each point, and the touch position is captured by the camera to provide the corresponding coordinates $(x_{cam,i}, y_{cam,i})$ in the camera frame. The difference between the projected and detected points in their relative frame can be seen in Figure 4.10.

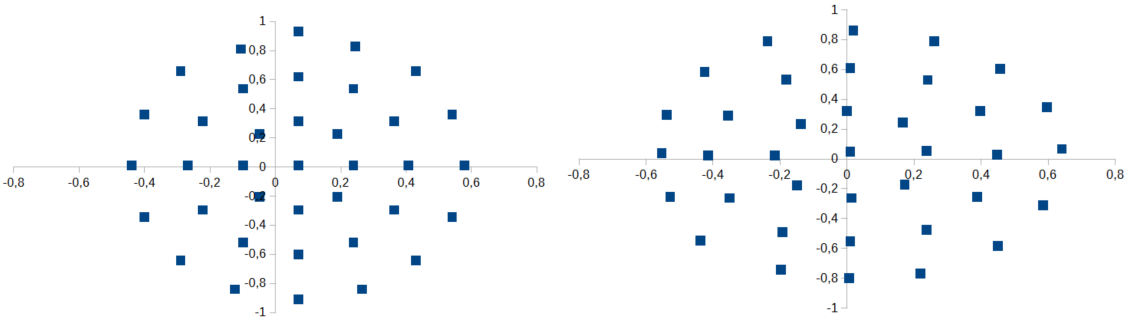
In order to extrapolate the mapping function to any pixel, a bilaplacian recon-



(a) (b) (c)

***Figure 4.9:*** *Finger tracking: (a) Background image with no fingers. The illumination is almost uniform on the sphere except close to the opening where we do not expect any tracking and at the pole where the LEDs are specularly reflected. (b) Sensed image with touching fingers. (c) Retrieved blobs, containing touch related information*

struction [ALP14] is performed, with the sampled points as constraints and the other pixels as unknowns. This method evaluates a 2D function that passes through the calibrated points and that has a zero second derivative everywhere else. Typically, 30 to 50 points are enough to have good tracking performance on the whole sphere.



(a) (b)

***Figure 4.10:*** *Representation of the sampled points. a) Projected points* $(x_{proj}, y_{proj})$*; b) Detected points* $(x_{cam}, y_{cam})$

### 4.5.5 Correction of chromatic aberrations

The optical system has some chromatic aberrations that degrade the image quality. Several factors explain their presence: firstly, we use only low-cost lenses so

aberrations are inherent to them, secondly we use 3D printed mounts so some misalignments are inevitable and lastly, we use the whole aperture area of the lenses, far from the Gauss condition, making aberrations worst.

Fortunately, the projector is composed of three monochromatic lasers whose spectra do not overlap to each other. This allows a color correction by channel that would not have been possible for light sources having a broad spectrum. As shown in Figure 4.11, we correct the chromatic aberrations by applying the inverse color shifting to the input image. To do so, a user samples the color correction over the sphere by translating red and blue channels onto the green one (chosen as reference) at several points. As the aberrations are continuous, we perform in a preprocess, like the tracking calibration, a bilaplacian reconstruction in order to generate a color correction map.

Each image displayed on the sphere is first rendered into a frame buffer object, and then a color correction shader is used on the GPU to correct the position of the red and blue channels, thus minimizing color aberration effects.

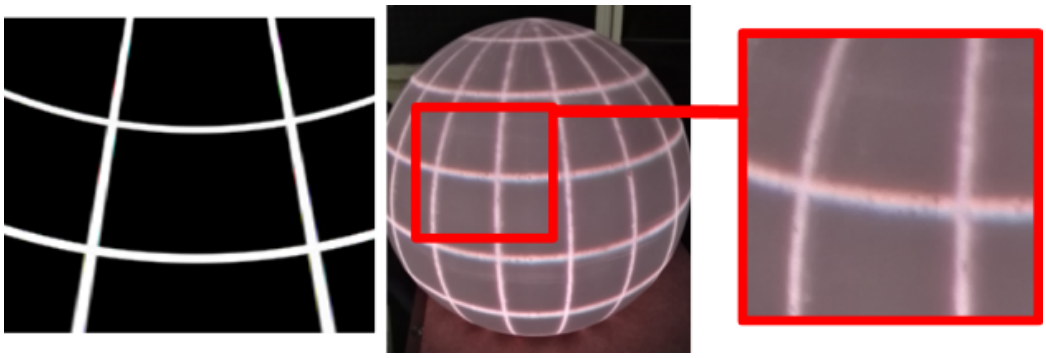
A correction of chromatic aberrations using theoretical models has failed because the projector's relative position with the lenses is a critical factor that cannot be controlled accurately enough, resulting in non-symmetric aberrations that we sample using our technique.

### 4.5.6 Application development for spherical displays

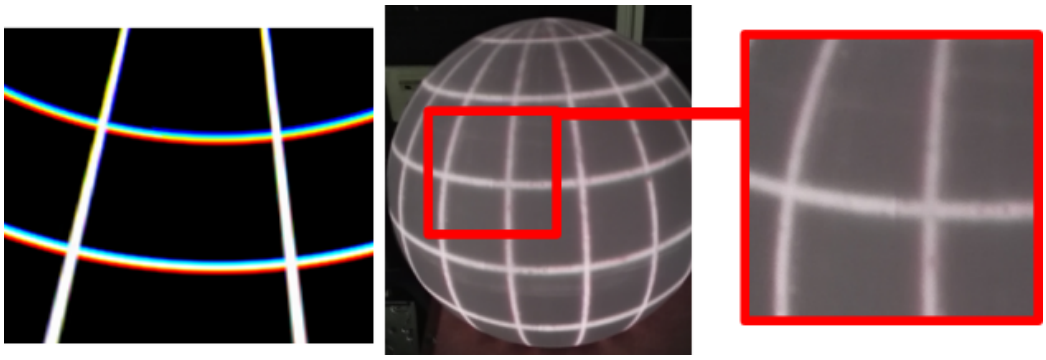
In addition to the basic environment ensuring a correct functioning of the display, we have developed a tool to ease the development of applications. Indeed, developers are more used to code applications for regular displays, and the requirements for spherical displays might not be evident. Our solution allows to completely disconnect the creation of the content and the display itself. For example, an application for the display can be developed and tested on a classical development environment without having a spherical display at hand. This work has been presented at the 2017 ACM International Symposium on Pervasive Displays [CRR+17] in Lugano.

As represented in Figure 4.12, the tool relies on three layers:

- The application layer, which is a regular desktop application such as a web page designed for a regular display. The developer can map specific multitouch events to keyboard events.
- The output processing layer applies the azimuthal transformation and color correction on the application layer. This creates directly the projected image.
- The input processing layer receives the user inputs, i.e. the coordinates of the blobs, and turns them into a regular input for the desktop application. For



(a)



(b)

***Figure 4.11:*** *(a) The uncorrected image results in visible chromatic aberrations b) By applying the inverse color shifting to the input texture, colors overlap correctly in the final image.*

example a single touch triggers a mouse click, and specific multitouch gestures trigger the associated keyboard events.

We believe that this work will contribute to spread the use of multitouch spherical displays by enabling graphic and web designers to easily develop compatible applications.

## 4.6 Prototype

In order to keep the low-cost aspect of this work, all the lenses have been ordered from SurplusShed[3], a low-cost optics provider. Moreover, the mounts for the

[3] `www.surplusshed.com`

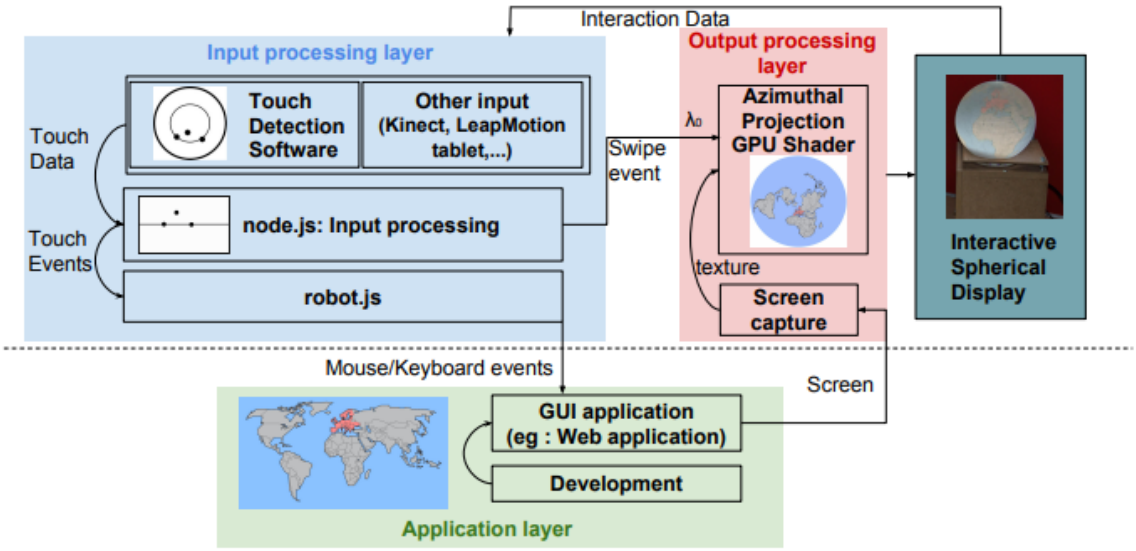



***Figure 4.12:*** *Illustration of our framework with the three different layers and their mutual interactions [CRR+17].*

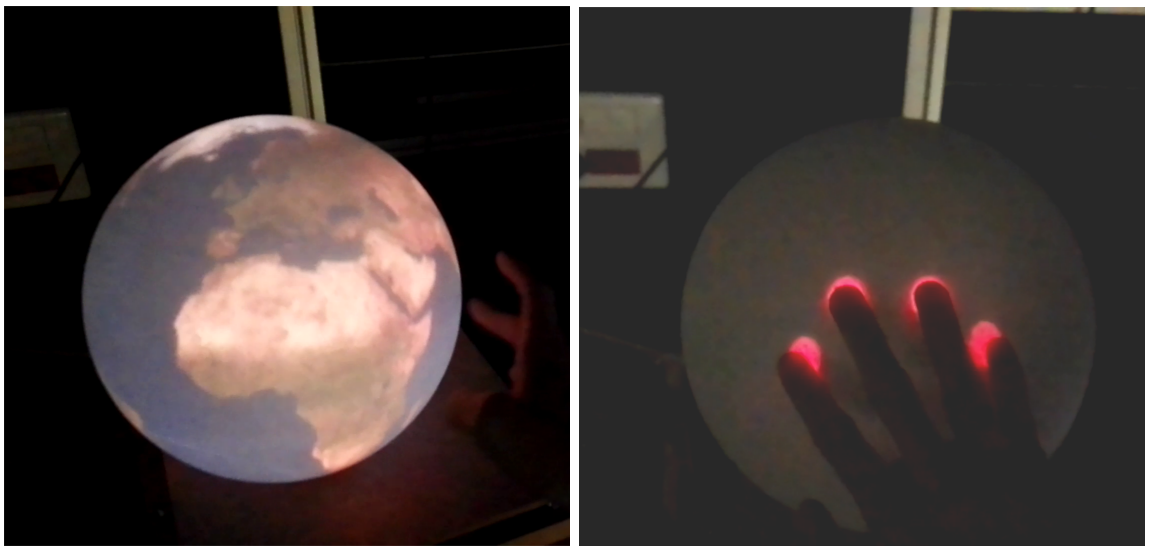
projector, lenses, cold mirror, and webcam are made by 3D printing. The different parts hold together thanks to 5mm metallic rods and shaft collars. As shown in Figure 4.3(a), the prototype holds vertically thanks to a homemade wood enclosure. We run the software on a laptop HP ZBook G3 with a 16Go RAM, a 2.60GHz processor and a NVidia Quadro M2000M graphic card, and results of the described software can be seen in Figure 4.13. Alternative Earth data can be seen in Figure 4.13.

In 2017, the cost of this project was less than 500€: the projector representing 340€, the fisheye lens 30€, the cold mirror 25€, and the other components costed less than 15€ each. We do not include the laptop in the calculation, and it is probably possible to run the display from a Raspberry Pi or similar embedded systems. In comparison, existing products such as Puffersphere™ [Ltd02] and Magic Planet™ [URF08], cost several thousands of euros.

## 4.7 Perspectives of improvements

### 4.7.1 Projection efficiency

As stated in Section 4.5.2, 55% of pixels are used. This is due to the aspect ratio of the projector: the projected image is rectangular with a 16:9 aspect ratio, and we only collect a circular portion of it. We currently use a circle whose diameter fits



(a) (b)

***Figure 4.13:*** *(a) Picture of our final prototype showing a terrestrial globe; (b) Demonstration of the multitouch capabilities.*

(a) (b) (c)

***Figure 4.14:*** *Examples of other Earth data (textures sources: nasa.gov). (a) Global transportation; (b) Night lights; (c) Global temperatures in April 2003.*

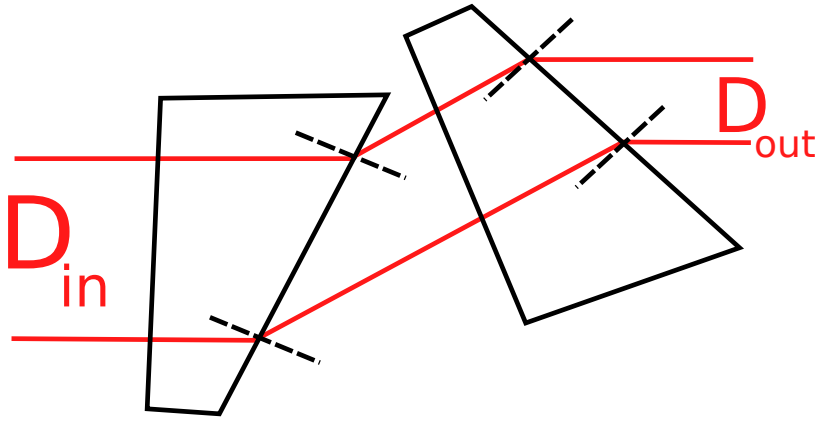



***Figure 4.15:*** *A pair of prisms can shape a beam in one dimension without affecting the other.*

roughly the vertical expansion of the image. This has two main consequences: a loss of resolution and brightness. It is possible to use an anamorphic optical system to squeeze the image into a 1:1 aspect ratio. In the latter, one can achieve up to 70% of the total resolution. A pair of prisms [JC89] can be inserted between the projector's output and the first lens L1, in order to squeeze the width until it matches the height. Figure 4.15 shows how an anamorphic prisms pair can shape a beam simply by refracting the rays on planar interfaces with a specific angle. Alternatively, assuming that the first lens has a sufficient aperture to collect the whole width, the anamorphic prism pair can be located between L1 and L3 to increase the height until it matches the width.

Another solution would be to use an afocal cylinder lenses pair that would change the magnification in one direction as with the prisms. However, it can be impractical to insert them in this path because the distance between them would also be imposed for the focus-free property to be preserved, as explained in Section 4.4.5.

Another approach would be to use a controllable MEMS-laser assembly, so that more degrees of freedom are available to optimize the energy loss. In particular, we might be able to run the laser on a square trajectory instead of a rectangular one, or even a circular shape so that the image would better fit the circular aperture of the lenses.

### 4.7.2 Frustrated Total Internal Reflection (FTIR)

We currently use a rear diffused illumination setup for the finger tracking. An alternative approach is the FTIR, in which infrared light is guided in the thickness of the projection surface and escape at the fingers positions (see Section 2.3.2) can

also be considered. The main advantage of the FTIR technique over rear diffused illumination is that the background image would have much lower intensity because no light is reflected toward the camera when no fingers are present. This would improve the robustness and precision of our tracking.

On planar surfaces, as illustrated in Figure 4.16(a), if a ray hits the interface with an angle greater than the critical angle, then the ray propagates with the same angle on both interfaces. As far as we know, the FTIR technique has never been used on spherical surfaces, so we explain briefly why this is possible. To this end, consider Figure 4.16(c):

- If the ray hits the exterior surface with an angle greater than the critical angle, several cases can happen:
  - (in orange) The reflected ray can hit the interior surface, in this case the interior angle is greater than the outer one, so the ray still reflects. Because of symmetry, the angles are equal two by two, and so the ray can propagate.
  - (in green) The ray can be reflected toward the exterior surface. In that case, by symmetry, the angle is kept and the propagation is possible.
- (in blue) If the ray hits the inside surface first, the reflected ray can be either transmitted or reflected. If reflected, the angles are the same two by two, so the propagation is ensured.

The problem with this technique is that it requires the surface to be transparent to allow light to propagate inside. At the same time, the surface must be diffuse so that the projected light is visible. For planar surfaces, this is generally done using a diffusing sheet attached to the surface: a tiny air gap in-between ensures the total reflection property while the projection is done on the diffuser. We have successfully implemented FTIR on a transparent acrylic sphere (Figure 4.16(d)), but did not find a way to make it diffuse at the same time: a diffusing sheet cannot be easily applied to a non-developable surface, projection sprays break the total internal reflection, and sandcasting alters the normals of the surface.

### 4.7.3 Towards a 3D spherical display?

We have also studied whether this display can be adapted to a 3D display. First, let us consider what kind of effects should a 3D spherical display reproduce. We can imagine for example a planet with satellites orbiting above it, the 3D information would be a great help for experiencing distances better than with a 2D display. However, the question of positive parallax arises: if information is intended to be displayed in front of the display, then depth inconsistencies may occur as the user turns

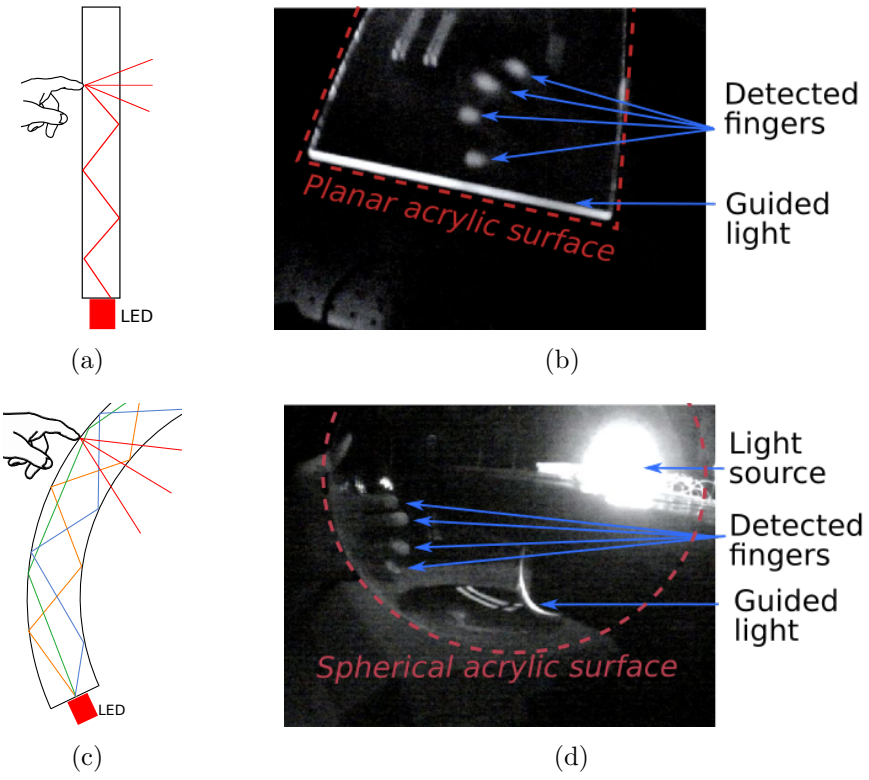



***Figure 4.16:*** *Illustration of FTIR on (a-b) planar surface: faces are parallel so the angles remain the same; (c-d) spherical surface: a ray can propagate by bouncing on both interfaces or only the outer one*

around the globe. Indeed, it is not possible to provide a positive parallax at grazing angles because no light is generated from above the surface, so depth inconsistencies may occur as the user turns around the globe. The curvature of the globe limits the parallax to negative, that is to say that such a display would only reproduce a 3D object located inside the sphere.

Ferreira et al. [TKC+] track the location of the user to provide perspective-corrected content on a spherical display, giving a 360° motion parallax as a strong depth cue to perceive a 3D object located inside the sphere. However, each eye sees the same parallax. Bolton et al. [BKV11] have the same approach but with a stereoscopic projector and shutter glasses. This can be considered a stereoscopic 3D spherical display, and an autostereoscopic 3D spherical display would give the same results without the need of tracking or glasses.

Even though such a display would share some similarities with volumetric displays such as Pespectra™ [FNH+02] reviewed in Section 3.4, both approaches are totally different because here the light comes from outside the volume instead of directly from the displayed voxels. Our display hardware is also hardly compatible with a holographic display. As for light-field or autostereoscopic displays, everything is about addressing different rays from every points on the surface.

Full parallax would be mandatory because, compared to a flat display, the horizontal position would also be crucial to perceive the object without distortion. To view the whole globe, a 180° FOV is necessary, but a smaller FOV can be considered, as this would simply limit the size of the 3D object to display.

For the integral imaging approach, a very high resolution would be required under the lenses array, and crosstalk would be inevitable at grazing angles. Using other approaches such as directional diffusers, the large range of required input angles seems difficult to address from the inside without occlusion. A 2D array of picoprojectors and fisheye lenses combined with one or several hemispherical mirrors and a custom directional diffuser might help in addressing different ray directions as illustrated in Figure 4.17. This is subject to future work.

## 4.8 Conclusion

We have introduced a more affordable spherical multitouch display than previous approaches. Its principal components are the projection through a fisheye lens via a laser-beam steering projector and optical finger tracking for multitouch. We offset the use of inexpensive material by software calibrations and corrections. An example of the flagship application of spherical multitouch applications, an interactive globe, can be seen in Figure 4.3(a).

The systems the most related to ours are the work from Benko et al. [BWB08]

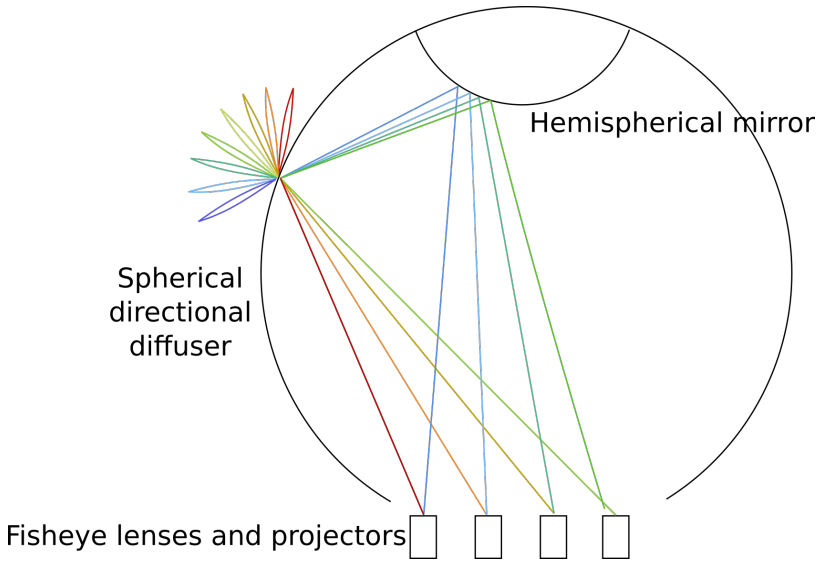



***Figure 4.17:*** *Sketch of a potential solution for a 3D spherical display. An array of fisheye lenses and projectors sends different rays on the sphere. A hemispherical mirror can increase the number of input directions, but of course occludes the north pole. A spherical directional diffuser with a custom optical function might scatter rays into specific directions to cover the FOV.*

and Patel [Nir11]. Compared to Benko et al. [BWB08], we only require off-the-shelf hardware components, and in particular, not a specifically tailored fisheye lens. Compared to Patel [Nir11], our system is multitouch and, thanks to our optical system, covers the entire sphere (except the opening neck).

We have tested our spherical multitouch display with several users, experts as well as non-experts, that were enthusiastically able to use our system without any difficulty. The main drawback of the proposed system is the lack of brightness that limits its use in controlled lighting environments. First, replacing the sphere by a spherical diffuser specially designed for rear projection can improve the transmittance quality and thus brightness and contrast. But this is also inherent to the projector, which has only 30 lumens. Despite that, LBS projectors are still a very interesting tool and we will continue to use them in this thesis. Indeed, they have a small size, provide a great contrast and color gamut, they are focus-free, and the laser nature of the projected light is also interesting in many aspects. We also expect their brightness to increase in the future.

Finally, we saw that extending our solution to a 3D spherical display would be interesting, however, the difficulty of generating multiple views all around a globe is very challenging. We decided instead to focus on generating multiple viewpoints from a planar surface to create a 3D planar display. To that extent, we investigate other types of imaging devices, and particularly wedge light guides that we introduce in Chapter 5.

# Chapter 5

# Wedge Camera for minimally invasive archaeology

In this chapter, we first introduce wedge light guides, which are optical light guides that have been used in the literature for both imaging and displays [TPZM00, TLEB13]. We then present a new imaging device prototype that is based on a wedge guide and its applications in the field of archaeology. Such a guide is also one of the core parts of the 3D display described in the following chapter. This work has been published in the Journal on Computing and Cultural Heritage [CTRG19].

## 5.1 Introduction

Imaging techniques have long played a major role in the acquisition, visualization and conservation of cultural heritage. In confined spaces, the imaging distance may not be enough for classical imaging techniques to work at their best. Indeed, taking a picture of a side of a narrow slit for example may be difficult, if not impossible, to achieve with a satisfying angle, focus or imaging area, and this work addresses this issue. The need for a minimum imaging distance might seem obvious: we all open a book in order to read it and we stretch out our arm in order to take a selfie. Archaeology is one of several disciplines like surgery and aircraft inspection where specialists seek to learn as much as they can with minimal disruption, so ensuring an imaging distance is not always possible in confined spaces.

As an example, in the archaeology of buildings, a standing structure can be studied and documented with both laser and photogrammetry [HG02], but there are inaccessible areas that require other dedicated devices such as radar measurements [BR04] (notably for underground analysis) or endoscopes [Bec15]. Endoscopes are indeed the most straightforward way to image places that are inaccessible by regular cameras,

but they still require that there be a chamber, i.e. a minimal distance between the surface to be inspected and the sensor. Even if we assume that this minimum imaging distance can be achieved in some specific confined spaces, the imaging area would be very small and numerous pictures would be necessary to get a full view.

The system presented in this chapter is between an endoscope and a document scanner: it can image inaccessible areas like endoscopes but larger images can be obtained without any imaging distance just like a document scanner. Unlike a document scanner, distant objects can also be imaged. Our system is based on the association of a regular camera and a wedge light-guide developed by Travis et al. [TPZM00, TLEB13], and we explore the utility of such a device in the specific field of archaeology.

This chapter is organized as follows. In Section 5.2, we present the wedge guide technology, core of both our imaging system for constrained environments and our 3D display described in the following chapter. Then, in Section 5.3, we explain the hardware and software requirements for the prototype, and we analyze its characteristics. In Section 5.4, we introduce the different applications and on-site results. Finally, in Section 5.5, we present directions for future work, before we conclude in Section 5.6.

## 5.2 Wedge light guides

### 5.2.1 Principle

As depicted in Figure 5.2, a wedge light guide is roughly composed of two parts: an expansion part and a wedge part. Let us consider a ray shone into the entrance of the expansion part as shown in Figure 5.1. The faces of the expansion part are parallel to each other like in a regular slab, and so as long as the angle of incidence $\theta_0$ is greater than the critical angle $\theta_c$ the ray will be guided through total internal reflection and will keep the same angle $\theta_0$.

In the wedge part, one face is tilted relatively to the other, so that at each bounce on the tilted surface, the angle of reflection decreases. If $\alpha$ denotes the angle of the tilted surface and $i$ the number of bounces experienced by the ray in the wedge part, then $\theta_i = \theta_{i-1} - \alpha$. Eventually the critical angle is reached and the ray is no longer reflected but instead passes into air.

The exit angle $\theta_{out}$ of the ray is near 90° because it just has reached the critical angle. More formally, the last ray inside the wedge has an angle $\theta_{last}$ with the normal so that

$$\theta_{last} \in [\theta_c - \alpha; \theta_c]$$

and

$$n \sin \theta_{last} = \sin \theta_{out}$$

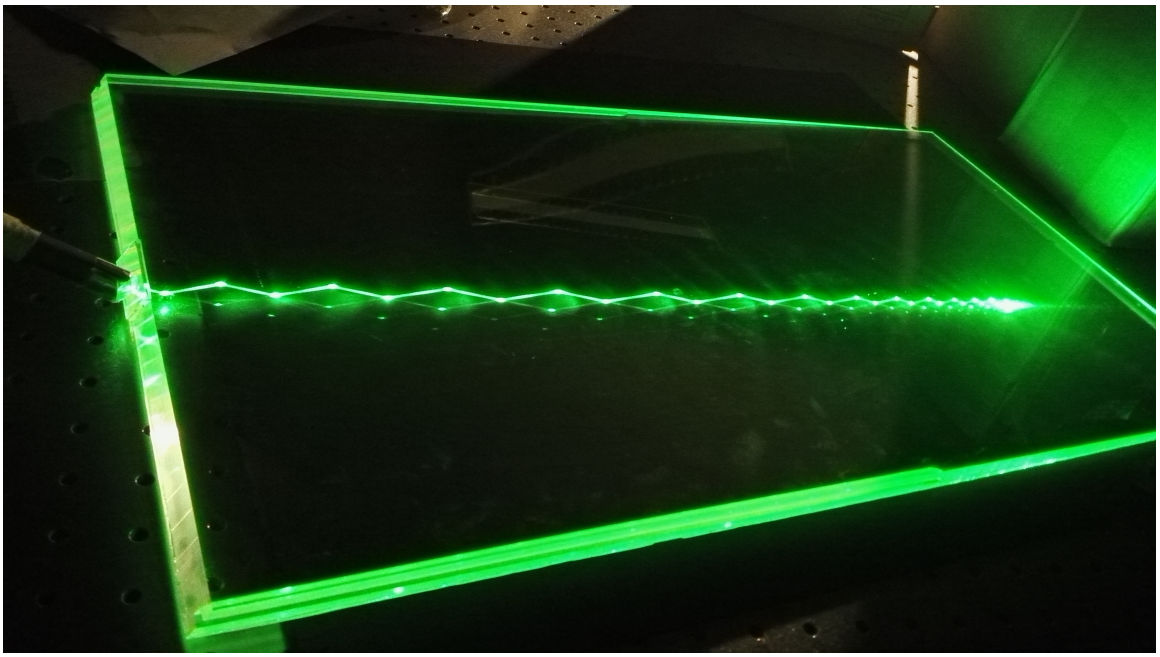


***Figure 5.1:*** *Picture of a laser beam propagating in a wedge guide*

so

$$\theta_{out} \in [\sin^{-1}(n \sin(\theta_c - \alpha)); 90°]$$

With $\alpha \approx 5°$, we obtain $\theta_{out} \in [74°; 90°]$.

Presented as such, the rays are as likely to leave from the top side as from the bottom. Patent WO/2003/013151 [Tra03] describes a method to force rays to undergo the same number of reflections by making the top surface slightly curved. This surface is later optimized with a ray-tracing software in order to enforce output rays to be collimated, so the angle interval of $[\sin^{-1}(n \sin(\theta_c - \alpha)); 90°]$ is actually much tighter. We will not detail further these steps because we did not manufacture wedge guides ourselves. Instead, we use off-the-shelf wedge guides that are designed upon the cited patent.

### 5.2.2 Beam shaping

Let us now consider what happens to a bunch of rays, as those generated by a projector. Referring to Figure 5.3, the wedge guide lies on the XY plane, and we will explain clearly how the rays behave in both XY and XZ planes.

Figure 5.3(a) illustrates how the angle of injection determines the distance travelled by the ray before it leaves the guide. The dashed red line represents a ray that

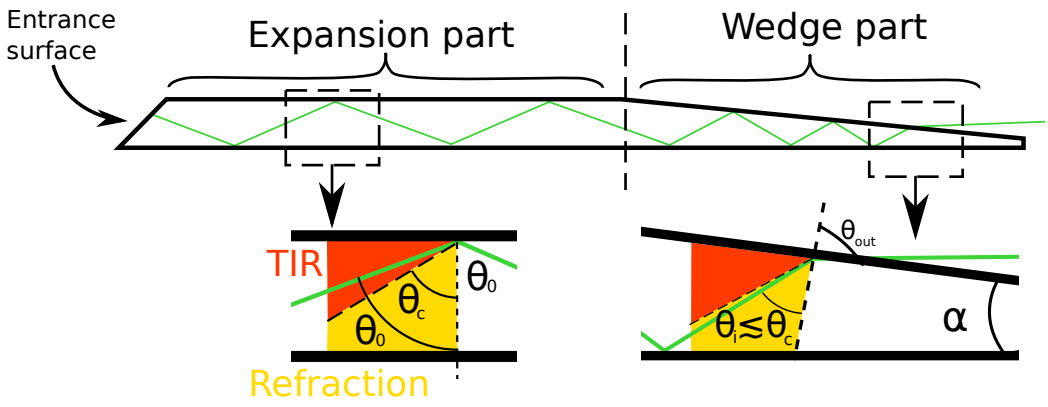



***Figure 5.2:*** *Principle of a wedge guide. In the expansion part, a ray propagates through total internal reflection. In the wedge part, the angle of incidence lowers at each bounce until it leaves the guide.*

is launched so as to be near the critical angle, and then the ray needs only a few reflections within the wedge part before it reaches the critical angle, and so will not travel far. If, however, the ray is launched at a small angle to the slab axis, such as represented by the plain green line, then many reflections will be needed within the wedge part before the ray exits. A wedge guide then turns a difference of angle $\Delta\theta$ into a difference of length $\Delta D$, or reversely. The smaller the angle $\alpha$, the larger $\Delta D$ gets. In this plane, the rays leave the guide collimated and with a grazing angle close to 90° with the exit surface's normal.

Seen from above (Figure 5.3(b)), the role of the "expansion part" becomes clearer: it allows the horizontal expansion of the image. The longer this part, the larger the image gets before exiting the guide. When the rays exit, they still diverge in the XY plane. In this figure, the plane delimited in dashed lines exits at the beginning of the wedge part, while the plain lines propagate until the end. On the one hand, if the projected angle is large enough to fully cover the wedge surface, then the beam exits with the same width but a different angle of divergence. On the other hand, if the beam does not totally cover the wedge surface, then it will leave the guide with an increasing width with the distance, while the angle of divergence stays the same.

To sum up, a diverging beam projected through a wedge guide exits with a peculiar shape (Figure 5.3): it gets collimated in the vertical direction and diverges in the horizontal direction. Moreover, it diverges from the wedge with a grazing angle, at an almost horizontal direction.

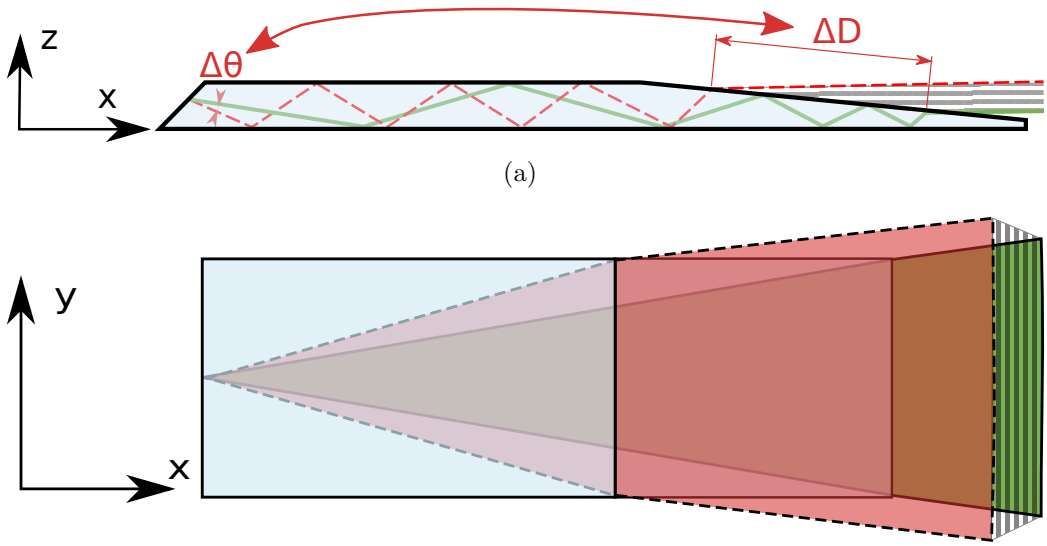



***Figure 5.3:*** *Top and side views of how a beam propagates and exits a wedge guide: collimated in one direction (a) and diverging in the other (b)*

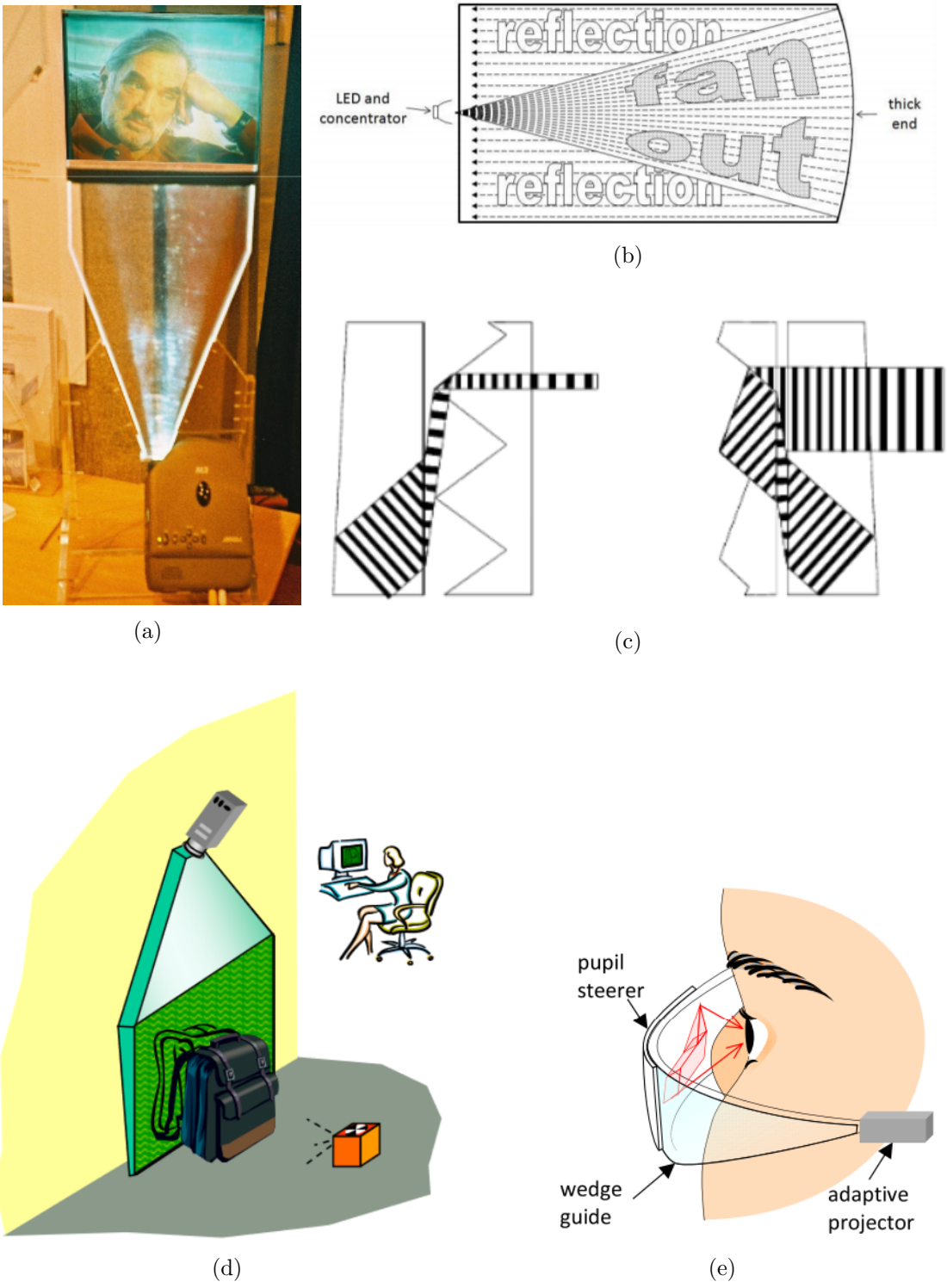
### 5.2.3 Applications of wedge guides

From the above description, the immediate application is to use a wedge guide associated with a projector, and this was indeed the earliest application of wedge guides [TLB04]. A "rear projection television" can be built by folding up the required projection distance in the guide as shown in Figure 5.4. The horizontal expansion of the image is done through the expansion part, and the vertical expansion is done through the wedge part.

It can also be used as a backlight unit for a LCD panel [TLEB09]. In this case, the wedge design can be made more compact. As shown in Figure 5.4(b), the expansion part and the wedge part can be the same. An array of micro-mirrors at the thick end make the ray travel back and forth before exiting the guide, and the rays are collimated in both directions thanks to the curvature of the thick end.

If several light sources are distributed along the thin edge, then a directional backlight can be made. By lighting up sequentially the LEDs at the thin end and switching the LCD accordingly, a time-sequential autostereoscopic display has also been developed [MT05].

Since the propagation of rays can be reversed, wedge light guides were also studied as cameras [BLB+06]. The first application was to image objects that are directly on the wedge surface by placing a camera at the thick end and the object on the wedge part surface. This has the advantage of folding up the space required between




(a) (b) (c) (d) (e)

***Figure 5.4:*** *a) A projected image is guided in a wedge guide and visible on a diffuser located on the wedge exit surface (source: [TLB04]); b) More compact wedge design allowing collimation in both directions (source: [TLEB13]); c) Two kinds of turning films acting in reflection (source: [TLEB13]); d) Wedge X-ray imaging device for baggage inspection (source: [TLB04]); e) A curved wedge guide for near eye displays (source: [TCG+18])*

a camera and an object, acting somehow like a document scanner. This means that it is possible, for example, to image a page in a book without opening it, simply by sliding the wedge between the pages.

It is important to note that if the object is not directly on the wedge surface, then the camera cannot capture it because the only rays guided toward the camera are those incident on the wedge surface with a grazing angle.

In order to change the direction of observation, two kinds of turning films were studied [TLEB13]. They consist of a 1D array of prisms, located at the wedge surface and acting like tiny mirrors. Either the rays are reflected by the tips of the prisms to a perpendicular direction, or by their entire surface but pass through the wedge thickness again before going to air. We will detail further the turning films in Section 5.3.2.

Basically, the idea of coupling a projector and a camera to allow human-computer interactions is also mentioned. For instance, if a diffuser is used to display a projected image, then an infrared camera might be able to see touch interaction on the diffuser surface (like in our previous chapter).

In the same way as the backlight unit of Figure 5.4(b), the geometry of the guide can be made more compact for either projection or sensing. The micro-mirrors array at the thick end, though, tends to severely blur the image by diffraction.

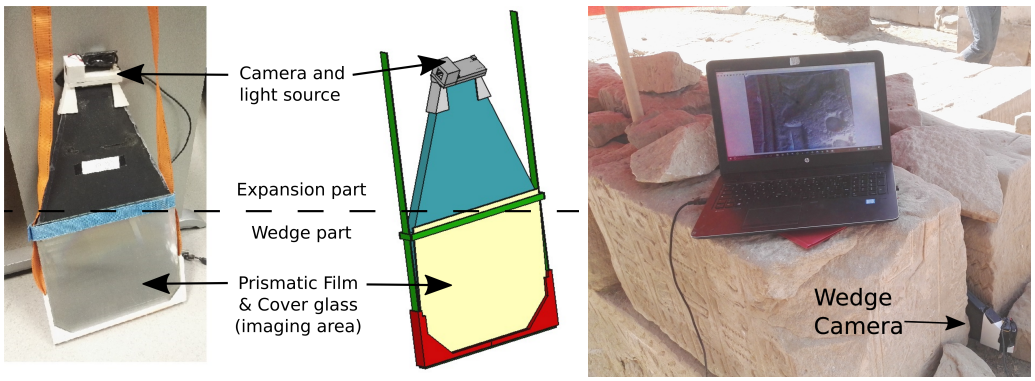
Another successful application of wedge guides as cameras mentioned in [BLB+06] is an X-ray imaging device for baggage inspection as illustrated in Figure 5.4(d).

More recent work on wedge guides studies its use for near-eye displays and mixed reality applications [TCG+18]. In this work, the wedge guide is not flat but curved, and intended to be worn as glasses as shown in Figure 5.4(e). A volume hologram scatters light toward the eye, and the whole FOV of the moving eye is covered by shearing the volume hologram.

## 5.3 Wedge camera design

### 5.3.1 Overview of the prototype

In this section, we describe the implementation of our wedge camera prototype specifically designed for use in an archaeological context. Our prototype, shown in Figure 5.5(a), features a wedge guide made of polymethyl methacrylate (e.g. Plexiglas$^{TM}$) and whose dimensions are 455mm x 215mm and a 25mm maximum thickness. The imaging area is of size 260mm by 215mm. The wedge was designed according to the principles outlined in the international patent WO03013151 [Tra03]. The surfaces are protected by black cardboard except at the imaging surface, which is covered by a prismatic film for deflecting light rays.




(a) (b)

***Figure 5.5:*** *(a) Annotated picture and scheme to identify the geometry and the different parts of our device. (b) The device used for inspecting rocks.*

A 3mm thick glass cover slip is glued onto the wedge entrance to ensure its flatness. At the wedge entrance, there is a mount made by 3D printing that holds a camera and the light source. The prismatic film and the cover glass are held against the wedge surface with another 3D printed mount and straps. These straps are also used to bring down the wedge into trenches.

After presenting how the prisms can deflect light rays in Section 5.3.2, we detail the resulting equivalent camera in Section 5.3.3 and the illumination system in Section 5.3.4. The prototype was connected to a computer software to visualize and record images corrected from distortion as we detail in Section 5.3.5.

### 5.3.2 Deflecting light rays

As explained in Section 5.2, the rays need to be deflected in order to observe a distant object, otherwise we observe rays arriving nearly parallel to the surface as shown in Figure 5.6(a). To this end, we propose to use a prismatic film, which is an optical surface that is composed of tiny prisms elongated in one direction. A prismatic film can turn the direction of the rays in different ways as illustrated in Figure 5.6. Note that in this figure, the red bars are showing the wavefront of a ray.

Figure 5.6(b) shows that the tips of the prisms can be used as mirrors to deflect the rays to a normal direction. The problem of this method is that the reflection happens over a tiny region, and so diffraction tends to blur the image as shown in Figure 5.6(e). As we explained in Section 2.1.4, a small slit diffracts light in a cone whose angular extent is inversely proportional to the width of the slit. This limits how small can be the pixels of an image projected through the slit and, equivalently,

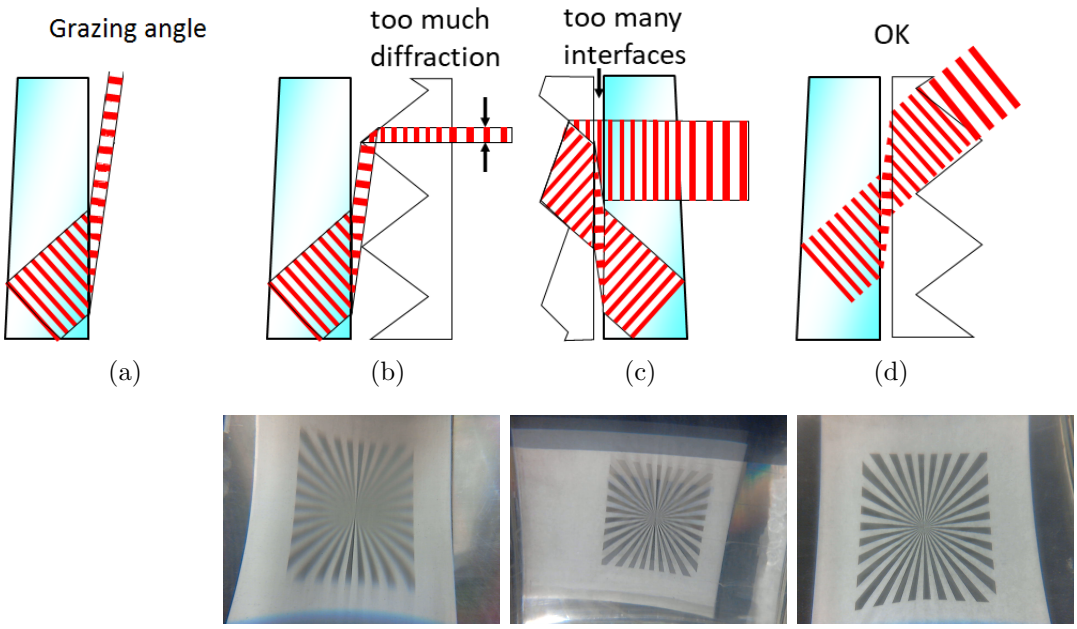



***Figure 5.6:*** *(a) Without a prismatic film, distant objects cannot be imaged; (b) configuration displaying strong aperture diffraction; (c) configuration displaying too much interface reflection; (d) configuration minimizing both of the latter at the cost of tilting the angle of observation away from the normal; (e) blurred image at about 7 cm with configuration b; (f) ghost image effect with configuration c; (g) sharp image at about 7 cm with configuration d.*

limits the resolution of the picture that can be imaged via the slit. Each prism of the prismatic film acts like a slit and so the bigger the prisms, the better is the resolution of images captured distant from the screen. However, arrays with big prisms have big transients between each prism and this means that an object close to the surface of the film cannot be imaged in its entirety. Note that this configuration was tested in an archaelogical context previously to this work, in Deir El Medina, Egypt, but the image quality was not satisfactory enough due to the diffraction.

A possible solution to reduce diffraction might be that the prism array acts in reflection in the manner shown in Figure 5.6(c), but the light then has to cross many boundaries before it leaves the system, and each boundary lowers the brightness and adds an unwanted reflection to the image. Anti-reflection coatings might help but would be affordable only in mass production.

While it is important for use in the human-computer interaction (e.g. gesture interaction) that rays are captured when incident perpendicular to the wedge surface, this matters much less for our application. So instead of using prisms as reflector, we use them in refraction with the flat surface close to the wedge and the tips facing the air as shown in Figure 5.6(d).

A thin air interface between the film and the wedge guide allows the total internal reflection to act as in a regular air-acrylic interface. Light escapes the guide at the grazing angle and immediately enters the material of the film at the critical angle that is approximately 43°. We chose a 45° prism array, so when the light meets the prism's surface, it travels straight on into air at the same angle because its incident angle on the facet is close to 0. In contrast to Figure 5.6(b), the aperture seen by the ray is no longer a portion of the prisms but the full prism aperture, so the diffraction effect is limited, and there are no extra surface reflections such as in Figure 5.6(c). In this way, the rays are deflected away from the normal but the acquired image is sharper as shown in Figure 5.6(g). The prismatic film we used for this experiment is the 25FD330 from Comar Optics[1], with a pitch of 0.3mm and a prism angle of 45°.

### 5.3.3 Equivalent camera description

Figure 5.7 shows the FOV of our wedge camera system compared to an endoscope or a regular camera. The required distance between the device and the object to get a large imaging area is much shorter with our wedge device than with a regular camera or endoscope. Moreover, the focal length of any endoscope lens defines a minimum distance between lens and object whereas in our system, this distance is folded up into the wedge guide and the object is in focus directly on the wedge surface.

[1]https://www.comaroptics.com/components/plano-optics/prisms/fresnel-prisms-and-beam-dividers

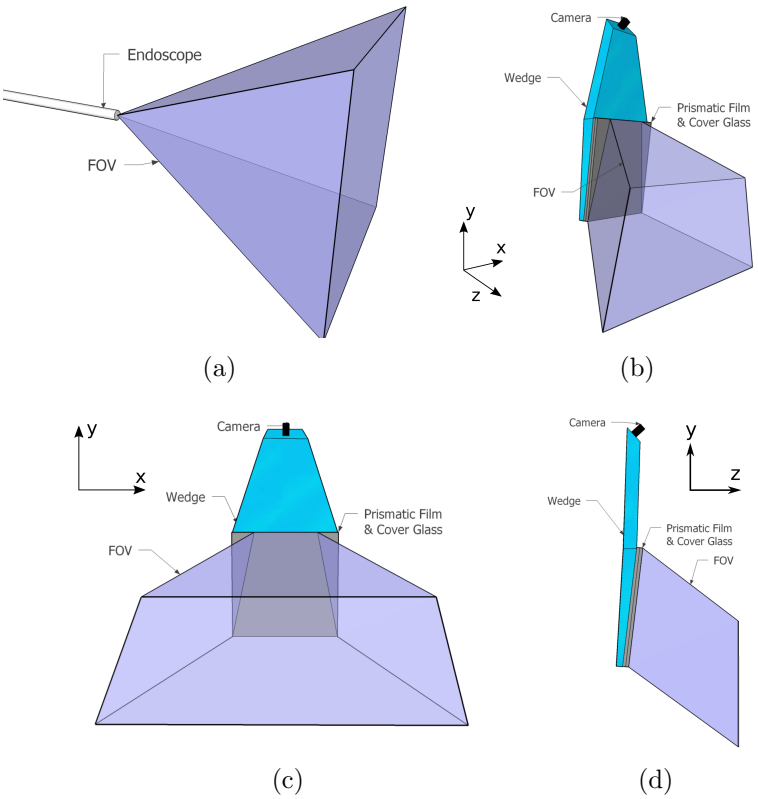



***Figure 5.7:*** *(a) Typical pinhole FOV of an endoscope or a regular camera. (b) FOV of the proposed device. (c,d) Front and side views of b.*

Let us analyze the resulting FOV in more detail. Once again, for the sake of clarity, the following explanations describe rays leaving the wedge but readers must keep in mind that for acquisition, rays propagate from the object, towards the wedge guide and the camera at the thick end.

As explained in Section 5.2, the wedge width (in $x$ direction in Figure 5.7(c)) allows rays to fan out inside the guide, and the prism array does not change this property when they leave the guide. Rays still fan out horizontally when leaving the wedge as in a regular perspective camera. The difference is that the FOV forms a sort a trapeze because rays that leave closer to the camera fan out differently from those leaving the wedge at the end as shown in Figure 5.7(c). This phenomenon is similar to a keystone distortion due to off-axis projection or sensing. Rays within the wedge are constrained in the $z$ direction and, thanks to the prismatic film, they all escape at the same angle close to 45° with respect to the wedge's normal. At the same time, they are parallel to each other in the $yz$ plane as shown in Figure 5.7(d),

so the vertical FOV is orthographic.

As a consequence, one can observe a 1D perspective: the images appear to be squeezed in $x$ direction with an increasing distance of the object from the device, whereas the vertical direction $y$ remains the same. This is illustrated in Figure 5.8 where we imaged the same checkerboard at different distances, from 0 to 30cm away from the device: the ratio $y : x$ changes from 1:1 at 0cm to about 1:2 at 30cm. Note that the images get blurrier with an increasing distance due to the diffraction by the prisms. In our use context, this is not an issue since we target the acquisition in confined spaces where the objects to image are close enough.

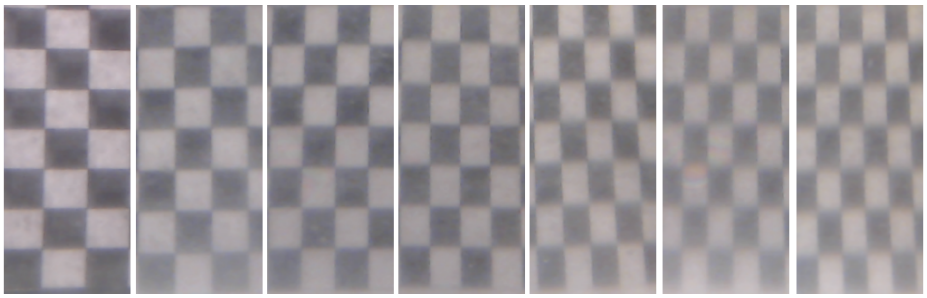


***Figure 5.8:*** *The effect of the 1D perspective on a checkerboard at (from left to right) 0cm, 5cm, 10cm, 15cm, 20cm, 25cm, 30cm away from the wedge surface.*

Our imaging geometry is not standard, but has similarities with the pushbroom camera model [GH97]. In a pushbroom camera, a linear sensor acquires a 1D image of a scene and moves perpendicularly to the sensor direction to build a 2D image. The integrated FOV of a pushbroom camera over a finite period of time hence resembles the described FOV, with the difference that our device does not show time-related effects (apart from an eventual rolling shutter of the camera).

We would like to emphasize the fact that only rays incident with specific directions are guided through the wedge, so most of them are not collected. This is actually how every camera works: the lens aperture selects rays incident at a specific location with specific angles defined by the FOV and the wedge is just a way to redirect light toward the camera aperture. The depth of field, exposure time and other camera-related parameters are defined by the choice of the camera. The required specifications are a 60° horizontal FOV in order to benefit from the maximum of imaging area, and an aperture that is smaller than the wedge entrance (25mm high). The camera should also have either a fixed focus with a large depth of field or an autofocus able to focus on the wedge surface and beyond. For the on-site experiments, we used the Logitech C310 webcam, however, many other cameras could be used as well. We later used the higher resolution Logitech C922 webcam, which produced even better results.

### 5.3.4 Illumination

The simplest way of illuminating a surface under inspection is to shine light from a point next to the camera. The surface will then receive uniform illumination whereas lamps placed at the sides of the wedge would illuminate the edges of the scene more than the center. Experiments in the laboratory have shown, however, that a fraction of the illumination reflects off the prism facets back towards the camera and dazzles the camera along a vertical line. A simple yet effective way to solve this is to rotate the prismatic film by approximately 15° about the surface normal as shown in Figure 5.9. The reflected light is then at 30° to the wedge axis so that the dazzling line moves off to one side and therefore out of the way of the surface under inspection. This does not disturb the camera rays that are still deflected at 45°, but adds some deviation to the side of the guide. Referring to Figure 5.7(c), with this prism configuration, the FOV is not symmetric but angularly shifted.

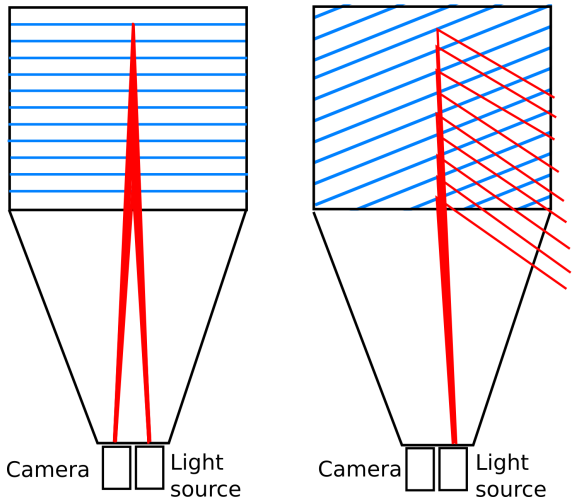



***Figure 5.9:*** *By rotating the prismatic film, the camera is no longer dazzled by the specular reflection of the light source.*

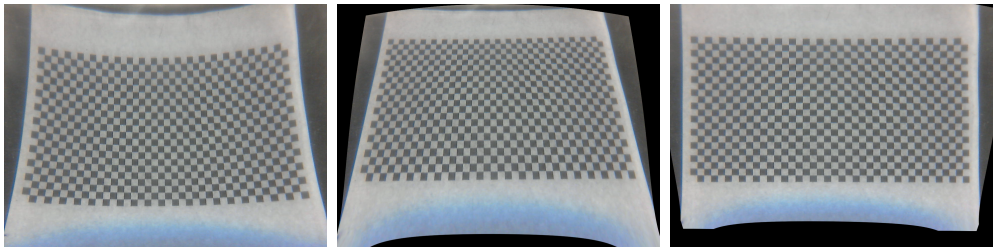
We just described a way to illuminate from the device itself. Other ways of illuminating a scene can also be imagined, such as enclosing the imaging area with an illumination frame. Compared to front illumination, this would exhibit more shadows on the inspected area but would be much less uniform. Illumination from the back face of the imaging area can also be considered.

In practice, we relied most of the times on sunlight to illuminate areas under inspection because it was sufficiently intense by workdays and provided nice shadows on the inspected areas. The proposed illumination system would be more useful in darker areas.

### 5.3.5 Software and calibration

The acquired raw images of the wedge guide cannot be used directly as there are strong distortions that arise because the wedge and the camera behave somewhat like a camera and a Fresnel lens in an off-axis configuration. Consequently, we propose to correct the acquired images in real-time with image processing, just as is done in other computational imaging systems. We developed a software solution in C++ with the OpenCV computer vision library that processes the stream of the camera in real-time. Hence, we can directly visualize and record the corrected acquired images.

The distortions consist of a pincushion distortion as well as a strong keystone distortion. For the calibration, in a pre-processing step, we use a series of images of a checkerboard in order to calculate the correction for the pincushion distortion (Figure 5.10(b)). Then, we identify the corners of the checkerboard in order to compute the homography matrix for the keystone correction (Figure 5.10(c)). After this calibration, the corrections are applied in real-time for any frame of the camera stream in order to provide the archaeologists with the correct visualization.



(a) (b) (c)

***Figure 5.10:*** *Software correction: (a) Raw image; (b) pincushion distortion correction; (c) pincushion and keystone distortion correction.*

## 5.4 Case study

We tested our wedge camera at two archaeological sites in Egypt near Luxor. Prior to this work, a wedge camera with prisms acting as reflectors was tested at Deir El Medina, but the diffraction was too strong and the image quality not satisfactory for objects at even a small distance. We took part of an archaeological mission in Medamoud to test our prototype on-site for various applications and get the feedback of expert archaeologists and researchers.

We identified three concrete applications where wedge cameras can aid inspection in confined spaces, and we present results and what we learnt in the following.

First, wedge cameras can image the underground through slim trenches (Section 5.4.1), second, underwater surveys can be done by imaging flooded trenches (Section 5.4.2), and third, areas in confined spaces such as structural crevices or between stones in a wall can be imaged (Section 5.4.3).

### 5.4.1 Trial trenches

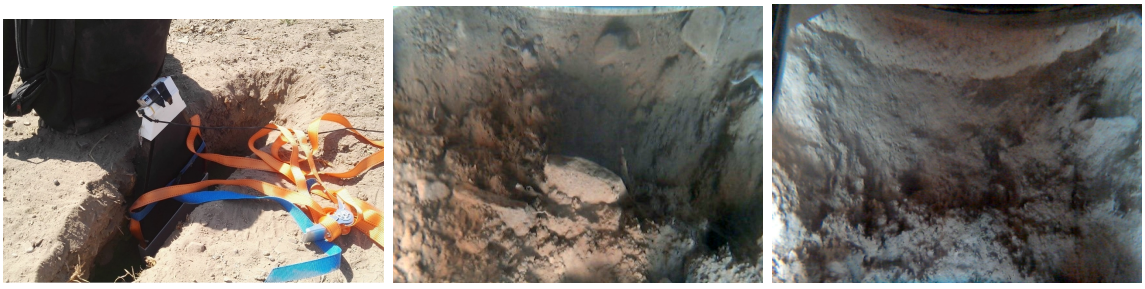
Trial trenches are usually dug before a full excavation is launched so that archaeologists can look at layers of soil beneath the surface and see if there are remains dating from layers corresponding to the epoch of interest. The trenches must be big enough for someone to get into, i.e. about 1m wide, typically 2m long and from one to several meters deep. Digging such trenches is time consuming, and, moreover, it is moderately invasive and this will bear on the decision of local authorities to grant permission to dig. With the wedge, it is possible to get visual information from a trench and be much less invasive.

Digging a deep trench that is only 5 to 10cm wide proved difficult in European clay soil and would no doubt be difficult in dry sand. However, on the site where the wedge cameras were tested in Medamoud, the soil is less prone to flow or crumble, and a simple trenching tool was enough to dig trenches to arm's length as recorded in Table 5.1. Rays undergo approximately 20 internal reflections within the wedge guide and so its surfaces must be especially smooth (roughness average of 1nm) and clean. It might therefore seem optimistic to have placed the wedge in trenches swirling with grit and dust but although there was some degradation of the image, it was not drastic.

| Trench | Length at the surface | Width at the surface | Length at the bottom | Width at the bottom | Depth |
|---|---|---|---|---|---|
| 1 | 75cm | 14cm | 28cm | 11cm | 70cm |
| 2 | 45cm | 18cm | 26cm | 7cm | 70cm |
| 3 | 55cm | 20cm | 25cm | 8cm | 63cm |

***Table 5.1:*** *Dimensions of the trenches.*

The wedge camera was inserted and it quickly became apparent that the soil is much darker than one might expect. Illumination that was perfectly adequate in the laboratory gave no useful image in the trench and the best images were obtained by allowing sunlight to leak down into the trench. Time and again, we found ourselves experimenting with different ways of illuminating the scene and perhaps this should not be too surprising because free-space photography requires much the same care with illumination.



(a) (b) (c)

***Figure 5.11:*** *(a) The wedge down a slim trench; (b-c) Images of walls of trenches acquired from the device.*

Once the illumination had been optimized, images were produced that were much the same as would be obtained through free space. However, archaeologists commented that the differences in reflectivity and texture that delineate different strata are difficult to see unless the sides of the trench walls are carved flat and smooth. Giving walls such a finish is not so difficult in a wide trench but is more challenging when the wall cannot be inspected at the same time that it is being perfected. It might be possible to create narrow trenches with the surface finish needed for useful inspection, for example by using a plough-share.

The authors note that whereas permission was given for conventional trial trenches to be dug only in specific locations on site, trenches adequate for the wedge camera were sufficiently narrow that we were allowed digging them anywhere on site. Moreover, it happens that a ground survey is necessary in areas that are constrained by a structure, so the wedge camera might find many applications for observing the underground in specific conditions.

### 5.4.2 Underwater surveys

It is not uncommon, particularly in Egypt, for the water table to be high so that traditional surveys cannot be performed because they would rapidly be flooded. Our wedge camera system can address this issue as it can work underwater.

The wedge principle as described in Section 5.2 is based on the value of the critical angle, which is the limit between total internal reflection and refraction. This value depends on the indices of refraction $n$ of both optical media: the acrylic that makes the wedge ($n = 1.49$) and air ($n = 1.00$), which is the exit medium. The wedge that we use is designed to work in air, so if the wedge surface is in direct optical contact with the water ($n = 1.33$), the critical angle is reduced, and the rays travel less far within the wedge (Figure 5.12(a)). This phenomenon is clearly unwanted as

the image would be squeezed in one direction.

It follows that an air interface is desirable between the wedge and the water but the water itself will refract light the same way the prismatic film described above does, so the prismatic film may not be needed if the air interface is provided (Figure 5.12(b)).

In our test, however, we used a flexible plastic enclosure and could not ensure the air interface when put inside water. We therefore kept the prismatic film so that the space between prisms acts as an air interface, and we used the system depicted in Figure 5.12(c).

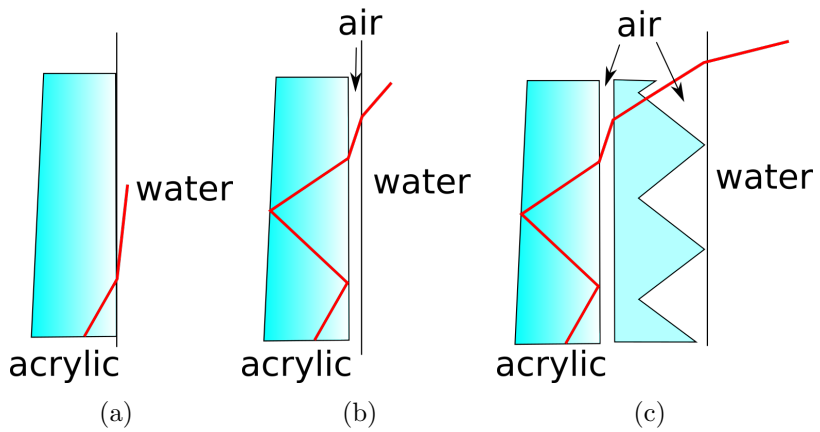



***Figure 5.12:*** *a) The wedge directly in water would make the rays experience less bounces. b) When the air interface is ensured, the prismatic film may not be needed to redirect light rays. c) Our test setup with the prismatic film and a transparent plastic bag, with air interfaces.*

A 10 cm wide trench was dug beneath the water table (Figure 5.13(a)) and then left overnight for the mud to settle. The wedge camera was wrapped in the flexible transparent sheet and lowered into the trench. The captured images were adequate even if rapidly degraded by the upwell of silt, but as in Section 5.4.1, it was difficult to see any changes in reflectivity or texture, perhaps because experts are more used to work in open air. Some tests were also performed to inspect the walls of a flooded chamber (Figure 5.13(b)). We could achieve an image quality that was as good as if the pictures were taken in open air.

We therefore think that our wedge light guide system may also find application in underwater archaeology.

(a) (b) (c)

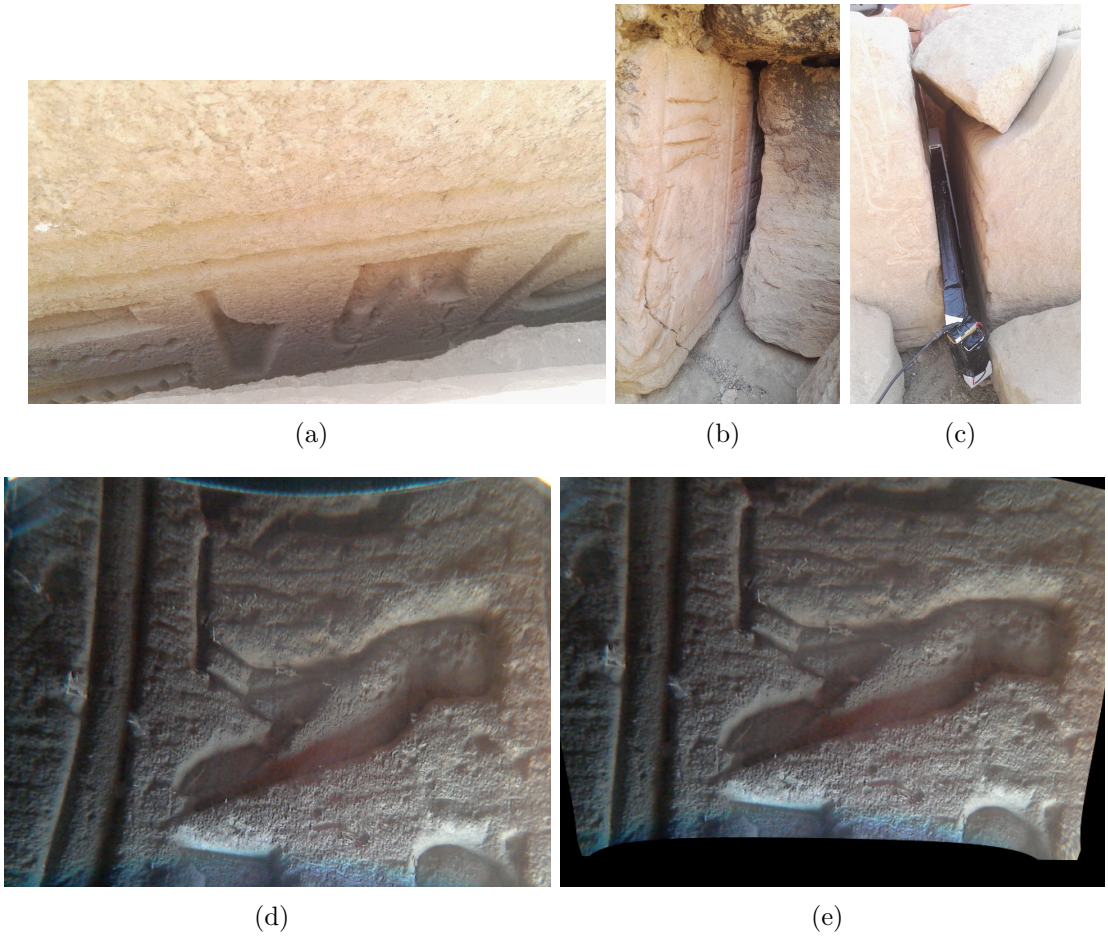
***Figure 5.13:*** *(a) An overflooded trench where the wedge can be used; (b) a wedge being used for observing the side of a flooded area; (c) image acquired underwater.*

### 5.4.3 Structural crevices

Crevices are common in ancient structures where a wall is rarely so smooth that there is no gap between their stones. The interest in these gaps arises because it was not unusual for temples to be disassembled and rebuilt with the stones of previous temples on the same site. Decorations from these earlier temples were often unwanted and so the decorated surfaces were placed out of view adjacent to other stones. These decorations cannot be observed without considerable disruption (such as removing temple stones) unless a wedge is used. Figures 5.14(a-b-c) show situations where the wedge camera can be useful to inspect these constrained areas. The device can notably help a lot in the field of epigraphy, where writings are not always easily accessible. As an example, Figures 5.14(d) and (e) show respectively the uncorrected and corrected image of a hieroglyph taken from the crevice of Figure 5.14(a) with the wedge camera. The wedge was placed at a few centimeters away from the wall yet we could achieve a large imaging area with good image quality.

Most crevices between the stones of a temple do not of course contain images on the opposing surfaces. Proving that they do not is useful in itself and the wedge camera was used to do this in several locations such as that of Figure 5.14(c). Once again, experimenting in order to optimize illumination was important in order to be sure that there really were no images to be seen. The crevice of Figure 5.14(b) was unfortunately slightly too narrow for the wedge used in these trials. Archaeologists suggested that a smaller, phone-sized wedge could be useful for quick inspection on more potential places. They emphasized that it is not only inscriptions that interest them, for example much can be learned about techniques of construction from marks left by the devices used to move the stones and by the procedures used to align them. There was a degree of enthusiasm that the richness of information given by the device

might lead to discoveries.



(a) (b) (c)

(d) (e)

***Figure 5.14:*** *(a,b,c) Typical examples of wall crevices that could be inspected with a wedge; (d) Raw image of a hieroglyph; (e) Image corrected from distortion.*

## 5.5 Future work

This project and our device are opening new directions of research, development and applications. In this section, we explain the prospective work and this could give an idea of its full potential.

### Image quality

Firstly, we observed that the image quality was degraded when the device was used on-site. The sand that was in contact with our glass protective layer inevitably

degraded the image quality. A prototype with a better protection and isolation from dust can greatly improve the results.

Note that the images presented here were a bit modified to improve the contrast ratio and brightness. It could be interesting to develop new filters to offer archaeologists an improved experience with high quality images. Moreover, the calibration process could be improved by also taking into account color correction.

Correcting the blur induced by diffraction, especially for distant objects, with deconvolution algorithms [VGP15] might also be a good way of improving image quality and extend the maximum imaging distance. The point spread function for such algorithms would be elongated perpendicularly to the prism direction, and might show non-uniformity along the prism array.

The resolution can also be improved by using super-resolution techniques [YSL+16].

### Size of the device

Our device is relatively large and it was suggested that a phone-sized device might be more useful for quick inspection of hidden areas. There is no limit to design a smaller wedge, so a smaller device might be investigated, for example with an A5 active area and a thickness of 10mm. In this case, the images should appear in real-time on a portable device like a smartphone. Of course, the downside of a small size is that the imaged area would be smaller as well, and this lead us to consider real-time stitching.

### Stitching

The image size is currently limited to the size of the device. Whether we use a small or large device, it is interesting to reconstruct the profile of a full wall or trench. Images corrected from distortion can be stitched together [S+07] as shown on our panorama attempt shown in Figure 5.15. This implies some software and hardware challenges such that ensuring that the illumination is uniform within the frame and coherent between several frames. Real-time stitching would bring even more benefit as the unreconstructed areas can be specifically addressed.

### Stereo camera

It is possible to put several cameras at the end of the wedge guide, so that different viewpoints of the same scene can be taken simultaneously. This can be used to extract depth information from parallax and generate depth maps of objects under inspection, as with any stereo camera. Our first tests are encouraging, however, some work still needs to be to done to fully evaluate the accuracy of such a method. A projector

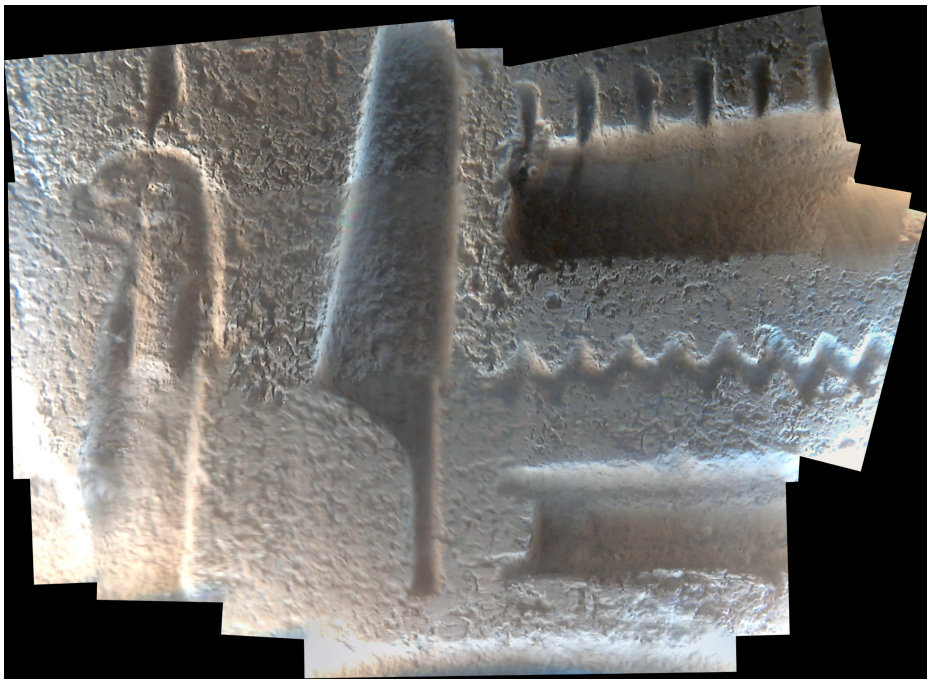


***Figure 5.15:*** *First attempt of image compositing for the reconstruction of a hieroglyph wall*

might be associated with one or several cameras to perform 3D reconstruction from structured illumination [Ger12].

### 3D reconstruction

Moreover, as images at different angles can be taken either simultaneously or by moving the device, we think it should be possible to reconstruct a 3D mesh from them by photogrammetry. This is not direct though: our device calibration is not standard, stable illumination of the scene is hard to achieve, and tilting the device in a trench can be hard so that only few angles might be accessible. To this end, using a better camera is mandatory, with a better resolution and probably a global shutter because it may be difficult to hold the device still. Changes in the core of photogrammetry algorithms may also be necessary, notably adapting the calibration process and changing the camera model to a "pusbroom-like" model.

Looking further into the future, it might be useful to combine the wedge light guide with a close range laser scanner. Indeed, the laser light could propagate in the wedge from a laser source to the scene back to a detector to evaluate depth information with time of flight techniques [GG07].

## 5.6 Conclusion

In this chapter, we presented a novel hardware optical device, a wedge camera associated with software processing that can be used in archaeology to get images where traditional cameras fail. We tested and discussed three concrete applications.

Before using our wedge camera, specialists in the digging of trial trenches were doubtful that our light guide would bring them much benefit because of the lack of pure visible information from the ground, and they were reluctant to lose the ability to touch and interact with the sides of a trial trench. However, after testing, they agreed that the wedge could help inspecting areas where traditional surveys do not fit. They also mentioned that it might be easier to get authorization for surveys when conducted with a wedge because it keeps disruption to a minimum.

The underwater application interested them because it means that one could observe trenches that are deeper than usual, even below the water line. Wedges might be used for the inspection of lakes and also in submarine archaeology.

There was more interest, even enthusiasm, in using the wedge camera to look into crevices. We have concentrated on the use of the wedge in Egyptology because of their need to inspect temples, but it can likely be used in many other applications such as in the archaeology of buildings, the exploration of caverns, among others, even beyond archaeology.

Finally, this work was the chance for us to experiment and understand deeply wedge guides. The properties described in Section 5.2 will be useful for the next chapter of this manuscript, particularly the shape of the beam and the question of light rays redirection.

# Chapter 6

# Autostereoscopic transparent display using a wedge light guide and a holographic optical element

We reviewed state-of-the-art 3D displays in Chapter 3 and saw that a wide variety of approaches are possible to design 3D displays for various applications. Among them, autostereoscopic displays are basically based on conventional displays associated with a view separation element and conventional multi-view rendering. LCD-based autostereoscopic displays may suffer from a lack of resolution because each pixel is addressed to a single view. They can be time-multiplexed to take advantage of the full resolution, but then some flicker is generally visible. Projection-based autostereoscopic displays can overcome this limitation by sending overlapping images to different locations without dividing the image resolution, but they can be bulky because of the projection distance. Wedge guides are interesting as they can compact the projection distance.

In this chapter, we introduce a new multi-projection based autostereoscopic display that relies on quite similar hardware and software issues as already explored in Chapters 4 and 5. As the spherical display introduced in Chapter 4, our projection-based display relies on LBS projectors, and they are coupled in a wedge light guide that we introduced in Chapter 5. Wedge guides are indeed well adapted for a multi-projection system, as several images can be projected from different directions within a confined space. Moreover, they have very good transparency. In this chapter, we investigate how wedge guides can be used to design a display that is both autostereoscopic and transparent, and we demonstrate a working prototype.

This work has been presented in a conference [CRGT19] and published in a journal [CRT+19]. A patent application is currently under revision.

## 6.1 Introduction

The utility of a transparent 3D display is evident: imagine a transparent window through which you can see both the real world and a 3D digital world that are superimposed. There is no need to wear any glasses or headset that requires preparation time, and reminds you that what you see is actually not real. The ability to superimpose stereoscopic images over a real-world scene without any wearable is an exciting perspective for augmented reality applications: such displays could be used as a showcase to augment products in shops, artefacts in museums, or even fishes in aquariums. It can also be an assistant in any task executed through a window such as driving, monitoring processes, and so on.

We describe in this chapter the design and implementation of a new, full-color, transparent autostereoscopic display prototype. More precisely, we review specifically related work on transparent 3D displays in Section 6.2. We then detail the general principle of our display in Section 6.3. The main idea is to use several LBS projectors with a wedge guide, and the view separation is done thanks to a custom Holographic Optical Element (HOE) that we detail in Section 6.4. We focus on describing the viewing zones of our display in Section 6.5. We then detail in Section 6.6 the software requirements in order to make a functional prototype with a good 3D experience. Section 6.7 presents the achieved results, and the perspectives of improvement are explained in Section 6.8, before we conclude in Section 6.9.

## 6.2 Previous transparent and 3D displays

We experience more and more transparent displays in our daily life and different technologies are used to create them. Among them, basic technologies for flat panel displays (see Section 2.2.2) can be adapted to transparent displays.

Indeed, LCD displays for example are intrinsically transparent: each panel described in Figure 2.11 is transparent itself, only the high power backlight is not. Instead of modulating the backlight, transparent LCD displays modulate ambient light, which has much less power, resulting in a low brightness display. The transparency of LCD is also quite low as polarizer layers limit the amount of unpolarized light passing through them to 50%. On the opposite, LED displays generate their own light, and so they do not require a backlight. They can be transparent simply by fixing the diodes on a transparent surface with a space between each. This is similar to the OLED technology, where the light is actively generated by using a transparent active TFT matrix. OLED panels currently exhibit about 40% transparency [PSK+18].

All of the described technologies have repeated micro structures, either TFT films, color filters or diodes, which leads to diffraction effects. This results in severe blurring


(a) (b) (c)

***Figure 6.1:*** *Transparent 2D displays: a) OLED digital signage b) Diffraction effect from a transparent OLED display (source: [THdJ+15]) c) A reflection-type transparent HUD in the Peugeot 3008 car (source: club-3008.com)*

of the see-through image. In particular, a diffraction pattern may be visible for direct light sources as shown in Figure 6.1(b). The pixel structure can be altered to reduce the diffraction effect [KLL+13, THdJ+15].

A transparent display can also be made with a regular display and a transparent combiner. For example, the Meta2 [Pul17] headset used a half-mirror to reflect the light from a LCD display. The same principle can be used for head up displays (HUD) in cars [MCGV18] as illustrated in Figure 6.1(c). The so-called "holographic" pyramid display, the DreamHoc™ [Rea09] (see Section 3.6.1) also uses semi-transparent combiners with regular 2D displays to create floating images. A combiner made of retroreflective material associated with a projector was also developped [RMDGR16] where the transparency was achieved either by drilling holes in the material or with a semi-transparent glass assembly.

As is, a 2D transparent display is generally not adapted for augmented reality applications because the digital content needs to be view-dependent in order to always align on the real objects to augment, as illustrated in Figure 6.2. The DreamHoc™product creates the illusion by delocalizing the screen at a virtual location (done naturally by the mirror reflection) and positioning objects to augment at about this plane position. Tracking the user's eyes positions relatively to the display and changing the image accordingly can correct alignment and perspective in real-time [LOIB13, PW16]. However, as long as only one image is generated, and as the user's left and right eyes have different positions themselves, a conflict between both eyes may arise depending on the distance of the virtual objects.

Head-mounted displays (HMDs) do not have the same alignment issues because the eye is fixed relatively to the screen. Instead of tracking the user's eye, the screen position and orientation are tracked to update the digital content in real time. Exisiting HMDs can either be monoscopic [SDB12, OMHC14] or stereoscopic [hol19, Pul17, BMA16].

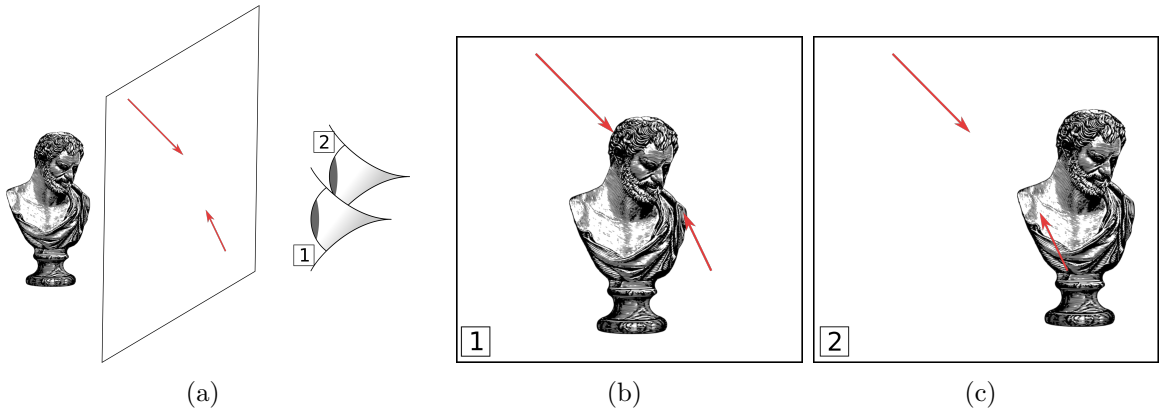



***Figure 6.2:*** *A 2D transparent display augmenting a real object (a) aligns well with the digital data for the correct position, or (b) generates a misalignment for the second eye and other viewing positions.*

A conceptually different approach is to simulate a see-through effect from a video stream and to augment real content on an opaque regular display, like a phone or tablet. This has been widely democratized by augmented reality applications for smart phones. These approaches only provide a 2D projection of both the real and digital worlds, however, they may provide software motion parallax. Mixed Reality headsets, as described in Section 3.2.2, are also based on this principle but are able to provide a 3D view of the real world through binocular depth cues and motion parallax.

Offering stereoscopic images to the user enhances 3D perception by exploiting binocular disparity. We already reviewed autostereoscopic displays in Section 3.3. Recall that this technology is based on generating several views and optically addressing them towards different viewing zones. Most of the existing approaches are based on parallax barriers or lenticular arrays [Rob03, All98], random holes [NF09], or compressive light-fields [WLHR12a]. If more than two views are presented, the display can also exhibit other depth cues like motion parallax [HR02], enhancing even more 3D perception.

The autostereoscopic displays presented above use conventional optics to separate the views, thus they are hardly adaptable to transparent displays. Takaki et al. [TY15] adapted lenticular-based displays to create a see-through flat panel display based on integral imaging with several layers of lenticular arrays and a transparent diffuser, but it suffers from diffraction effects introduces by the pixel pitch and the lenticular pitch, as well as low transmittance. The parallax barrier method with a transparent LCD was also developed [HLgP+16] but the same remarks apply. The random hole display was also adapted to a transparent display [KMCS12a], but the use of LCD panels

inevitably blurs the real objects by diffraction. Hong et al. [HPL+16] proposed the design of a reflective transparent anisotropic diffuser and a multi-projection system based on it. The diffuser consists of a nanometric metallic layer in an index-matching oil, so it is quite hard to manufacture. In addition, it has a low reflectivity (the authors have achieved 35%) so the resulting brightness might be quite low. Besides, their display includes free-space projection, and this is often unwanted due to possible occlusion and bulkiness. Lee et al. [LgPM+15] proposed a compact, non-transparent, multi-projection system using a wedge light guide, a regular vertical diffuser, and Fresnel optics to create multiple viewing zones (see Figure 6.3(c-d)). In this work, the light guide is transparent, however, the whole display cannot be directly adapted to a see-through display because the Fresnel optics and diffuser would distort the real world behind them.

HOEs are interesting to perform view separation while ensuring a good transparency as explained in Section 2.1.5. Notably, some waveguide-based displays use HOEs, as for example the Hololens™, where several HOEs are used to perform in-coupling, pupil expansion, and out-coupling in a waveguide [KC17]. The ability of HOEs to realize almost any optical function makes them suitable for multi-projection systems. Holografika [BFA+05] designed a holographic screen for a full parallax opaque display using a dense projector arrangement. Olwal et al. [OLG+05] designed the ASTOR display (see Figure 6.3(a-b)) with two projectors and a holographic combiner that scatters the light of each projector in separate viewing zones. Their display also involves free-space projection. As seen in Figure 6.3(b), it is bulky and not standalone: it cannot be moved unless the projectors and their support are moved accordingly. The fabrication of the involved HOEs is often eluded in the aforementioned literature, and our intention is to also clearly describe this process.

## 6.3 Overview

### General Principle

Having reviewed and compared existing techniques for both 3D and transparent displays, our goal is to propose a solution that overcomes some of the limitations mentioned in Section 6.2. More specifically, we want to design a 3D transparent display that does not require any glasses or headset, has a good transparency, a good brightness, produces a limited diffraction effect on the see-through images, is standalone and does not require free-space projection.

The principle of our solution, as depicted in Figure 6.4(a), is to first couple beams from multiple LBS picoprojectors into a transparent wedge guide. We denote by $N_{proj}$ the number of projectors. The output color image of each projector is then scattered

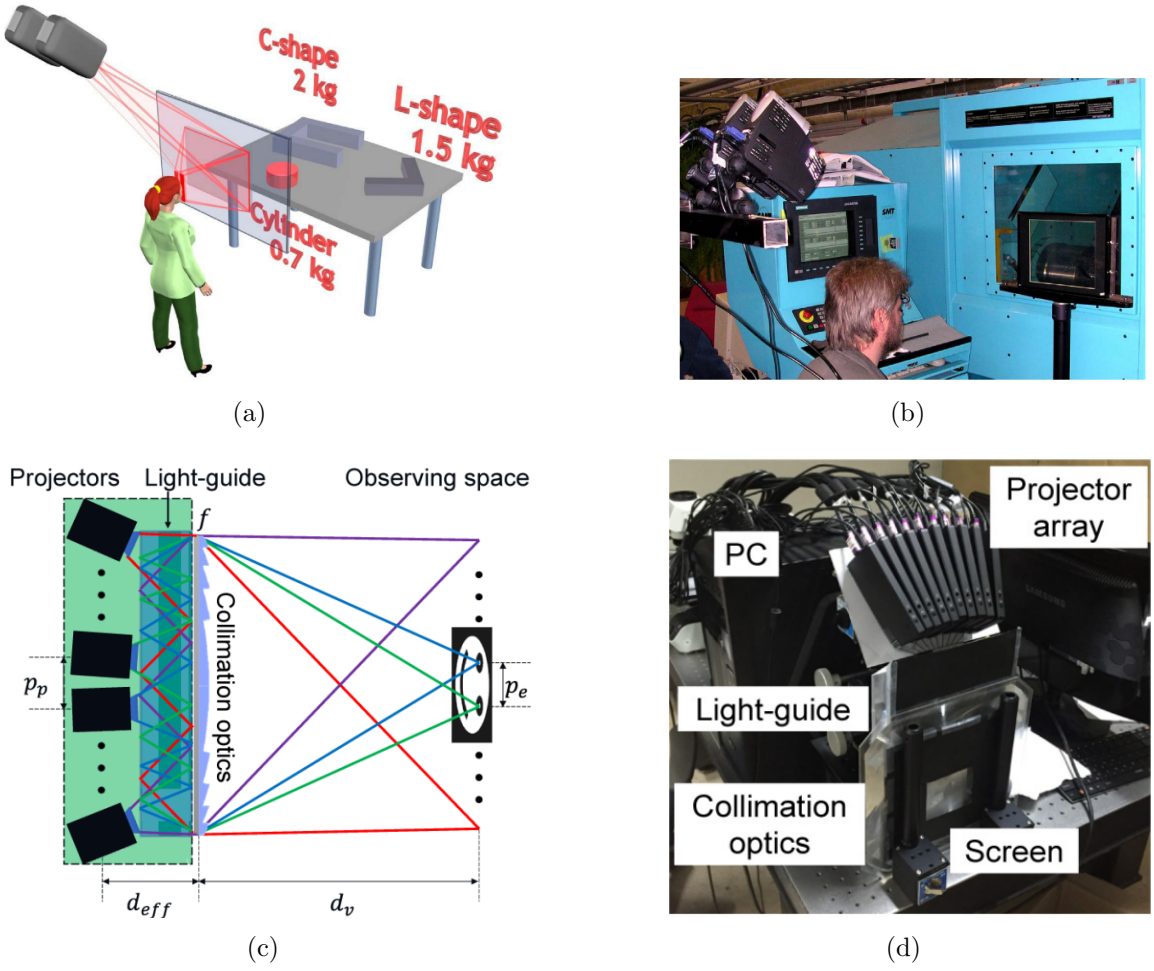
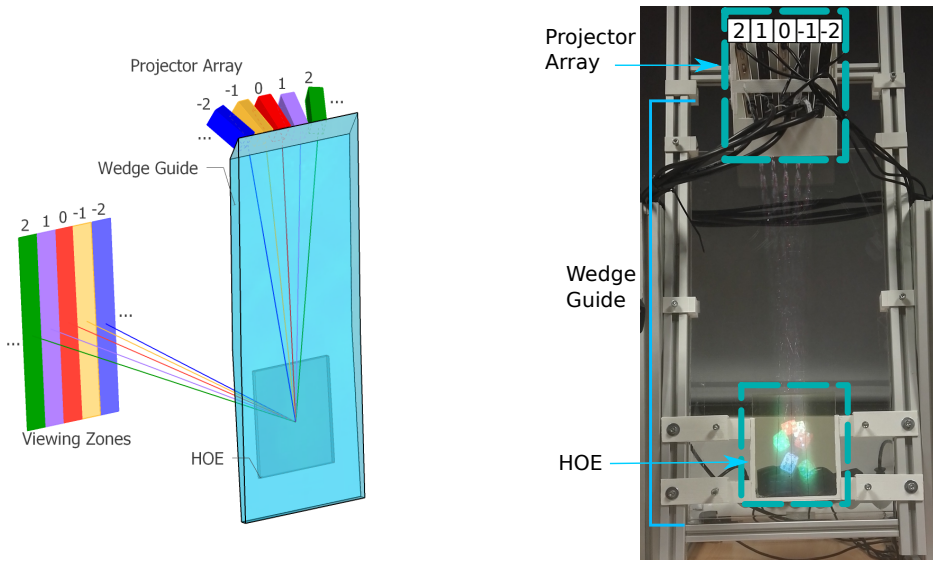



***Figure 6.3:*** *The most related works to ours: (a-b) The ASTOR display is based on two projectors and a reflection HOE that scatters light towards each eye. Note that the display is bulky and cannot be moved easily because of the independent parts. (images from [OLG+05]); (c-d) A multi-projection autostereoscopic display using a wedge light guide and collimation optics. Note that (c) is the top view of (d). It cannot be directly adapted to a see-through display because of the involved optics.(images from [LgPM+15])*




(a) (b)

***Figure 6.4:*** *a) Overview of the prototype: 5 projectors labeled -2..2 are coupled into a wedge light guide, and a HOE scatters the light towards independent viewing zones -2...2; b) Photograph of our prototype from the viewing zone labeled 0.*

toward separate viewing zones using a transparent HOE, making $N_{proj}$ images visible simultaneously at different locations. If the spacing between viewing zones corresponds to the interpupillary distance, then a stereoscopic image is perceived without glasses.

### Hardware setup

Consider Figure 6.4(b) for a picture of the hardware setup. We use an off-the-shelf wedge guide designed according to the patent [Tra03], with a size of 480x250mm and a thickness varying from 10mm to 1mm. $N_{proj} = 5$ LBS projectors are located at the entrance of the wedge guide. They are spaced by 17mm, and their angles are adjusted so that the guided light converges toward the HOE. We use the Sony MP-CL1A laser picoprojectors[1] whose operating wavelengths are $\lambda_R = 640\ nm$, $\lambda_G = 520\ nm$ and $\lambda_B = 450\ nm$.

The HOE is located at the imaging area of the wedge guide. One can observe that the HOE lets the light pass through it without deformation and that a digital image of a dice scene is superimposed with the real world. The HOE, whose size is almost

[1] `www.sony.com/fr-ma/electronics/projecteur/mp-cl1a` (Accessed on 09/18/2019)

equivalent to the imaging area, has a width $W_{hoe} = 10cm$ and height $H_{hoe} = 13cm$. The dice scene is visible because the camera is located in the viewing zone labeled 0 in Figure 6.4(a). Each viewing zone is located at about $D_{obs} = 50cm$ from the display, has a width of $W_{vz} = 3cm$ and a height of $H_{vz} = 10cm$. All the components hold together thanks to an aluminium frame and 3D printed components.

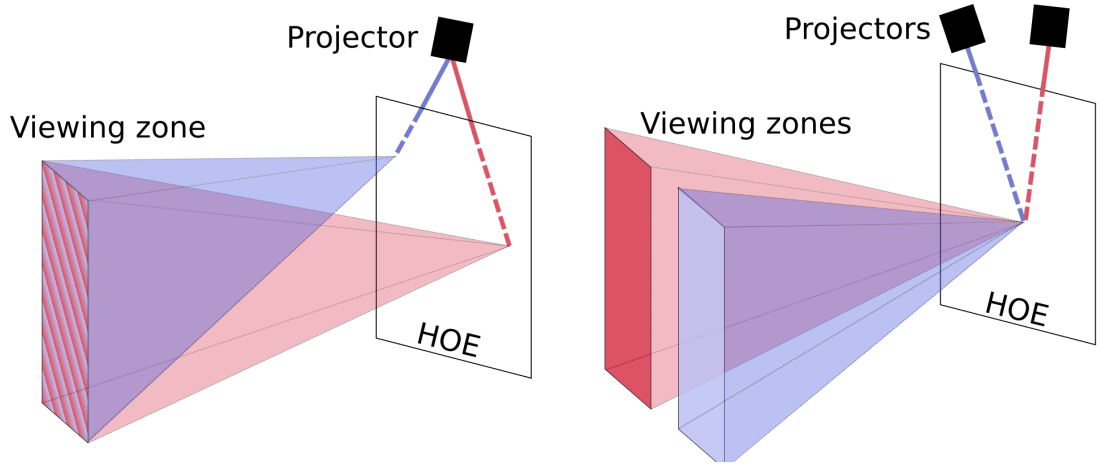
The projectors are connected to a laptop computer HP ZBook G3 with a 16Go RAM, a 2.60GHz processor and a NVidia Quadro M2000M graphic card. It only has two HDMI ports, so we added an external graphic card Matrox TripleHead2Go Digital SE[2] to handle the five projectors.

## 6.4 Holographic optical element

The HOE that we use is a transmission volume hologram for the following reasons. First, we decided to use a volume hologram instead of a thin one because we wanted it to exhibit angular selectivity as we explained in Section 2.1.5. This way, ambient light can pass through it without being affected, and the HOE can be transparent without any undesired diffraction effects. Second, we have chosen a transmission hologram over a reflection hologram because the latter would have required it to be located at the back face of the wedge guide, adding potential ghosting images as we observed with one of the prismatic film arrangement described in Section 5.3.2. Moreover, transmission volume holograms are known to be more selective in angles than transmission type [Dav13], and our first requirement is that ambient light should not be diffracted.

The HOE is clamped to the wedge exit surface but is not optically coupled with it, meaning that there is a tiny air interface between the HOE and the wedge. This allows the rays to propagate and leave the guide in the same manner as described in Section 5.2.2. As also explained in Section 5.2, a diffuser must be located at the wedge part so that an image projected through the guide can be visible, and in our display, the HOE acts as a diffuser.

In this section, we start by describing in Section 6.4.1 the desired optical function of the HOE and our approach to realize it. We then detail the recording setup in Section 6.4.3.




(a) (b)

***Figure 6.5:*** *Desired properties of the HOE. a) For a single projector, each pixel must be scattered toward a single viewing zone. b) For a single point of the HOE illuminated by several projectors, the two pixels must be scattered toward different areas.*

### 6.4.1 Role of the HOE

#### Scattering light of a single projector

The first function of the HOE is to scatter the light of a single projector toward a viewing zone. The projector scans a laser beam, representing a single pixel, over the entire surface of the HOE. This beam is incident on the HOE with different angles according to the position on the HOE. Compared to our previous work with the prisms (Section 5.3.2), each ray leaving the guide at the HOE must scatter its energy toward the full viewing zone and not only deflect the ray, as shown in Figure 6.5(a). This is done by recording a rectangular diffuser of the size and location of the viewing zone as we will detail in Section 6.4.3.

#### Scattering light of several projectors

Secondly, for each point on the HOE, different angles of incidence must result in different viewing zones as shown in Figure 6.5(b). To achieve this, we record the optical function on a material with an angular bandwidth that is sufficiently wide to enclose the difference of incidence of each projector. The typical efficiency

[2] `www.matrox.com/graphics/fr/products/gxm/th2go/digital_se/` (Accessed on 09/18/2019)

curve [Kog69] with respect to the deviation from the Bragg angle is shown in Figure 6.6. Note that the Bragg angle is a 3D angle that depends on the position of the HOE, but every point of the HOE is subject to this angular dependency with respect to its Bragg angle.

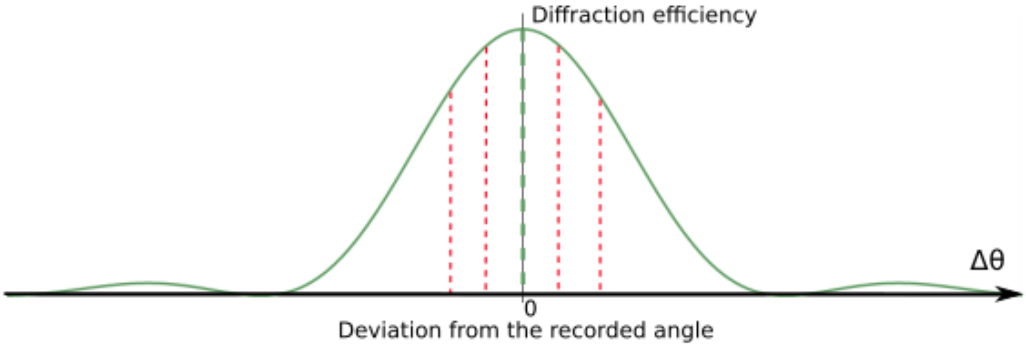



***Figure 6.6:*** *Typical diffraction efficiency of a volume hologram depending on the deviation from the Bragg angle and the incidence angles of the 5 projectors of our display. The idea of our work is to locate the central projector (green dashed line) at about the Bragg angle and the others (red dashed lines) within the angular bandwidth.*

Rays that are incident on the hologram at the Bragg angle $\theta_B$ reconstruct exactly the recorded viewing zone with the maximum efficiency. If the central projector is exactly positioned, then this is the case for all the rays coming from it. Rays that are incident on the hologram with an angle $\theta = \theta_B + \Delta\theta$ within the angular bandwidth reconstruct a viewing zone that is angularly shifted by $\Delta\theta$ and with a lower efficiency. This is the case for projectors $i$ with $i \neq 0$ as represented in red dashed lines. If the projectors are properly arranged, then the reconstructed viewing zones can be adjacent to each other.

At this point, it is interesting to recall that rays that are incident on the hologram with an angle outside the angular bandwidth are not diffracted because the efficiency has dropped to zero. This is the case with light rays that do not come from the inside of the wedge, and consequently the HOE appears transparent to ambient light.

### 6.4.2 Potential alternative approach

An attentive reader might have considered another solution based on angular multiplexing. Indeed, the most straightforward approach would be to angularly multiplex the hologram as many times as the desired number of projectors. This means that there would be an optical function for each projector, with a very narrow angular bandwidth so that they do not overlap each other, as illustrated in Figure 6.7.

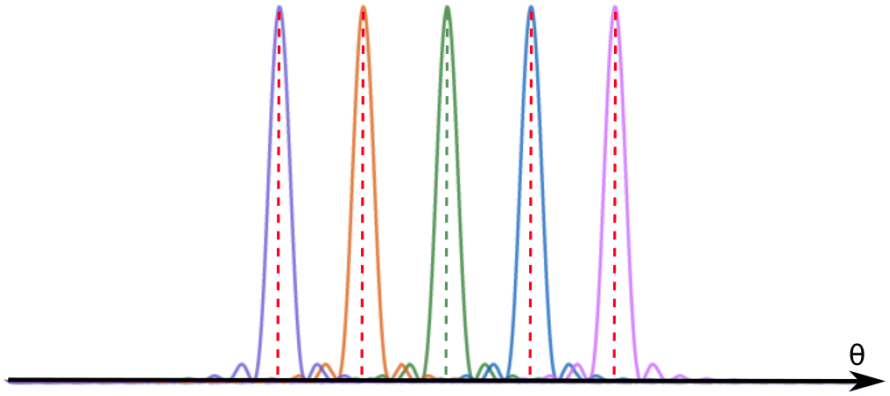



***Figure 6.7:*** *The alternative approach would be to angularly multiplex the hologram for each projector positions (in dashed lines).*

We have chosen not to follow this approach though, for several reasons. First, the recording setup described in Section 6.4.3 would have been much more complex because it would have consisted in lighting, either simultaneously or sequentially, the hologram from many angles. Besides, each grating would have a small angular bandwidth, so the required accuracy for both the recording and the reconstruction would have been a lot higher. Finally, the diffraction efficiency would have been decreased by a factor $1/N^2$ [PBW88] where $N$ is the number of multiplexed holograms.

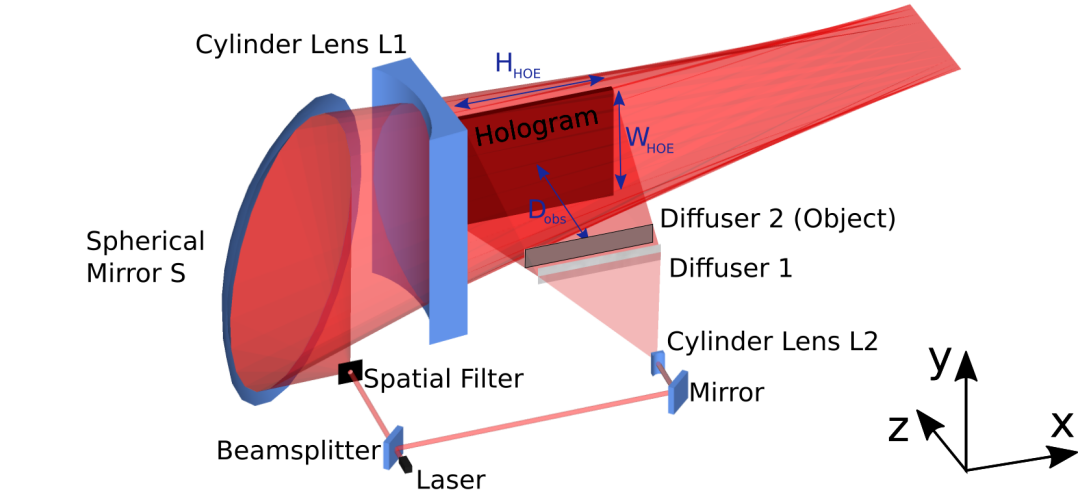
Instead, we show in this work that it is possible to record a hologram for a single projector and use the large bandwidth to reconstruct other viewing zones, without recording each projector position. Our approach also has the advantage of giving more freedom on the placement of the projectors after the recording process.

### 6.4.3 Recording setup

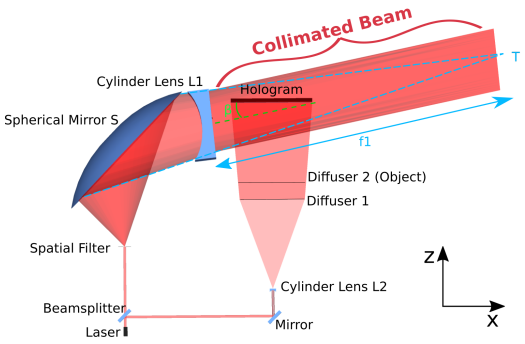
We fabricated the HOE on the material Ultimate 04 (U04) together with the Ultimate Holography company[3]. The HOE is optically recorded, this means that the material is exposed to interferences between a reference beam and an object beam as described in Section 2.1.5.

Let us consider the recording for the central projector position ($i = 0$). Figure 6.8 illustrates the recording setup, the used coordinate frames and notations. First, a laser beam is divided into a reference and object beam by a beamsplitter. The object that we want to recreate is a viewing zone, that is to say a rectangular window of dimensions $W_{vz} \times H_{vz}$ located at $D_{obs}$ in front of the hologram. The object beam is created by a diffuser D2 of these dimensions and position. To uniformly light the

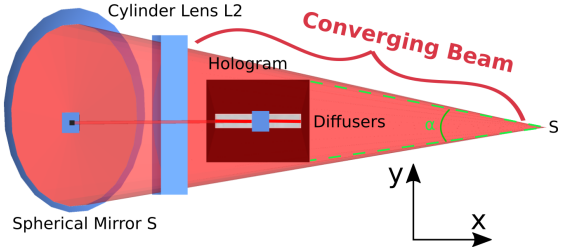
[3]www.ultimate-holography.com (accessed on 09/18/2019)




(a) Side view




(b) Top view




(c) Front view

***Figure 6.8:*** *Recording setup of the HOE. The reference beam is constructed with a spherical mirror to make the beam convergent and a cylinder lens to collimate it in one direction. The object beam is constructed by lighting a rectangular diffuser of the dimensions and position of the central viewing zone.*

diffuser, the laser beam is first turned into a line by a cylinder lens L1. The focal length of L2 does not matter since the aim here is simply to create a line on a first diffuser D1. This line is then scattered by D1 to uniformly light D2, which in turn lights the hologram.

This diffuser corresponds to the object that will be reconstructed by the reference beam. Defined this way, it is clear that the reconstruction has to result in a *real* image of the object where rays converge, rather than a virtual image. Thus, the **reference beam** must be the *complex conjugated* of the incident beam, i.e. the counterpropagating wave, as explained in Section 2.1.5, notably Equation 2.11 and Figure 2.8. Let us insist a bit on this point: if we recorded directly the incident beam, then, at the reconstruction process, a virtual rectangular diffuser would be visible at $D_{obs}$ behind the hologram to a user looking through it. This is not what we want: instead we want a real rectangle that is visible on a screen, so that when a user puts an eye in it, the scattered light from the projector is visible on the plane of the hologram.

The beam that is incident on the HOE is the beam that exits the wedge guide: it is collimated in the propagation direction of the guide and diverges in a parallel plane as explained in Section 5.2.2. As a result, the reference beam must be collimated and converging. Moreover, it must be incident on the hologram with a grazing angle corresponding to the exit angle of the wedge guide (about 8°).

The reference beam is created as follows. First, it passes through a spatial filter that spreads it uniformly toward a large concave off-axis spherical mirror S1 that converges the beam. Of course, an astigmatism can be noticed due to the off-axis configuration, this means that instead of focusing to a single point, the beam converges toward a sagittal image S, created by the largest dimension, and a tangential image T created by the smaller dimension (see Figures 6.8(b) and 6.8(c)). The hologram is located at 400mm from the sagittal point, so that the converging angle $\alpha$ (see Figure 6.8(c)) corresponds approximately to the measured average diverging angle from the wedge part at the expected HOE location ($\alpha \approx 15°$).

Since the objective is to collimate the beam in the XZ plane (see Figure 6.8(b)), the tangential image is located at the focal plane of a diverging cylinder lens L2. The collimated direction is incident with a grazing angle $\beta \approx 8°$ on the hologram (see Figure 6.8(b)). The collimation must be done before the beam reaches the HOE, so this imposes a minimal focal length of about -400mm. In addition, its size must be wide enough to cover the apparent size of the hologram in the converging beam. We were unable to find a lens with the required height, so we glued together 8 cylinder lenses (122-0234E from EskmaOptics[4]) of focal length $f_1 = -500mm$, width 25.4mm

[4]http://eksmaoptics.com/optical-components/lenses/plano-concave-cylindrical-lenses-122-0234e/

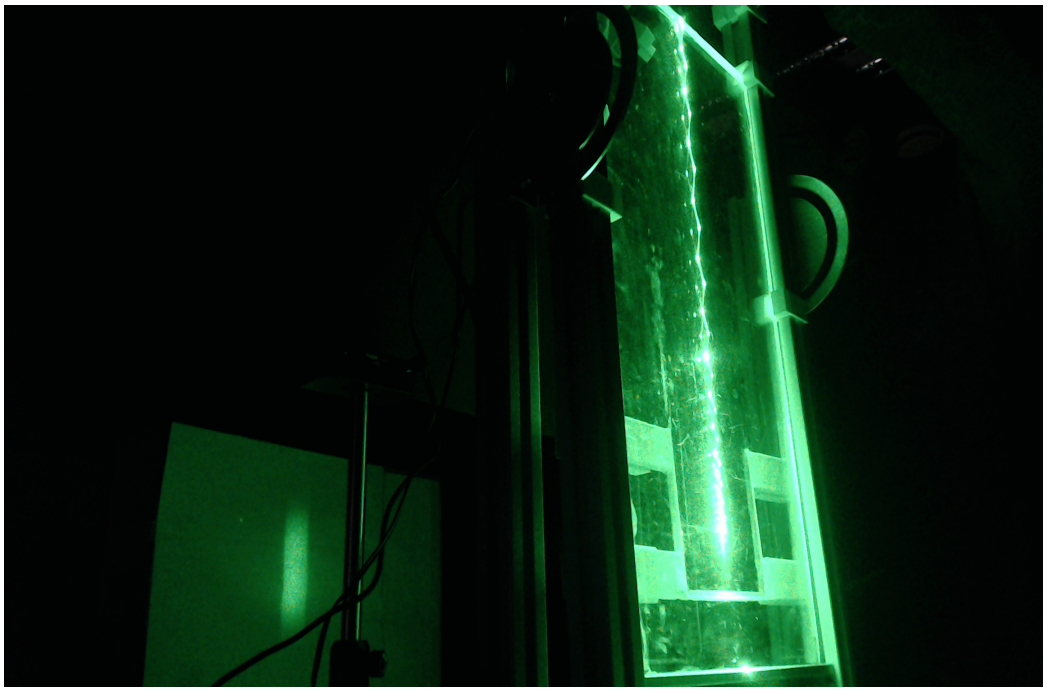


***Figure 6.9:*** *Reconstruction of a full viewing zone with a green laser*

and height 25.4mm, resulting in a 203.2mm high cylinder lens.

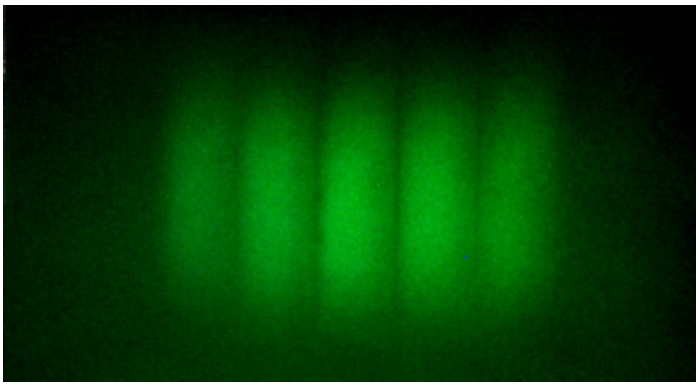
The first HOE that we recorded was monochromatic at 532nm. Then, a full-color HOE was recorded in a single shot by replacing the laser by a combination of 3 lasers at respectively 457nm, 532nm and 639nm. Note that the recording wavelengths do not exactly match the wavelengths of the projectors, but this did not cause any problem, apart from probably a small loss of efficiency.

## 6.5 Viewing zones

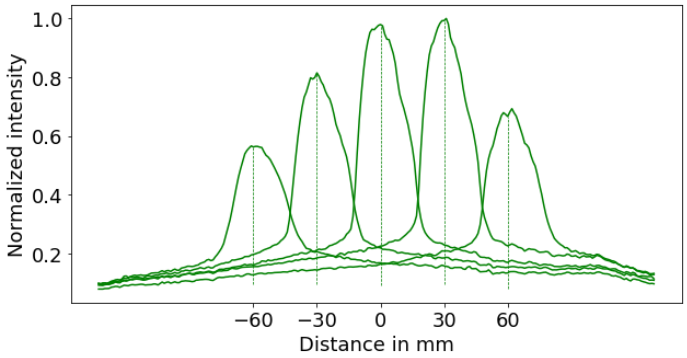
### 6.5.1 Reconstruction of the viewing zones

We first show the reconstruction of the viewing zones with the monochromatic HOE. Figure 6.9 shows a picture of a 30mW green laser beam that propagates in the whole system: from the entrance of the wedge guide to the resulting viewing zone created by the HOE. This would represent a single pixel of a LBS projector. A diffuse screen (a white painted wooden board) is located at 50cm in front of the HOE and allows the observation of the reconstructed viewing zones. It can be noticed that the full viewing zone is reconstructed, and indeed we insist on the fact that every point on the HOE spreads its energy toward the full viewing zone and not only a portion

of it. If the HOE is fully lit, then the intensity is simply higher because every point of the HOE lights the entire viewing zone.



(a)




(b)

***Figure 6.10:*** *(a) Photograph of the 5 reconstructed viewing zones with the projectors on a screen at distance $D_{obs}$ from the HOE. The image is corrected from perspective. (b) Normalized intensity profile along the horizontal direction*

Consider now the following configuration: the five projectors are located at the entrance of the wedge guide, and their spacing is chosen so that the viewing zones are adjacent to each other. They project a full screen green-only image on the monochromatic HOE, and the diffuse screen is located 50cm in front of the HOE. Figure 6.10(a) shows a picture of the reconstructed viewing zones on the screen, and Figure 6.10(b) shows a normalized intensity profile of the viewing zones. Note that we built the curve with photographs of each viewing zones independently. The photographs are taken from the same location, and with the same camera settings. We corrected them from keystone distortion (see Section 2.3.2) and the pixel values are corrected by the camera response function in order to get the relative intensity values. The horizontal

axis is built with respect to markers of known distances. We can observe that the viewing zones are well defined, and their sizes match the size of the recorded diffuser.

The relative intensity of each viewing zone can be seen in Figure 6.10(b). As predicted by the coupled-wave theory [Kog69], the intensity lowers from the central positions with a $sinc^2$ modulation when the angle of incidence diverges from the Bragg angle. We can see from this image that the optimal Bragg angle is not located exactly at the central projector position but between two views. We do this is to prevent the central viewing zone from being too bright with respect to the outermost projectors addressed to the second eye.

The outermost projectors have an efficiency ratio of 60% with respect to the central one. Hence, $N_{proj}$ can be more than 5 before extinction of the diffraction efficiency. We evaluate at $N_{proj,max} = 9$ the maximum number of projectors with a minimum intensity ratio of 20% with respect to the central one. Uniformity in intensity can be achieved computationally by lowering the intensity of the brightest viewing zones so that they match the darkest ones. Uniformity within viewing zones would require knowing the user's eyes exact positions, and this is part of future work (see Section 6.8).

The same observations can be made with the full-color hologram, but it exhibits some chromatic issues that are specifically discussed in the next section.

### 6.5.2 Color issues

Even though the monochromatic display behaves exactly as planned, we have faced unexpected results with the wavelength-multiplexed HOE.

Figure 6.11(a) is a picture of a viewing zone created on the $D_{obs}$ distant screen by a fullscreen white image on the projector. Figures 6.11(b), 6.11(c), and 6.11(d), respectively, show the contributions of the red, green and blue wavelengths at $D_{obs}$.

Figure 6.11(e) introduces notations of each observed area, and the contribution of the three wavelengths is explained below:

- The areas $R_R$, $G_G$ and $B_B$ are sharply imaged and superimposed at the screen position, as desired. An eye located in this area can thus perceive a full color image. The ideal white balance can be achieved either computationally or by optimizing exposure times for each wavelength at the recording step.
- The red wavelength does not reconstruct any other image.
- The areas $G_R$ and $B_G$ are overlapping and appear blurred at the screen location. They are actually images that are reconstructed 15cm further than $D_{obs}$ (at $D_{G_R,B_G} = 65cm$ from the HOE). These images are undesired crosstalk images that are respectively created by the green wavelength diffraction in the red grating, and the blue wavelength diffraction in the green grating.

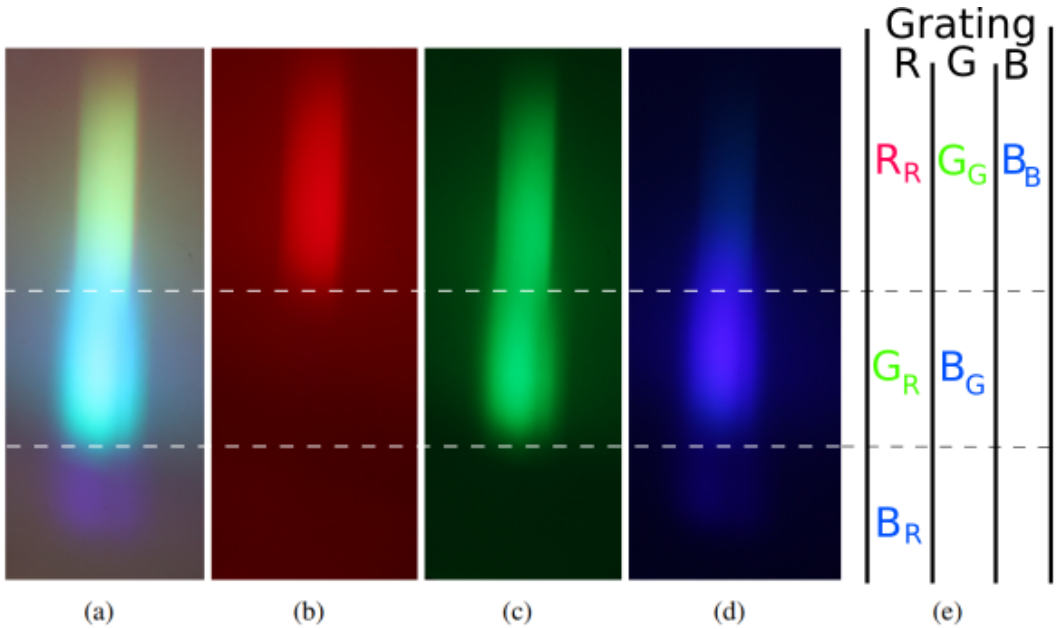



***Figure 6.11:*** *Photographs of a single viewing zone created with the wavelength-multiplexed hologram lit by (a) Red, green and blue wavelengths; (b) Red only; (c) Green only; (d) Blue only; (e) Wavelength-dependent contributions of each grating: $G_R$ represents the green wavelength diffraction in the red grating, and so on. $R_R$, $G_G$ and $B_B$ are the expected viewing zones while $G_R$, $B_G$ and $B_R$ are unwanted crosstalk images created by a wrong grating.*

- The image $B_R$ is located 85cm behind the diffuse screen ($D_{B_R} = 135cm$ in front of the HOE) and corresponds to the order created by the blue wavelength diffraction in the red grating.

At the time of the recording, we did not measure the wavelength dependency of the diffraction efficiency. We observe that the areas $G_R$ and $B_G$ are brighter than $G_G$ and $B_B$, respectively. This is why we believe that the red grating is more efficient than the green one, which is itself more efficient than the blue one. Moreover, the green and red grating have a very large bandwidth. This is unusual and this might come from the fact that we recorded the hologram at a grazing angle, and the recording material was never used this way before.

We measure the RGB area to be about $75mm$ high, and this is the effective viewing zone height where an eye can perceive a color image. In the current state of the prototype, if an eye is located too low in the viewing zone, the user would see an image that does not contain red. We discuss in Section 6.8.1 how to avoid such color effects.

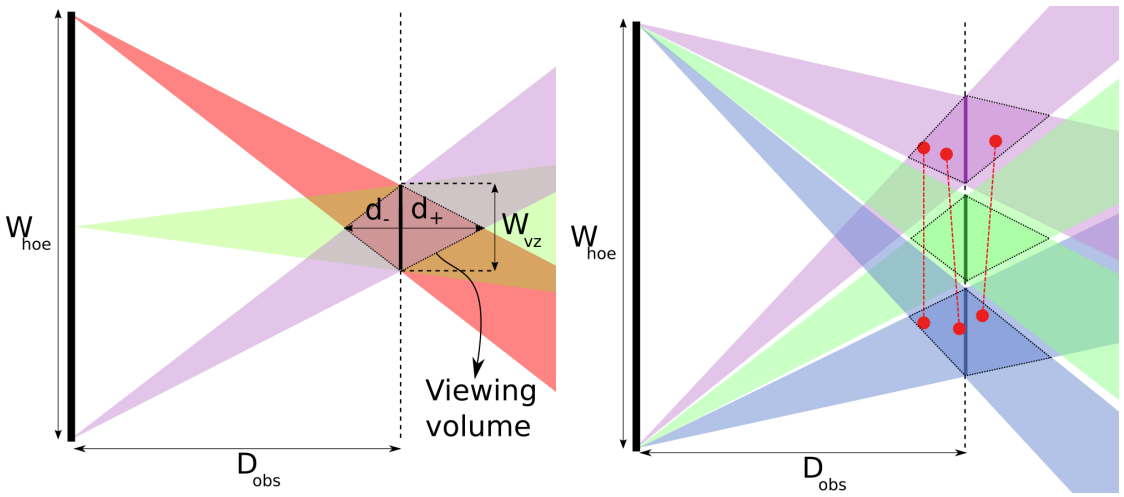



***Figure 6.12:*** *Schematic top view of the display showing the reconstructed viewing volume. (a) For a single position, the cones of emission of the center part and edges of the HOE, represented in different colors, define a viewing volume for a single eye; (b) Three projectors reconstruct different viewing volumes, and each eye has to be in the corresponding volume to see a correct stereoscopic image.*

### 6.5.3 Viewing volume

For now, we have described each viewing zone as a 2D plane that is located at a precise distance $D_{obs}$ from the display. In practice, our prototype does not strictly require the user to be located at this exact distance to be able to see the correct images. Indeed, a "viewing volume" is defined in space.

Figure 6.12(a) shows a schematic top view of our display showing one viewing zone. Three points on the HOE emit cones, represented in different colors, toward the viewing zone that has a width of $W_{vz}$ at a distance $D_{obs}$. In a plane perpendicular to the drawing, the cone is also elongated along $H_{hoe}$. The cones created by the edge points define an intersecting volume that has a diamond shape. We define $d_+$ and $d_-$ the maximum distances between the viewing zone and the corner of the viewing volume, as illustrated in the figure. Geometric considerations give the distances with respect the other parameters:

$$d_+ = \frac{W_{vz} D_{obs}}{W_{hoe} - W_{vz}} \tag{6.1}$$

$$d_- = \frac{W_{vz} D_{obs}}{W_{hoe} + W_{vz}} \tag{6.2}$$

$$d_{tot} = d_+ + d_- = \frac{2 D_{obs} W_{hoe} W_{vz}}{W_{hoe}^2 - W_{vz}^2} \tag{6.3}$$

With the numerical values, this gives a total tolerance of $d_{tot} = 33cm$ around $D_{obs}$, and this tolerance is greater with a smaller image size. This value is relevant for the center of the central viewing zone and decreases linearly as the eye moves laterally in the viewing zone. The extreme case happens when the eye is located at an edge of a viewing zone, then it has to be at exactly $D_{obs}$. This results in a more or less intense transition between the views. Note that if the eye is outside this viewing volume, then some pixels may be visible but not the full image. The visible pixels lie on a vertical line that gets narrower with an increasing distance.

Let us consider what happens with several projectors and two eyes. Figure 6.12(b) shows the cones emitted by the edge pixels toward three viewing zones, each represented with a different color, and each viewing volume is enclosed in dotted line. In red, several pairs of eye position are represented with red dots linked by a $d_{IPD} = 6cm$ long dashed line. Each eye is located inside the viewing volume of its intended viewing zone, so the drawn pairs represent correct locations where an autostereoscopic image can be observed. Moreover, if a single eye is located outside a viewing volume, then several images might be visible in separate vertical lines.

The concept of "viewing volume" cannot be strictly defined for both eyes because it depends on too much factors. But the idea that we want to convey is that there is a comfortable tolerance around $D_{obs}$. In practice, a correct autostereoscopic point of view can be found quite easily.

Considerations in the plane perpendicular to the drawings are not particularly relevant because $H_{vz}$ is greater than $W_{vz}$, hence it does not limit the viewing distance range.

## 6.6 Software

Our software is implemented in C++ and OpenGL. It consists of $N_{proj}$ full-screen windows that are sent to the different projectors, and a control window that is visible on the computer screen. GLSL shaders are used to perform parallel computations on the graphic card in order to achieve real-time results.

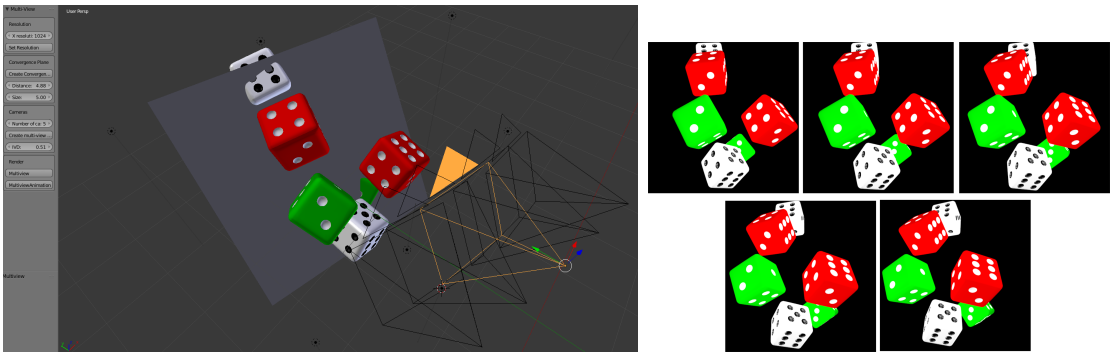


***Figure 6.13:*** *a) Screenshot of the Blender UI; b) 5 generated views of a dice scene showing parallax.*

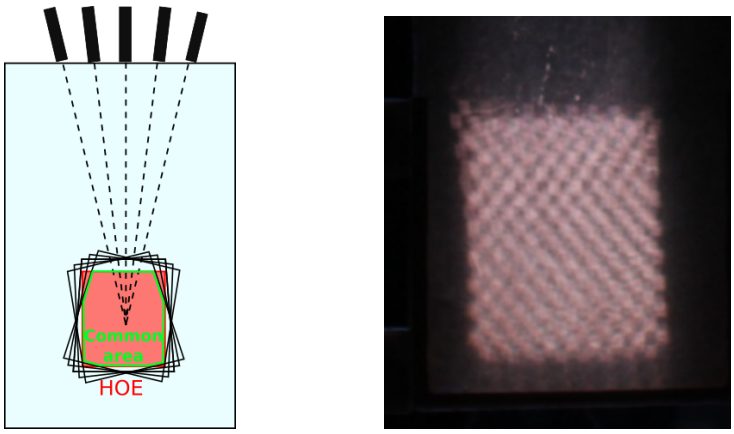
In this section, we explain the important aspects of the developed software. In particular, in Section 6.6.1, we focus on how the displayed images are generated. Then, we explain in Section 6.6.2 the required transformation that has to be applied on each image before being displayed and the calibration process to evaluate this transformation.

### 6.6.1 Rendering of the images

The input images are different views of a 3D scene showing different parallax. In a first implementation, the views are pre-rendered using Blender and stored into separate files. To this end, we have developed a Blender plugin in Python that helps to dispose virtual cameras in a 3D scene and that renders the different views. A screenshot of the user interface of our plugin can be seen in Figure 6.13(a).

The user sets up the position of a convergence plane that represents the imaging window (i.e. the HOE) and that is not rendered. All cameras are directed toward the center of this plane. Objects that are located between the plane and the cameras have positive parallax, that is to say they appear in front of the display. Reciprocally, objects behind the plane have negative parallax. Our plugin can export either a static multi-view rendering or an animation. For the animated rendering, all frames are stored in files and displayed with a controllable framerate on our display. Figure 6.13(b) shows an example of five renderings of a dice scene from five cameras, showing parallax.

We also implemented a real time rendering version, where a virtual scene with five cameras is loaded by the software and rendered directly at the execution time. This allows more user interactions such as exploring the 3D scene. This solution is based on the QML library of Qt that allows to quickly modify a 3D scene. For now,




(a) (b)

***Figure 6.14:*** *a) The projected images do not align directly on the HOE surface, and define a common area; b) Direct projection of 5 undistorted checkerboards, illustrating the misalignment.*

we render simple 3D scenes at an interactive rate and use a keyboard interaction. In the future, we will further enrich its functionalities.

In both of the above approaches, the generated images are loaded as OpenGL textures and cannot be directly displayed. Instead, they have to undergo an image transformation that depends on the hardware configuration and that is described in the following section.

### 6.6.2 Calibration and image correction

Each projector generates its own image through its own optical path. As shown in Figure 6.14, images overlap on a common area but they are obviously not directly aligned. We computationally pre-distort projected images so that they perfectly align on the HOE plane. This is done by correcting both the wedge optical distortion (pincushion and keystone) and the positional relative projectors' offsets.

In previous work by Lee et al. [LgPM+15], the optical path of all projected rays is modeled to compute the image transformation. Compared to their work where a wedge guide is considered as having planar faces, we use a guide with a slightly curved surface so the modeling is different. In addition, we do not have access to the exact geometry of the wedge guide neither to the exact positioning and angle of the projectors with respect to the entrance. Parameters of the model would be hard to set up so that they match exactly the hardware, and small errors would be visible. Instead, we designed a new calibration process that finds the required

image transformation without any a prior knowledge of the hardware configuration (see Figure 6.15) This means that the calibration process is still relevant if the HOE or projectors are moved, or if the wedge guide is changed.

The principle of our calibration is to evaluate a coordinate mapping between each projected OpenGL window and a common reference frame. We use this mapping to pre-distort every frame so that they overlap in the imaging area. The calibration has to be done only once, assuming that the prototype elements remain fixed between each other.

The calibration procedure is illustrated in Figure 6.15 and described below:

1. We replace the HOE with a regular diffuser (e.g. a piece of paper), so that the images projected by each projector are all visible from a single viewpoint.
2. A geometrically-corrected camera images the diffuser, and a rectangular subset of the common area is defined by the user by cropping the camera frame. The cropped area is designated as the calibrated area and defines a reference frame, in which a 2D reference grid is drawn.
3. For each projector and for each grid point, we project a point in projector space that is automatically moved until it reaches the target grid point in the reference frame. The general principle of this algorithm is explained later in this section.
4. For each projector, a list of calibrated points that are imaged on the reference grid is stored. At this point, the five projected grids align on the reference grid.
5. We then run a Python script that extrapolates the calibrated points to any other projector point: for each pixel, we perform a bilinear interpolation in a quad formed by the closest calibrated points. The script generates an image of the resolution of the projector that stores, for each pixel of the OpenGL window $(x_{proj}, y_{proj})$, the corresponding coordinates $(x_{ref}, y_{ref})$ in the reference frame in two color channels (green and blue). The pixel is set to black if it is not part of the calibrated area.

The algorithm that we use in step 4 for moving a point toward the reference corner is detailed in the following. We start by a point in the center of the image that is supposed to be visible in the reference frame. We detect it in the reference frame and compute the error with the target point. This detection is done on the camera frame with a regular HSV color segmentation with the OpenCV library. Then, we select randomly a direction X or Y and move the points in this direction with a certain step (which can be negative). If the error decreases, then the step remains the same. If the error increases, then the step is reversed and halved.

It is a Las Vegas type algorithm [LSZ93] and as such, the solution is certain to be found, however it may take an arbitrarily long time because the directions of displacement are taken randomly. To increase the speed, we track simultaneously three

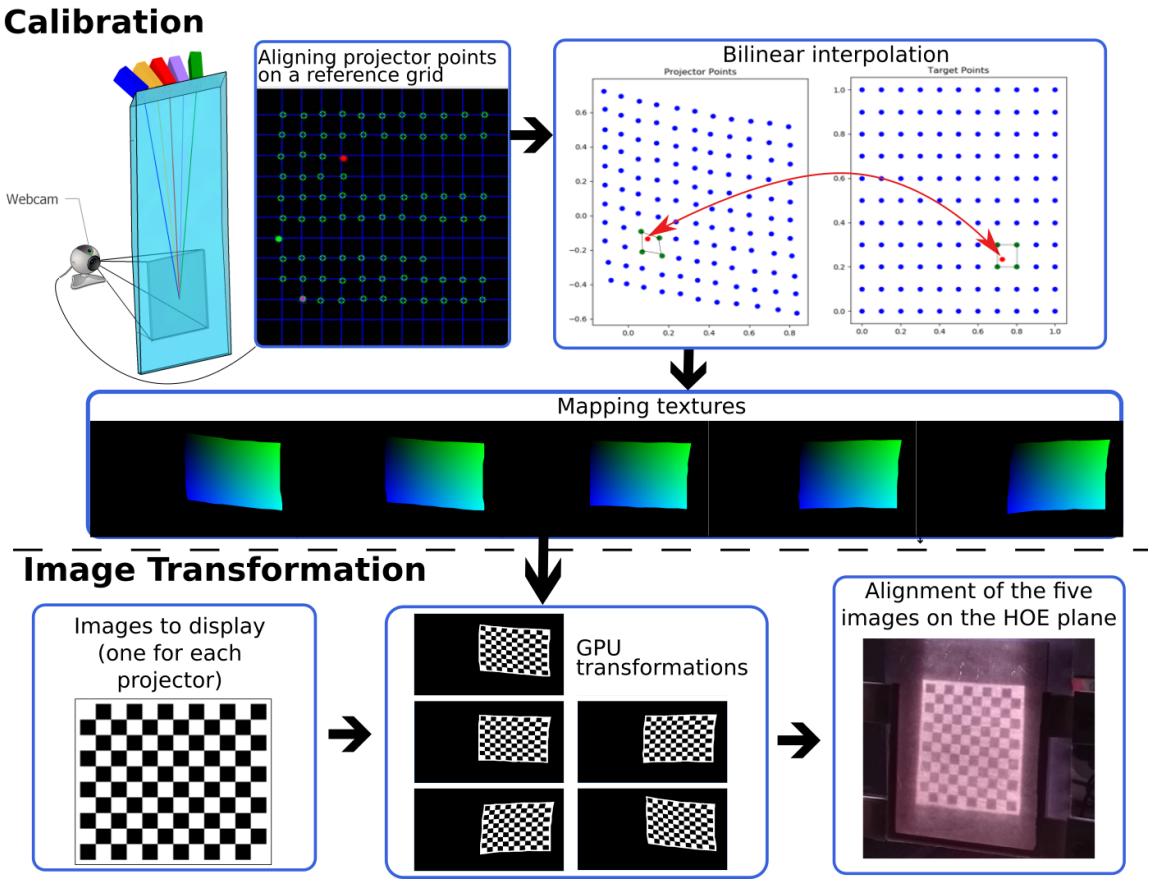



***Figure 6.15:*** *Calibration procedure and image transformation of five checkerboard patterns that align on the imaging area*

points by their color in each camera frame, as shown in Figure 6.15. In addition, we use a priori knowledge for the steps values: the prior displacement values in projector space are used for the next ones as initial steps, then the search is focused around the presupposed position. The ambient light of the room can have an influence on the detected colors so, to increase robustness, we display the three different tracking colors in full screen at the beginning of each projector's loop to retrieve the corresponding HSV range. A 11x11 points grid for the reference frame turned out to be enough to generate good results, and two minutes per projector are required on average to get the list of calibrated points.

At the end of the calibration process, $N_{proj}$ mapping textures are generated. The five images to display are sent over a fragment shader along with the mapping texture of the target projector. The shader reads the green and blue values of the mapping texture at the current pixel position. This gives the corresponding coordinate in the reference frame, which is directly the coordinate of the input texture. Figure 6.15 shows how a checkerboard texture is distorted before being sent to each projector, and how the five projected images are superimposed on the calibrated area. Of course, when running the display, we replace the checkerboards with the five different generated textures as described in Section 6.6.1.

## 6.7 Results of the prototype

Most of results were presented in the relevant sections of in this chapter, notably with the performance of the HOE and the software. In this section, we evaluate the full-color display as a whole.

### 6.7.1 Capabilities and limits

**Image quality** Figure 6.16 shows five photographs of the displayed images in each viewing zone. Recall that the images are rendered with a slightly different perspective that can be observed in the figure, and corrected with the procedure described in Section 6.6.1. Tiny vertical black lines are noticeable, and they are simply due to our large cylinder lens built from gluing 8 smaller lenses as explained in Section 6.4.3. Indeed, the sides of the lenses were not polished and the interfaces have cast shadows on the hologram. The images might appear a bit hazy, and this is mainly due to the scattering of the light guide interfaces. The quality of the wedge guide was a bit degraded over time because such a large optical surface is hard to maintain crystal clear and some scratches were made during all the experimentations. We believe that the above artifacts can be avoided in a more industrial manufacturing process.

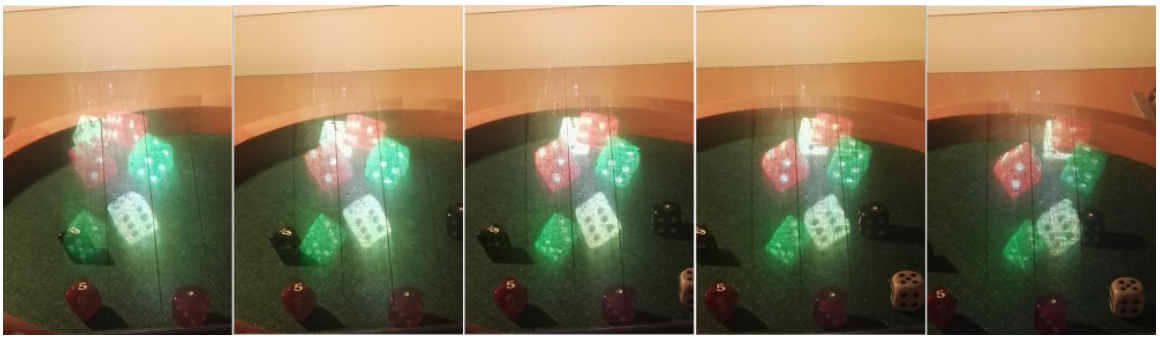


***Figure 6.16:*** *Five photographs of a rendered 3D scene taken from each viewing zone, showing transparency and the multi-view rendering. Note for example how the farther white dice is hidden behind the red dice on the right image, and visible in the left one.*

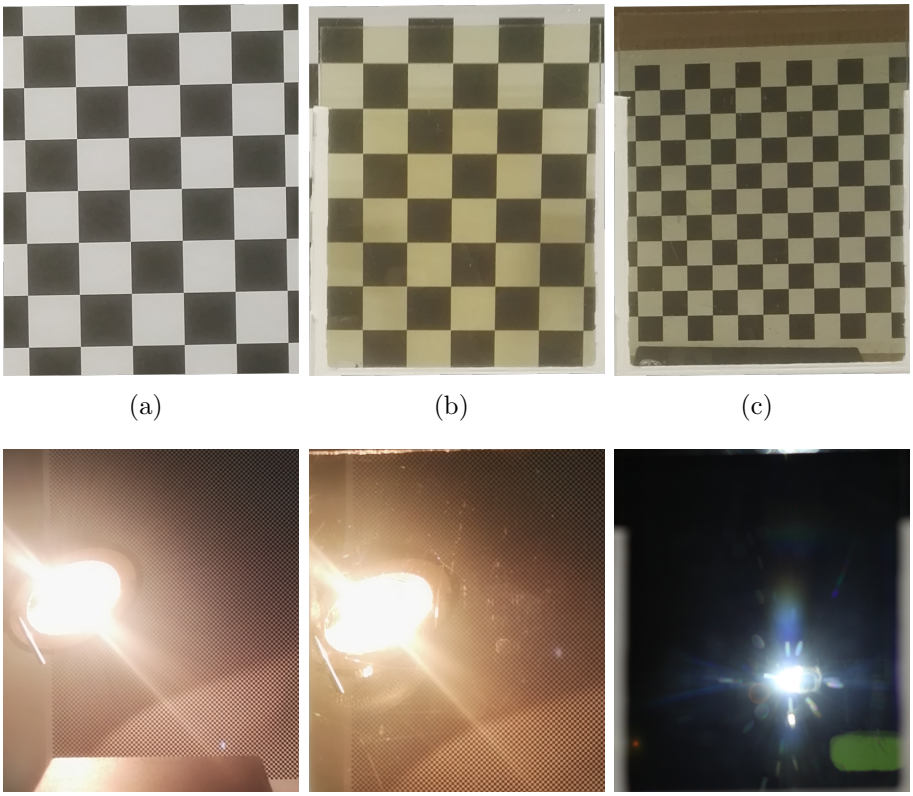
**Transparency** The hologram gives a "yellowish" hue to what is behind, as can be observed in Figure 6.17(b), (c) and (e). This is inherent to the HOE material, and more research on this point could improve transparency. In particular, the development process of the hologram includes a bleaching step that can be optimized.

We measured a transmission at normal incidence through the guide and hologram of about 85% with a 532nm green laser, being higher than for transparent OLED displays (≈40%). We do not observe any blurring of the image due to diffraction, but direct strong light sources are still diffracted in some extent by the gratings (Figure 6.17(f)). Compared to other transparent displays, we think that our solution improves the quality of transparency but further comparative studies are required to truly quantify the benefits.

**User position** We observe that a tolerance of about $\pm 10cm$ around $D_{obs}$ is acceptable to view the entire images comfortably, and the measured horizontal field of view is 17°.

**Depth cues** In terms of depth cues, our displays exhibit all of the monocular depth cues of a 2D display, as well as binocular disparity, convergence, and motion parallax, resulting in a convincing 3D effect. However, the vergence-accomodation conflict is not solved. On the contrary, we can almost consider that transparency makes it worse because the eye cannot focus on a real object located too far behind the display and its augmentation at the same time. This issue and perspectives to solve it are discussed in Section 6.8.

Of course, it is hard to describe the effect created by our 3D display through description, images or even videos. The concrete experimentation of the display is



(a) (b) (c)

(d) (e) (f)

***Figure 6.17:*** *Transparency of our display. (a-c) Photography of a checkerboard just before the display (a), just behind the display (b), and 50cm behind the display (c). (d-e) Photograph of a light source without the display (d) and with the display; (e) Strong and direct light sources still get diffracted*

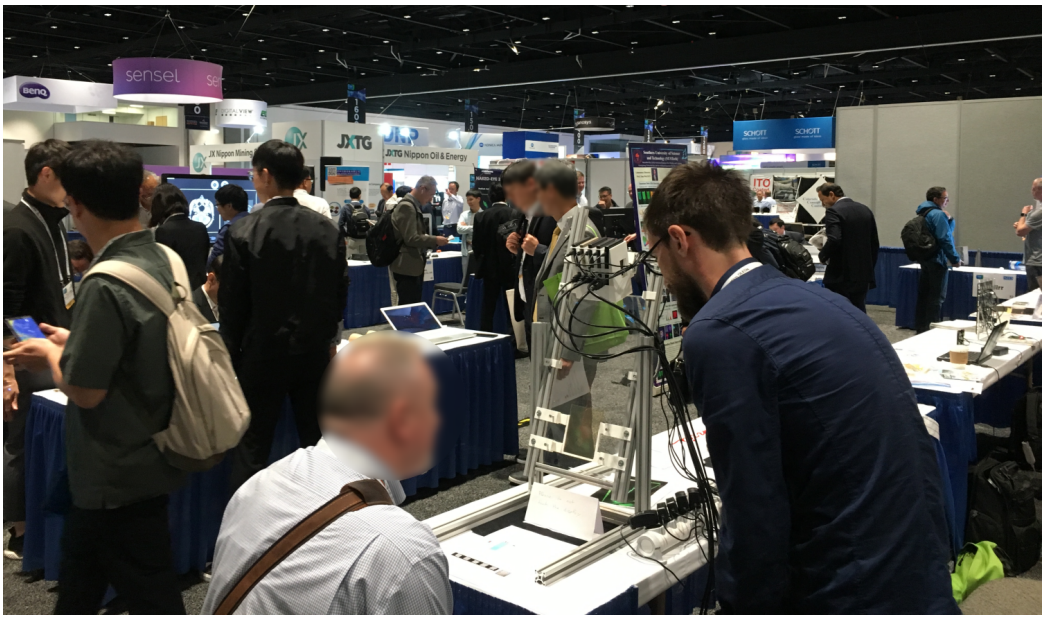



***Figure 6.18:*** *Public demonstration during three days at the I-Zone of the 2019 SID Display Week event, in San Jose, California*

necessary to appreciate its full potential as we will also see in the following.

### 6.7.2 User feedback

Our prototype was selected for a three-day demonstration at the I-Zone (Innovation Zone) of the exhibition hall of the 2019 SID Display Week event, in San Jose, California (Figure 6.18). This was the chance to present our solution directly to the display community. Most of visitors were amazed by the display, because nothing could be seen from the hallway until they sit down and the 3D color image appears suddenly. The stereoscopy gives a great effect, and we also noticed that motion parallax was also really important. Indeed, people enjoyed moving their head from left to right to appreciate the depth of the scene, and some of them were really surprised to learn that no tracking was used.

We approximately indicated what was the optimal observation distance (with a free-hands estimation) and observed that most of the visitors found the stereoscopic points of view pretty easily, although the audience was composed of people from different backgrounds, geographical areas, ages and so on. We also had to specify the color issues to some people, because, depending on their height, they were naturally located in the $B_G$, $G_R$ area described in Section 6.5.2. We asked them to move up their head so that they could appreciate the full colors.

## 6.8 Future work

### 6.8.1 Improvements on the HOE

**Color correction** The most important improvement of the HOE is the correction of chromatic aberrations described in Section 6.5.2, and several approaches to solve them are possible. We could use the same material and design a solution based on the observations. For example, recording another red viewing zone that overlaps with $G_R$ and $B_G$ would allow full color images in this area. This would also have a pleasant side effect of extending the height on the viewing zone.

Another way of improving color rendering would be to increase the thickness of the hologram to reduce the spectral bandwidth and avoid color crosstalk. The problem with this approach is that the angular bandwidth would be reduced as well, and the maximum angle of incidence separating the outermost projectors would be lower. With a constant projector pitch, this implies that $N_{proj,max}$ would be lower, and so the field of view would also be smaller. With this approach, it might be required to also multiplex the hologram angularly to extend the field of view.

**Image location** Depending on the application, it could be interesting to study how the image plane could be relocated behind the HOE plane, so that it is closer to the real objects and thus reduces the accomodation-vergence conflict for augmented reality applications. A possible solution might be to put an imaging lens between the diffuser and the hologram in the recording setup. In further research, we might study as well the recording of multiple imaging planes.

**Size of the HOE** For now, the HOE has a size of 10x13cm, but for some applications it might be interesting to record a larger HOE. The process could be a bit more complicated because of the larger optics that would be required to create the converging beam. On the contrary, a smaller HOE, for mobile applications for example, can be obtained by slightly adjusting the same recording setup.

**Setup parameters** Depending on the targeted application, the other parameters of the recording setup can also be changed: $D_{obs}$ can be changed at will, and $W_{vz}$ can be decreased to produce a denser horizontal parallax with a denser distribution of image sources.

### 6.8.2 Improvements on the display

In the following, we show several perspectives of improvement of the other parts of the display that should be further investigated.

**Eye tracking** First of all, we can consider how eye tracking would improve our display. For now, the perceived image does not change with head movement as long as the eye stays within the same viewing zone. It is possible to update the virtual camera position with respect to the real eye position of the observer. This would improve the resolution of the horizontal parallax while still providing different images to the left and right eyes, and also add vertical parallax. Moreover, the tracking can also be used to reduce the transition effects between views by adjusting the intensity of the viewing zones and within the viewing zones. Color aberrations can also be compensated by software with an eye-tracking approach.

**Field of view** Adding more projectors up to $N_{proj,max}$ would improve the FOV from 17° to 30° and offer more viewing zones. However, having more projectors might not be a practical solution due to the increasing cost and the consumed power.

**Optomechanical system** Instead of having one projector per viewing zone, we would like to study whether it is possible to reduce the number of projectors by an optomechanical system. The drastic example would be to have a single projector at the entrance of the wedge guide and mechanically move it laterally at high frequency, while synchronizing the rendering with its position. Systems including a few projectors and scanning mirrors might be a more practical solution. We can also consider coupling the projected light in a vibrating optical fiber as it is done in [SJMS12]. However, higher frequency projectors might be required to generate multiple viewpoints over time without flickering.

**Denser horizontal parallax** In the spirit of adding viewpoints, the views can also be made denser. Smaller viewing zones can be recorded, and then a denser distribution of projectors would result in a better horizontal angular resolution. Our display can even be adapted to a super multiview display (see Section 3.3.5). The idea would be to have viewing zones that are smaller than the eye pupil (let us say $W_{vz} = 1\ mm$) and either have a high density of projectors or move the projector at high frequency (e.g. by vibrating it) in order to generate several perspectives per eye.

**Specially shaped wedge guide** Finally, note that we used an off-the-shelf wedge guide that was not originally designed for this application, and it might be interesting to manufacture a dedicated wedge guide that is optimized for several projector positions. In particular, having a curved entrance instead of a rectangular one should improve the reconstruction quality by ensuring that the distance between the light source and the hologram is constant. Changing the size of the wedge guide with respect to the intended application and the size of the HOE should also be investigated.

In particular, there is no physical limitation to design a phone-sized display.

## 6.9 Conclusion

We have demonstrated the feasibility of a compact, autostereoscopic transparent display with multiple projectors and a custom HOE. The HOE is wavelength-multiplexed for full-color efficiency, but not angularly multiplexed. The multiple viewing zones are reconstructed with several projector positions due to the high angular bandwidth of the volume hologram. Although some chromatic issues are visible, we have explained their origin and provided perspectives to correct them. Hence, the proof-of-concept is demonstrated.

Our current prototype has $N_{proj} = 5$ views but it is theoretically able to generate up to $N_{proj,max} = 9$ views. The views are located 50cm in front of the display, they are 3cm wide and 10cm high. These values are fixed once the HOE is recorded; they result from our choices and can be changed in the recording step. Having narrower viewing zones would provide a denser horizontal parallax but would also require much more projectors. An optomechanical system might allow us to have more viewing zones than projectors. Eye tracking can improve horizontal motion parallax and add vertical motion parallax.

Such display has great potential for augmented reality applications such as augmented exhibitions in museums or shops, head-up displays for vehicles or aeronautics, and industrial maintenance, among others.

# Chapter 7

# Conclusion

Displays have imposed their use in every part of our society, and thinking outside the scope of the traditional display technologies is essential to make society evolve. By changing our relationship with the digital data, new applications and user experiences are made possible. In this thesis, we notably introduced the major challenges and technologies related to 3D displays. The ultimate 3D display replicating all aspects of the real world is not yet to arrive, but it is important that we push toward this ideal and explore all the possibilities.

We clearly described the realization of three different projects, each one resulting in a working prototype, and the pathway that lead to each contribution, and we summarize them in Section 7.1. We introduce the future work of each project in Section 7.2, and we describe general perspectives of research on 3D displays in Section 7.3.

## 7.1 Summary of the contributions

Spherical displays offer an unconventional way of representing virtual data. Even though commercial versions exist, they remain expensive and restrained to specific uses: chances are high that our readers never experienced one. This is unfortunate, because they offer very interesting properties that release some creative and perceptual constraints of conventional displays. After reading this manuscript, we hope that any curious person interested in exploring the possibilities offered by such displays is now free to do so. We have described a method to design a spherical display from scratch, made out of off-the-shelf and low-cost optical elements. The principle is to use a focus-free laser projector in combination with an optical system to cover the full sphere. At the same time, the optical system allows to image the spherical projection surface, making optical finger tracking possible and thus providing a direct way of

interaction. Of course, software considerations are mandatory for the prototype to be useful, and they are also described. The projected images undergo a GPU transformation to be adapted to this completely different display. The optical finger tracking is done through computer vision algorithms, and the projected images are modified according to the detected touch points. We mainly focused on applications for geographical or spherical data, but many others can be derived. Recall that the display part has the size of a ball lamp, so we could imagine it being a personal assistant somewhere on your desk.

Using wedge light guides also allows the creation of innovative displays, and we clearly explained their principle and applications. We first exploited their properties for creating an imaging device, together with a camera and custom software. In certain configurations, an object of interest can be in a constrained location so that a classical camera cannot reach a sufficient imaging distance in front of it. Achieving such imaging distance might be impracticable if not impossible in some cases: try to imagine a simple strategy to inspect below your fridge. Through the implementation of our prototype, we propose a solution where the imaging distance is folded up into a wedge guide, allowing the guide to be simply slid in front of the object of interest. Compared to previous work, we are able to image objects directly on the surface but also in front of it. We explained how the rays behave and showed that they needed redirection to provide normal viewing conditions, and we do this with a prismatic film in a new configuration. The wedge guide introduces optical distortions that are corrected by software. We illustrate the utility of our prototype in the particular field of archaeology, in which constraints are omnipresent and experts want to document as much as possible by causing minimal disturbance. We validated and tested various on-site configurations: for imaging underneath the soil through a slim trench, flooded constrained areas and through building crevices. We outlined that using wedge guides can solve issues linked to the imaging distance of cameras. Considering light travelling the other way, we then naturally got interested in their use with projectors for display applications.

We finally proposed a hardware and software implementation of a transparent autostereoscopic display. Metaphorically, such a display is a regular window that allows a direct perception of the real world as well as 3D digital content at the same time. Devices that achieve the same goal at the moment require a user to wear a heavy headset. Current autostereoscopic displays generally require conventional optics to separate the views and for this reason, they cannot be directly applied for a transparent display. Our idea to overcome this emerged directly from the results of our previous projects and the expertise that we acquired: it combines laser projectors

and a transparent wedge guide in order to have a compact projection device capable of superimposing several images on the same plane. In addition, we recorded a volume HOE to separate each image into independent viewing zones while ensuring the transparency of the whole system. We specifically described the role of the HOE and our method to record it. As for the spherical display, a software part is mandatory for both the calibration of the display and the rendering of the images. As a result, the combination of both the hardware and software provides a good depth illusion through binocular disparity, convergence and motion parallax. Our implementation has some issues but we proposed directions to overcome them, and we think that our concept can eventually lead to a commercial product. Specifically, our display is able to show 3D content over the real world without any headset, making it a very useful tool for augmented reality.

## 7.2 Future work

In these three years of research, we have explored many paths. Through three distinct but complementary projects, we have developed new concepts and realized working prototypes out of them. Each of these projects has their own perspectives of improvements and opens the way to future research as we detail in this section.

**Multitouch spherical display** The spherical display prototype can be improved in several ways. First, the resulting brightness should be increased in order to allow its use in a bright environment. For this, we should investigate other projection surfaces and identify the best one. Taking advantage of the whole luminous flux of the projector with an anamorphic system might also help in this direction. In parallel, we will continue our research regarding the implementation of the FTIR technique to improve the robustness of the tracking and ensure that it is still compatible in brighter environments.

This document as well as the associated publications will guide people in the replication. If a sufficient number of people get involved, then a community of passionates might emerge, sharing open source software and applications between each other. We can even think of a dedicated Operating System, from which many applications could be accessed.

We have mainly focused on the development of the prototype, and it would be interesting to further investigate user-oriented research. Several user studies are already available in the literature, notably to investigate the benefits of spherical displays for user interaction [PSS14, BKV12, BW09, BWB08], and our prototype can serve as a basis to contribute to the user interaction and perception literature on this subject.

Thinking outside the scope of spherical displays, most aspects of our method (e.g.

laser projection, optical sensing) can also be can be applied to other types of surfaces. We will also investigate free-form multitouch displays based on this experience.

**Wedge camera** Our wedge camera prototype suffers from several drawbacks. First, we found that our illumination system was not suitable for our cases studies, but changing the light source while keeping the same design can solve this. Second, we will investigate further the pushbroom camera model in order to provide a better image correction, particularly for distant objects that appear squeezed in one direction. This step is a mandatory step for reconstructing a 3D surface through a photogrammetry approach. Similarly to our wedge display where several projectors are located at the entrance, we could work with several cameras and take advantage of the parallax to extract depth information on-the-fly. The multiscopic stream could also be rendered directly on a VR headset or on a portative autostereoscopic display. Adding a laser projector next to the camera(s) might also be useful for 3D reconstruction from structured lighting techniques. It might also be interesting to investigate whether Holographic Optical Elements would be compatible for acquisition purposes, in order to replace the Fresnel prisms and achieve a 90° redirection.

We showed that the concept of a wedge camera finds useful applications in the field of archaeology, and we wanted to communicate our work specifically to this community. However, its utility is more general and we think that such a product should be available in the market as it could potentially find applications in many fields. With this aim in mind, the apparatus should be made smaller to be more handy. As it was suggested to us, a smartphone-size would allow quick and effective inspections. However, this would make the imaged area smaller, and it is important to have a global vision of it. For this reason, a live stitching algorithm would greatly extend the possibilities. To help the stitching in areas where visible features (e.g. SIFT features) are not necessarily salient, the optical flow can be analyzed and inertial sensors like accelerometers could provide information regarding the relative motion of the camera. The live feedback, corrected from distortion, is an essential aspect, but the fact that a computer is required is limiting. Ideally, an embedded small screen would be useful, or a Bluetooth unit to cast the feedback toward a distant device like a phone or tablet.

**Transparent autostereoscopic display** Regarding the transparent autostereoscopic display, the correction of the chromatic issues is the first step to consider. Indeed, even though we have achieved a comfortable viewing zone size, the appearance of unwanted viewing zones should be corrected. To this, we need to find a way to reduce the spectral bandwidth of the HOE while keeping its angular bandwidth to a similar level. Angular multiplexing each projector position would potentially be

required, then the recording setup should be adapted to rotate the reference beam with respect to the hologram. Secondly, the viewing zones are quite large and this introduces visible transitions. Reducing the size of the viewing zones at the recording step can be interesting but would require more projectors to distribute them continuously. Alternatively, we can have a temporal multiplexing approach involving to mechanically move a high-speed projector, or an image of it through an optical system (i.e. vibrating fiber optics). Ideally, the eye pupil should be addressed by several viewing zones to achieve a Super Multi-View display.

The developed prototype will be used as a basis for further studies. In the short term, we plan to investigate the benefits of eye tracking. This way, the brightness of the display can be made more uniform, and the user could experience full parallax. This is also a way to get rid of the unwanted viewing zones by shutting down the projectors for users in these locations. Several users could be tracked at the same time as long as they are in independent viewing zones. We can also investigate more specifically augmented reality applications. These investigations can first focus on augmenting a fixed object of known geometry. Ultimately, it is possible to add cameras and analyze the real scene to augment through computer vision, thus providing real time dynamic augmentation. Interaction with the display is also an essential step and now, it is achieved on a regular keyboard. Injecting infrared light in the wedge guide could make it touchable with FTIR techniques in particular, but interacting with 3D data through touch interaction on a 2D surface may not be appropriate. Instead, gesture interactions might be better adapted and would allow for example the grabbing of virtual objects as in an AR headset. Again, this is possible by adding cameras and computer vision algorithms, or such dedicated commercial systems like LeapMotion™.

## 7.3 Perspectives of research

The most important message that we wanted to convey through this thesis is that the association of optics and computer science is essential for building the next generation display technologies. These two communities are too often separated, and cross-disciplinary communications between each introduce intermediaries, misunderstandings, and this eventually slows down the research. We saw that the limitations of 3D display technologies depend on both hardware and software and it is important that these issues are addressed conjointly, with a common goal, rather than separately. This is true for the development of 3D displays but many other fields are concerned: for example computer graphics/vision, cameras, healthcare, astronomy, and many more.

Autostereoscopic displays are currently the most mature 3D display technology and this is why we have oriented our research in this direction. Their compatibility with the current hardware and software makes them practical to implement at the moment. Their relative ease of development is highly determined by our current GPUs, which are mainly dedicated to single view rendering. I see them as an "intermediate" solution, filling short-term expectations until more dedicated hardware and software emerge. Yet, I do not think that the multiple view point rendering approach is the right one for a long-term view. On the opposite, light-field displays focus more at reproducing smooth parallax, and this is currently done at the cost of image quality and higher computational costs. Still, the light-field approach seems to me to be more long-sighted as it really targets the reproduction of the 3D scene and not some discrete views of it.

I think that the use of LBS projectors should be also largely investigated toward the light-field approach. Indeed, they allow a direct control of the flux in a specific direction, and this is also the goal of a light-field display. We can imagine for example a light-field display made of a 2D matrix of individual scanning mirrors addressing a controlled luminance.

For now, GPUs are the best way to compute a light-field, but this is not their primary goal. The development of dedicated Multiview Processing Units [Bur17] might allow real-time light-fields computing in the near future.

Augmented Reality is a very interesting tool to improve the understanding of both real and digital data, however, I think that the lack of true 3D augmentations is limiting right now.

On the one hand, we can investigate distant displays that allow a 3D augmentation of the real world with naked eyes, and this is the direction that we chose to follow through the Chapter 6. In particular, we showed that wedge guides, laser projectors and holographic optical elements might have a great future with this aim in mind. Our solution is not perfect yet, and I notably think it should be able to reproduce a full light-field in order to solve the conflicts between the real and virtual worlds, but it opens the way to new directions of research.

On the other hand, we can focus on improving the headset-based solutions. Current AR headsets offer exciting possibilities, but they still have limitations that make them impracticable for many applications. We could imagine lightweight AR glasses that are integrated in our corrective glasses, or even contact lenses, and render a perfect light-field that is aware of the real world. If such a solution exists, then what room would be left for any other display?

# Bibliography


[3Di17] 3D is 'No Longer The Default' Says IMAX — Screen Rant. `https://screenrant.com/imax-3d-movies-2d-screenings/`, 2017. (Accessed on 09/08/2019).

[3DT19] 3D TV is Dead - What You Need To Know. `https://www.lifewire.com/why-3d-tv-died-4126776`, 2019. (Accessed on 09/08/2019).

[AB91] Edward H. Adelson and James R. Bergen. The plenoptic function and the elements of early vision. In *Computational Models of Visual Processing*, pages 3–20. MIT Press, 1991.

[ABF+06] Tibor Agócs, Tibor Balogh, Tamás Forgács, Fabio Bettio, Enrico Gobbetti, Gianluigi Zanetti, and Eric Bouvier. A large scale interactive holographic display. *Proceedings - IEEE Virtual Reality*, 2006:57, 2006.

[AJ01] Hidenobu Arimoto and Bahram Javidi. Integral three-dimensional imaging with digital reconstruction. *Optics letters*, 26(3):157–159, 2001.

[All98] Pierre Allio. Autostereoscopic video device and system, September 15 1998. US Patent 5,808,599.

[ALP14] Ken Anjyo, John P Lewis, and Frédéric Pighin. Scattered data interpolation for computer graphics. In *ACM SIGGRAPH 2014 Courses*, page 27. ACM, 2014.

[BAGM07] J Blasco, N Aleixos, J Gómez, and E Moltó. Citrus sorting by identification of the most common defects using multispectral computer vision. *Journal of food engineering*, 83(3):384–393, 2007.

[BBC+07] Massimo Bertozzi, Alberto Broggi, Claudio Caraffi, Mike Del Rose, Mirko Felisa, and Guido Vezzoni. Pedestrian detection by means of far-infrared stereo vision. *Computer vision and image understanding*, 106(2-3):194–204, 2007.

[BBdH05] Ralph Braspenning, Eric Brouwer, and Gerard de Haan. Visual quality assessment of lenticular based 3d-displays. *2005 13th European Signal Processing Conference*, pages 1–4, 2005.

[BCP+08] Gareth P. Bell, Robert Craig, Robert Paxton, Gordon Wong, and Duncan Galbraith. 25.4: Invited paper: Beyond flat panels - multi layer displays with real depth. *SID Symposium Digest of Technical Papers*, 39(1):352–355, 2008.

[BDM+00] R. Borner, B. Duckstein, O. Machui, H. Roder, T. Sinnig, and T. Sikora. A family of single-user autostereoscopic displays with head-tracking capabilities. *IEEE Transactions on Circuits and Systems for Video Technology*, 10(2):234–243, March 2000.

[BDW06] Tim Bailey and Hugh Durrant-Whyte. Simultaneous localization and mapping (SLAM): Part II. *IEEE robotics & automation magazine*, 13(3):108–117, 2006.

[Bec15] Ronald G. Beckett. Application and limitations of endoscopy in anthropological and archaeological research. *The Anatomical Record*, 298(6):1125–1134, 2015.

[Bei81] Leo Beiser. Anaglyph stereoscopy, September 22 1981. US Patent 4,290,675.

[BFA+05] Tibor Balogh, Tamás Forgács, Tibor Agács, Olivier Balet, Eric Bouvier, Fabio Bettio, Enrico Gobbetti, and Gianluigi Zanetti. A scalable hardware and software system for the holographic display of interactive graphics applications. In *Eurographics (Short Presentations)*, pages 109–112, 2005.

[BFB+05] Tibor Balogh, T Forgács, Olivier Balet, E Bouvier, F Bettio, Enrico Gobbetti, and Gianluigi Zanetti. A scalable holographic display for interactive graphics applications. *IEEE VR 2005 Workshop on Emerging Display Technologies*, 01 2005.

[BKV11] John Bolton, Kibum Kim, and Roel Vertegaal. Snowglobe: A spherical fish-tank vr display. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 1159–1164, New York, NY, USA, 2011. ACM.

[BKV12] John Bolton, Kibum Kim, and Roel Vertegaal. A comparison of competitive and cooperative task performance using spherical and flat displays.

In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, CSCW '12, pages 529–538, New York, NY, USA, 2012. ACM.

[BL11] Pradipta Biswas and Pat Langdon. A new input system for disabled users involving eye gaze tracker and scanning interface. *Journal of Assistive Technologies*, 5(2):58–66, 2011.

[BLB+06] Sophie Boual, Timothy Large, Mark Buckingham, Adrian Travis, and Simon Munford. 72.3: Wedge displays as cameras. *SID Symposium Digest of Technical Papers*, 37(1):1999–2002, 2006.

[BMA16] Gary R Bradski, Samuel A Miller, and Rony Abovitz. Methods and systems for creating virtual and augmented reality, January 28 2016. US Patent App. 14/738,877.

[BR04] Leandro Bornaz and Fulvio Rinaudo. Terrestrial laser scanner data processing. In *XXth ISPRS Congress Istanbul*, 2004.

[BTV+17] Behnam Bastani, Eric Turner, Carlin Vieri, Haomiao Jiang, Brian Funt, and Nikhil Balram. Foveated pipeline for AR/VR Head-Mounted Displays. *Information Display*, 33(6):14–35, 2017.

[Buc08] Edward Buckley. 70.2: Invited paper: Holographic laser projection technology. *SID Symposium Digest of Technical Papers*, 39(1):1074–1079, 2008.

[Bur45] Harry Edwin Burton. The optics of Euclid. *JOSA*, 35(5):357–372, 1945.

[Bur17] Thomas L Burnett. 61-1: Invited paper: Light-field display architecture and the challenge of synthetic light-field radiance image rendering. In *SID Symposium Digest of Technical Papers*, volume 48, pages 899–902. Wiley Online Library, 2017.

[BW09] Hrvoje Benko and Andrew Wilson. Design challenges of interactive spherical user interfaces. In *Proceedings of CHI 2009*, 2009.

[BWB08] Hrvoje Benko, Andrew D. Wilson, and Ravin Balakrishnan. Sphere: Multi-touch interactions on a spherical display. In *Proceedings of UIST 2008*, 2008.

[CLK00] Robert T. Collins, Alan J Lipton, and Takeo Kanade. Introduction to the special section on video surveillance. *IEEE Transactions on pattern analysis and machine intelligence*, 22(8):745–746, 2000.

[CLRS15] P. Cigliano, V. Lippiello, F. Ruggiero, and B. Siciliano. Robotic ball catching with an eye-in-hand single-camera system. *IEEE Transactions on Control Systems Technology*, 23(5):1657–1671, Sep. 2015.

[CNH+07] Oliver S Cossairt, Joshua Napoli, Samuel L Hill, Rick K Dorval, and Gregg E Favalora. Occlusion-capable multiview volumetric three-dimensional display. *Applied optics*, 46(8):1244–1250, 2007.

[Cok08] Ronald S Cok. Method of manufacturing an OLED device with a curved light emitting surface, May 6 2008. US Patent 7,368,307.

[CPS+00] Hans J Coufal, Demetri Psaltis, Glenn T Sincerbox, et al. *Holographic data storage*, volume 8. Springer, 2000.

[CRG17] Thomas Crespel, Patrick Reuter, and Xavier Granier. A low-cost multi-touch spherical display: Hardware and software design. *SID Symposium Digest of Technical Papers (Display Week)*, 48(1):619–622, 2017. Online access: `https://hal.inria.fr/hal-01455523`.

[CRGT19] Thomas Crespel, Patrick Reuter, Xavier Granier, and Adrian Travis. Autostereoscopic transparent display using a wedge light guide and a holographic optical element. In *Digital Holography and Three-Dimensional Imaging 2019*, page M3A.4. Optical Society of America, 2019. Online access: `hal.inria.fr/hal-02129569`.

[CRR+17] Thomas Crespel, Brett Ridel, Clara Rigaud, Anke M. Brock, and Patrick Reuter. Code the globe: Interactive content for spherical multi-touch displays with simple webpages. In *Proceedings of the 6th ACM International Symposium on Pervasive Displays*, PerDis '17, pages 14:1–14:6, New York, NY, USA, 2017. ACM. Online access: `hal.inria.fr/hal-01523744`.

[CRT+19] Thomas Crespel, Patrick Reuter, Adrian Travis, Yves Gentet, and Xavier Granier. Autostereoscopic transparent display using a wedge light guide and a holographic optical element: implementation and results. *Applied Optics*, 2019. (Accepted, publication in progress).

[CTRG19] Thomas Crespel, Adrian Travis, Patrick Reuter, and Xavier Granier. Wedge cameras for minimally invasive archaeology. *Journal on Computing and Cultural Heritage*, 12(2):14:1–14:13, May 2019. Online access: `hal.inria.fr/hal-02096396`.

[CZF09] Cha Zhang, Zhaozheng Yin, and D. Florencio. Improving depth perception with motion parallax and its application in teleconferencing. In *2009 IEEE International Workshop on Multimedia Signal Processing*, pages 1–6, Oct 2009.

[Dav13] Brotherton-Ratcliffe David. Understanding diffraction in volume gratings and holograms. In *Holography-Basic Principles and Contemporary Applications*. IntechOpen, 2013.

[DDD+14] Song-Pei Du, Piotr Didyk, Frédo Durand, Shi-Min Hu, and Wojciech Matusik. Improving visual quality of view transitions in automultiscopic displays. *ACM Transactions on Graphics (TOG)*, 33(6):192, 2014.

[Des87] René Descartes. La dioptrique (1637). *Œuvres et Lettres*, 1987.

[DML+00] Neil A Dodgson, John R Moore, Stewart R Lang, Graham Martin, and Peter Canepa. A 50" time-multiplexed autostereoscopic display. *Proc. SPIE: Stereoscopic Displays and Virtual Reality Systems VII*, (January):24–26, 2000.

[DSG+08] Matthieu Debailleul, Bertrand Simon, Vincent Georges, Olivier Haeberlé, and Vincent Lauer. Holographic microscopy and diffractive microtomography of transparent samples. *Measurement Science and Technology*, 19(7):074009, 2008.

[EDF+16] Netalee Efrat, Piotr Didyk, Mike Foshey, Wojciech Matusik, and Anat Levin. Cinema 3D: large scale automultiscopic display. *ACM Transactions on Graphics (TOG)*, 35(4):59, 2016.

[Ein05] A Einstein. Einstein's proposal of the photon conceptfia translation of the annalen der physik paper of 1905. *Ann. Physik*, 17:132, 1905.

[ESC14] Jakob Engel, Thomas Schöps, and Daniel Cremers. LSD-SLAM: Large-scale direct monocular SLAM. In *European conference on computer vision*, pages 834–849. Springer, 2014.

[FCM09] Mark Freeman, Mark Champion, and Sid Madhavan. Scanned laser pico-projectors: Seeing the big picture (with a small device). *Opt. Photon. News*, 20(5):28–34, May 2009.

[FNH+02] G. E. Favalora, Joshua Napoli, D. M. Hall, R. K. Dorval, M. G. Giovinco, M. J. Richmond, and W. S. Chun. 100-Million-Voxel Volumetric Display. *SPIE's 16th Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls*, (April), 2002.

[Fre12] Cynthia Freeland. On being stereoblind in an era of 3d movies. *Essays in Philosophy*, 13(2):550–576, 2012.

[GCR15] Daphne Geerse, Bert Coolen, and Melvyn Roerdink. Kinematic validation of a multi-kinect v2 instrumented 10-meter walkway for quantitative gait assessments. *PLOS ONE*, 10:e0139913, 10 2015.

[Ger12] Guido Gerig. Structured lighting. *3D Computer Vision*, 2012.

[GG07] Carlo Bianchini Gabriele Guidi. TOF laser scanner characterization for low-range applications. volume 6491, pages 6491 – 6491 – 11, 2007.

[GH97] Rajiv Gupta and Richard I. Hartley. Linear pushbroom cameras. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(9):963–975, September 1997.

[GSB+65] D Gabor, GW Stroke, D Brumm, A Funkhouser, and A Labeyrie. Reconstruction of phase objects by holography. *Nature*, 208(5016):1159–1162, 1965.

[GSM+06] P. Gorrn, M. Sander, J. Meyer, M. Kroger, E. Becker, H.-H. Johannes, W. Kowalsky, and T. Riedl. Towards see-through displays: Fully transparent thin-film transistors driving transparent organic light-emitting diodes. *Advanced Materials*, 18(6):738–741, 2006.

[GWL67] Joseph Goodman and R W. Lawrence. Digital image formation from electronically detected holograms. *Applied Physics Letters*, 11:77 – 79, 09 1967.

[HBF+14] Yong Seok Hwang, Friedrich-Karl Bruder, Thomas Fäcke, Seung-Cheol Kim, Günther Walze, Rainer Hagen, and Eun-Soo Kim. Time-sequential autostereoscopic 3-D display with a novel directional backlight system based on volume-holographic optical elements. *Optics Express*, 22(8):9820, 2014.

[HG02] Klaus Hanke and Pierre Grussenmeyer. Architectural photogrammetry: basic theory, procedures, tools. In *ISPRS Commission*, volume 5, pages 1–2, 2002.

[HLgP+16] Jong-Young Hong, Chang-Kun Lee, Soon gi Park, Jonghyun Kim, Kyung-Hoon Cha, Ki Hyung Kang, and Byoungho Lee. See-through multi-view 3D display with parallax barrier. In Liang-Chy Chien, Sin-Doo Lee, and Ming Hsien Wu, editors, *Advances in Display Technologies VI*, volume 9770, pages 11 – 18. International Society for Optics and Photonics, SPIE, 2016.

[hol19] *Microsoft Hololens*, 2019. `www.microsoft.com/hololens`.

[HPL+16] Jong-Young Hong, Soon-Gi Park, Chang-Kun Lee, Seokil Moon, Sun-Je Kim, Jisoo Hong, Youngmin Kim, and Byoungho Lee. See-through multi-projection three-dimensional display using transparent anisotropic diffuser. *Opt. Express*, 24(13):14138–14151, Jun 2016.

[HR02] Ian P Howard and Brian J Rogers. *Seeing in depth, Vol. 2: Depth perception.* University of Toronto Press, 2002.

[HR08] Ian P. Howard and Brian J. Rogers. *Seeing in Depth: Volume 1: Basic Mechanics/ Volume 2: Depth Perception 2-Volume Set.* Oxford University Press, 2008.

[HSK+17] Jong-Ho Hong, Jae Min Shin, Gun Mo Kim, Hyejin Joo, Gyung Soon Park, In Bom Hwang, Min Woo Kim, Won-Sang Park, Hye Yong Chu, and Sungchul Kim. 9.1-inch stretchable AMOLED display based on LTPS technology. *Journal of the Society for Information Display*, 25(3):194–199, 2017.

[HSXS13] J. Han, L. Shao, D. Xu, and J. Shotton. Enhanced computer vision with microsoft kinect sensor: A review. *IEEE Transactions on Cybernetics*, 43(5):1318–1334, Oct 2013.

[Huy20] Christiaan Huygens. *Traité de la lumière.* 1920.

[HYFA00] Christiaan Huygens, Thomas Young, Augustin Jean Fresnel, and François Arago. *The wave theory of light: memoirs of Huygens, Young and Fresnel*, volume 15. American Book Company, 1900.

[IM80] Leonid Pinkhusovich IAroslavskii and Nikolai Stepanovich Merzliakov. Methods of digital holography. *Moscow Izdatel Nauka*, 1980.

[IYS93] Haruo Isono, Minoru Yasuda, and Hideaki Sasazawa. Autostereoscopic 3-D display using LCD-generated parallax barrier. *Electronics and Communications in Japan (Part II: Electronics)*, 76(7):77–84, 1993.

[IYS13] Tsubasa Ichikawa, Takuo Yoneyama, and Yuji Sakamoto. CGH calculation with the ray tracing method for the Fourier transform optical system. *Opt. Express*, 21(26):32019–32031, Dec 2013.

[JC89] Jeffrey A Jalkio and Steven K Case. Use of prisms to obtain anamorphic magnification, October 10 1989. US Patent 4,872,747.

[KA03] Janusz Konrad and Philippe Agniel. Artifact reduction in lenticular multiscopic 3D displays by means of anti-alias filtering. In *Stereoscopic Displays and Virtual Reality Systems X*, volume 5006, pages 336–347. International Society for Optics and Photonics, 2003.

[Kar12] Huda Karajeh. *Intermediate view reconstruction for multiscopic 3D display.* PhD thesis, Durham University, 2012.

[KBB+05] Martin Kaltenbrunner, Till Bovermann, Ross Bencina, Enrico Costanza, et al. TUIO: A protocol for table-top tangible user interfaces. In *Proc. of the The 6th International Workshop on Gesture in Human-Computer Interaction and Simulation*, pages 1–5, 2005.

[KC17] Bernard C. Kress and William J. Cummings. 11-1: Invited paper: Towards the ultimate mixed reality experience: Hololens display architecture choices. *SID Symposium Digest of Technical Papers*, 48(1):127–131, 2017.

[KHH17] Kota Kumagai, Satoshi Hasegawa, and Yoshio Hayasaki. Volumetric bubble display. *Optica*, 4(3):298–302, Mar 2017.

[KHL08] Hwi Kim, Joonku Hahn, and Byoungho Lee. Mathematical modeling of triangle-mesh-modeled three-dimensional surface objects for digital holography. *Appl. Opt.*, 47(19):D117–D127, Jul 2008.

[KLL+13] Chia-Wei Kuo, Ching-Huan Lin, Yi-Yang Liao, Yi-Hsiang Lai, Chin-Tang Chuang, Cheng-Nan Yeh, Jen-Kuei Lu, and Norio Sugiura. 8.3: Blur-Free Transparent LCD with Hybrid Transparency. *SID Symposium Digest of Technical Papers*, 44(1):70–73, 2013.

[KMCS12a] Abhijit Karnik, Walterio Mayol-Cuevas, and Sriram Subramanian. Mustard: A multi user see through ar display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 2541–2550, New York, NY, USA, 2012. ACM.

[KMCS12b] Abhijit Karnik, Walterio W. Mayol-Cuevas, and Sriram Subramanian. MUSTARD: a multi user see through AR display. In *CHI*, 2012.

[Kog69] Herwig Kogelnik. *Coupled Wave Theory for Thick Hologram Gratings*, pages 133–171. Nokia Bell Labs, 1969.

[KUY06] Hidei Kimura, Taro Uchiyama, and Hiroyuki Yoshikawa. Laser produced 3D display in the air. In *ACM SIGGRAPH 2006 emerging technologies*, page 20. ACM, 2006.

[LA02] A Loukianitsa and Putilin Andrey. Stereodisplay with neural network image processing. *Proceedings of SPIE - The International Society for Optical Engineering*, 05 2002.

[LCH08] Sheng Liu, Dewen Cheng, and Hong Hua. An optical see-through head mounted display with addressable focal planes. *Proceedings - 7th IEEE International Symposium on Mixed and Augmented Reality 2008, ISMAR 2008*, pages 33–42, 09 2008.

[Len17] Lenovo. Lenovo explorer. `https://www.lenovo.com/fr/fr/lenovo-explorer/`, 2017. (Accessed on 08/28/2019).

[Lev06] Tapani Levola. 7.1: Invited paper: Novel diffractive optical components for near to eye displays. *SID Symposium Digest of Technical Papers*, 37(1):64–67, 2006.

[LFL03] Thomas Ligon, Christopher Fedde, and Jonathan Lang. A self-contained spherical display system. In *ACM Siggraph 2003 Emerging Technologies*, 2003.

[LgPM+15] Chang-Kun Lee, Soon gi Park, Seokil Moon, Jong-Young Hong, and Byoungho Lee. Compact multi-projection 3d display system with light-guide projection. *Opt.*, 23(22):28945–28959, Nov 2015.

[LH96] Marc Levoy and Pat Hanrahan. Light field rendering. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 31–42. ACM, 1996.

[LH08] Taehee Lee and Tobias Hollerer. Hybrid feature tracking and user interaction for markerless augmented reality. pages 145 – 152, 04 2008.

[LH12] J. Laviole and M. Hachet. PapARt: Interactive 3D graphics and multi-touch augmented paper for artistic creation. In *2012 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 3–6, March 2012.

[lig05] Lightspace technologies. `www.lightspace3D.com`, 2005. (Accessed on 07/27/2019).

[Lin12] Tony Lindeberg. Scale invariant feature transform. 2012.

[Lip08] Gabriel Lippmann. Epreuves reversibles donnant la sensation du relief. *J. Phys. Theor. Appl.*, 7(1):821–825, 1908.

[Lip91] Lenny Lipton. Selection devices for field-sequential stereoscopic displays: a brief history, 1991.

[LJM+16] Seungjae Lee, Changwon Jang, Seokil Moon, Jaebum Cho, and Byoungho Lee. Additive light field displays: Realization of augmented reality with holographic optical elements. *ACM Trans. Graph.*, 35(4):60:1–60:13, July 2016.

[LOIB13] Jinha Lee, Alex Olwal, Hiroshi Ishii, and Cati Boulanger. Spacetop: Integrating 2d and spatial 3d interactions in a see-through desktop environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 189–192, New York, NY, USA, 2013. ACM.

[LP12] Junchang Li and Pascal Picart. *Digital holography*. Wiley Online Library, 2012.

[LSZ93] Michael Luby, Alistair Sinclair, and David Zuckerman. Optimal speedup of las vegas algorithms. *Information Processing Letters*, 47(4):173 – 180, 1993.

[Ltd02] Pufferfish Ltd. `pufferfishdisplays.co.uk`, 2002. (Accessed on 09/30/2019).

[LTL08] Chao-Te Lee, Chao-Hsu Tsai, and Hoang-Yan Lin. 32.2: The improvement of in-cell microretarder for stereoscopic lcd fabrication. *SID Symposium Digest of Technical Papers*, 39(1):448–451, 2008.

[LYK13] Changgeng Liu, Xiao Yu, and Myung K Kim. Phase aberration correction by correlation in digital holographic adaptive optics. *Applied optics*, 52(12):2940–2949, 2013.

[Max65] James Clerk Maxwell. VIII. A dynamical theory of the electromagnetic field. *Philosophical transactions of the Royal Society of London*, (155):459–512, 1865.

[MB97] James E Millerd and Neal J Brock. Holographic profilometry with a rhodium-doped barium titanate crystal and a diode laser. *Applied optics*, 36(11):2427–2431, 1997.

[MBM01] Edward M Mikhail, James S Bethel, and J Chris McGlone. Introduction to modern photogrammetry. *New York*, 2001.

[MC11] James Mentz and Samuel Caldwell. Method of stereoscopic synchronization of active shutter glasses, January 6 2011. US Patent App. 12/791,856.

[MCGV18] Marcos Maroto, Enrique Caño, Pavel Gonzalez, and Diego Villegas. Head-up displays (HUD) in driving. *CoRR*, abs/1803.08383, 2018.

[Mey01] Jean-Louis Meyzonnette. Bases de radiométrie optique. 2001.

[MGK17] Andrew Maimone, Andreas Georgiou, and Joel S. Kollin. Holographic near-eye displays for virtual and augmented reality. *ACM Trans. Graph.*, 36(4):85:1–85:16, July 2017.

[MHM+13] Daisuke Miyazaki, Noboru Hirano, Yuki Maeda, Siori Yamamoto, Takaaki Mukai, and Satoshi Maekawa. Floating volumetric image formation using a dihedral corner reflector array device. *Appl. Opt.*, 52(1):A281–A289, Jan 2013.

[MM05] Carlos H. Morimoto and Marcio R.M. Mimica. Eye gaze tracking techniques for interactive applications. *Computer Vision and Image Understanding*, 98(1):4 – 24, 2005. Special Issue on Eye Detection and Tracking.

[MMM15] Yuki Maeda, Daisuke Miyazaki, and Satoshi Maekawa. Volumetric aerial three-dimensional display based on heterogeneous imaging and image plane scanning. *Appl. Opt.*, 54(13):4109–4115, May 2015.

[Moo06] Christian Moore. `nuigroup.com`, 2006. (Accessed on 09/30/2019).

[MP04] Wojciech Matusik and Hanspeter Pfister. 3d tv. *ACM Transactions on Graphics*, 23:814, 08 2004.

[MPWL13] Mostafa Mehrabi, Edward M. Peek, Burkhard C. Wuensche, and Christof Lutteroth. Making 3d work: A classification of visual depth cues, 3d display technologies and their applications. In *Proceedings of the Fourteenth Australasian User Interface Conference - Volume 139*, AUIC '13, pages 91–100, Darlinghurst, Australia, Australia, 2013. Australian Computer Society, Inc.

[MSSM06] Daisuke Miyazaki, Kensuke Shiba, Koji Sotsuka, and Kenji Matsushita. Volumetric display system based on three-dimensional scanning of inclined optical image. *Opt. Express*, 14(26):12760–12769, Dec 2006.

[MT05] Christian Moller and Adrian R. L. Travis. Time-multiplexed autostereoscopic flat panel display using an optical wedge. In *IS&T/SPIE Electronic Imaging*, 2005.

[MWL+13] A. Maimone, G. Wetzstein, D. Lanman, M. Hirsch, R. Raskar, and H. Fuchs. Focus 3D: Compressive Accommodation Display. *ACM Trans. Graph.*, 32(5):1–13, 2013.

[MY00] Shigeru Murata and Norifumi Yasuda. Potential of digital holography in particle measurement. *Optics & Laser Technology*, 32(7-8):567–574, 2000.

[New84] Isaac Newton. *The Optical Papers of Isaac Newton: Volume 1, The Optical Lectures 1670-1672: Volume 1. The Optical Lectures 1670-1672*, volume 1. Cambridge University Press, 1984.

[NF09] Andrew Nashel and Henry Fuchs. Random hole display: A non-uniform barrier autostereoscopic display. In *2009 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video*, pages 1–4. IEEE, 2009.

[Nie18] Nielsen. Time flies: U.s. adults now spend nearly half a day interacting with media. `https://www.nielsen.com/us/en/insights/article/2018/time-flies-us-adults-now-spend-nearly-half-a-day-interacting-with-media/`, 07 2018. (Accessed on 10/05/2019).

[Nir11] Patel Nirav. Snow globe: Part one, cheap DIY spherical projection. `eclecti.cc/computergraphics/snow-globe-part-one-cheap-diy-spherical-projection`, 2011. (Accessed on 09/30/2019).

[Ocu16] Oculus. `https://www.oculus.com/`, 2016. (Accessed on 09/11/2019).

[OKH+16] Yoichi Ochiai, Kota Kumagai, Takayuki Hoshi, Jun Rekimoto, Satoshi Hasegawa, and Yoshio Hayasaki. Fairy lights in femtoseconds: Aerial and volumetric graphics rendered by focused femtosecond laser combined with computational holographic fields. *ACM Trans. Graph.*, 35(2):17:1–17:14, February 2016.

[OLG+05] Alex Olwal, Christoffer Lindfors, Jonny Gustafsson, Torsten Kjellberg, and Lars Mattsson. Astor: An autostereoscopic optical see-through augmented reality system. In *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 24–27. IEEE Computer Society, 2005.

[OMHC14] MI Olsson, MW Martin, JJ Hebenstreit, and PM Cazalet. Wearable device with input and output structures. *USA Patent No. US20130044042 A*, 1, 2014.

[PB06] Alex Poole and Linden J Ball. Eye tracking in HCI and usability research. In *Encyclopedia of human computer interaction*, pages 211–219. IGI Global, 2006.

[PBW88] Demetri Psaltis, David Brady, and Kelvin Wagner. Adaptive optical networks using photorefractive crystals. *Applied Optics*, 27(9):1752–1759, 1988.

[Pep90] John Henry Pepper. *True History of the Ghost: And All about Metempsychosis*. 1890.

[PJH16] Matt Pharr, Wenzel Jakob, and Greg Humphreys. *Physically based rendering: From theory to implementation*. Morgan Kaufmann, 2016.

[PKG+07] Tom Peterka, Robert Kooima, Javier Girado, Jinghua Ge, Dan Sandin, Andrew E. Johnson, Jason Leigh, Jürgen P. Schulze, and Thomas A. DeFanti. Dynallax: Solid state dynamic parallax barrier autostereoscopic VR display. *2007 IEEE Virtual Reality Conference*, pages 155–162, 2007.

[PPK00] Ken Perlin, Salvatore Paxia, and Joel S Kollin. An autostereoscopic display. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 319–326. ACM Press/Addison-Wesley Publishing Co., 2000.

[Pri70] Joseph Priestley. *A familiar introduction to the theory and practice of perspective*. J. Johnson and J. Payne, 1770.

[PSB13] Raviraj S Patkar, S Pratap Singh, and Swati V Birje. Marker based augmented reality using Android OS. *International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE)*, 3(5), 2013.

[PSK+18] Chan Il Park, Miryn Seong, Mi Ae Kim, Dojin Kim, Hyunju Jung, Mingu Cho, Sang Hoon Lee, Hyokang Lee, Sungjoon Min, Jaehyeong Kim, Minseok Kim, Jong-Hyun Park, Seyeoul Kwon, Binn Kim, Se June Kim, Weonseo Park, Joon-Young Yang, Sooyoung Yoon, and Inbyeong Kang. World's first large size 77-inch transparent flexible OLED display. *Journal of the Society for Information Display*, 26(5):287–295, 2018.

[PSS14] Ye Pan, William Steptoe, and Anthony Steed. Comparing flat and spherical displays in a trust scenario in avatar-mediated interaction. In *Proceedings of CHI 2014*, 2014.

[Pul17] Kari Pulli. 11-2: Invited Paper: Meta 2: Immersive Optical-See-Through Augmented Reality. *SID Symposium Digest of Technical Papers*, 48(1):132–133, 2017.

[PW16] Voicu S Popescu and Juan P Wachs. Simulated transparent display with augmented reality for remote collaboration, November 22 2016. US Patent 9,503,681.

[Rea09] RealFiction. Holographic display - world class 3D hologram display solutions. `https://www.realfiction.com/holographic-displays`, 2009. (Accessed on 07/31/2019).

[RG79] Brian Rogers and Maureen Graham. Motion parallax as an independent cue for depth perception. *Perception*, 8(2):125–134, 1979.

[RHS+11] Mohammad Osiur Rahman, Aini Hussain, Edgar Scavino, Hassan Basri, and M.A. Hannan. Intelligent computer vision system for segregating recyclable waste papers. *Expert Systems with Applications*, 38(8):10398 – 10407, 2011.

[RLKG+09] Randi J Rost, Bill Licea-Kane, Dan Ginsburg, John Kessenich, Barthold Lichtenbelt, Hugh Malan, and Mike Weiblen. *OpenGL shading language*. Pearson Education, 2009.

[RMDGR16] Brett Ridel, Lois Mignard-Debise, Xavier Granier, and Patrick Reuter. EgoSAR: Towards a Personalized Spatial Augmented Reality Experience in Multi-user Environments. *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pages 64–69, 2016.

[Rob03] David E Roberts. History of lenticular and related autostereoscopic methods. *Leap Technologies. Hillsboro*, 16, 2003.

[S+07] Richard Szeliski et al. Image alignment and stitching: A tutorial. *Foundations and Trends in Computer Graphics and Vision*, 2(1):1–104, 2007.

[SBD+08] Johannes Schöning, Peter Brandl, Florian Daiber, Florian Echtler, Otmar Hilliges, Jonathan Hook, Markus Löchtefeld, Nima Motamedi, Laurence Muller, Patrick Olivier, et al. Multi-touch surfaces: A technical guide. *Johannes Schöning, Institute for Geoinformatics University of Münster, Technical Report TUM-I0833*, 2008.

[SCP11] Daniel Shafer, Corey Carbonara, and Lucy Popova. Spatial presence and perceived reality as predictors of motion-based video game enjoyment. *Teleoperators and Virtual Environments - Presence*, 20:591–619, 12 2011.

[SDB12] Khaled Sarayeddine, Guilhem Dubroca, and Pascal Benoit. Informative eyeglasses, November 13 2012. US Patent 8,310,763.

[SH98] Armin Schwerdtner and Holger Heidrich. Dresden 3D display (D4D). In *Electronic Imaging*, 1998.

[SHBL+90] Pierre St-Hilaire, Stephen A Benton, Mark E Lucente, Mary Lou Jepsen, Joel Kollin, Hiroshi Yoshikawa, and John S Underkoffler. Electronic display system for computational holography. In *Practical Holography IV*, volume 1212, pages 174–182. International Society for Optics and Photonics, 1990.

[SHBL+93] Pierre St-Hilaire, Stephen A Benton, Mark E Lucente, John D Sutter, and Wendy J Plesniak. Advances in holographic video. In *Practical Holography VII: Imaging and Materials*, volume 1914, pages 188–196. International Society for Optics and Photonics, 1993.

[SHPW07] Hongyue Sun, David C Hendry, Michael A Player, and John Watson. In situ underwater electronic holographic camera for studies of plankton. *IEEE Journal of Oceanic Engineering*, 32(2):373–382, 2007.

[SHS+08] Phil Surman, Klaus Hopf, Ian Sexton, Wing Kai Lee, and Richard Bates. Solving the 3D problem - the history and development of viable domestic. In *Three-Dimensional Television*, pages 471–503. Springer, 2008.

[SJMS12] Brian Schowengerdt, Richard Johnston, C. Melville, and Eric Seibel. 47.4: Invited paper: 3d displays using scanning laser projection. *SID Symposium Digest of Technical Papers*, 43, 06 2012.

[SLH07] A. Schwerdtner, N. Leister, and R. Haeussler. 32.3: A new approach to electro-holography for TV and projection displays. *SID Symposium Digest of Technical Papers*, 38(1):1224–1227, 2007.

[SLH+08] Armin Schwerdtner, Norbert Leister, Ralf Haeussler, Stephan Reichelt, Gerald FÃ¼tterer, and Alexander Schwerdtner. 25.2: Eye-Tracking Solutions for Real-Time Holographic 3-D Display. *SID Symposium Digest of Technical Papers*, 39(1):345–347, 2008.

[SNMI10] Tomoyoshi Shimobaba, Hirotaka Nakayama, Nobuyuki Masuda, and Tomoyoshi Ito. Rapid calculation algorithm of fresnel computer-generated-hologram using look-up table and wavefront-recording plane methods for three-dimensional display. *Opt. Express*, 18(19):19504–19509, Sep 2010.

[SNS+18] D. E. Smalley, E. Nygaard, K. Squire, J. Van Wagoner, J. Rasmussen, S. Gneiting, K. Qaderi, J. Goodsell, W. Rogers, M. Lindsey, K. Costner, A. Monk, M. Pearson, B. Haymore, and J. Peatross. A photophoretic-trap volumetric display. *Nature*, 553(7689):486–490, jan 2018.

[Sny87] John Parr Snyder. *Map projections–A working manual*, volume 1395. US Government Printing Office, 1987.

[Sol13] Angelo G Solimini. Are there side effects to watching 3D movies? a prospective crossover observational study on visually induced motion sickness. *PloS one*, 8(2):e56160, 2013.

[SOT+04] Shiro Suyama, Sakuichi Ohtsuka, Hideaki Takada, Kazutake Uehira, and Sigenobu Sakai. Apparent 3-D image perceived from luminance-modulated two 2-D images displayed at different depths. *Vision Research*, 44(8):785 – 793, 2004.

[SR98] Lev S Sadovnik and Alexander Rizkin. 3-D volume visualization display, June 9 1998. US Patent 5,764,317.

[SRB10] Deqing Sun, Stefan Roth, and Michael J Black. Secrets of optical flow estimation and their principles. In *2010 IEEE computer society conference on computer vision and pattern recognition*, pages 2432–2439. IEEE, 2010.

[SS16] Hagen Seifert and Quinn Smithwick. 360-degree multi-viewer autostereoscopic tabletop display with omnidirectional dynamic parallax barrier and novel time-multiplexed directional backlight. *Electronic Imaging*, 2016(5), 2016.

[SSK+18] Yusuke Sando, Kazuo Satoh, Takahiro Kitagawa, Makoto Kawamura, Daisuke Barada, and Toyohiko Yatagai. Super-wide viewing-zone holographic 3D display using a convex parabolic mirror. *Scientific reports*, 8(1):11333, 2018.

[STU+00] S. Suyama, H. Takada, K. Uehira, S. Sakai, and S. Ohtsuka. 54.1: A Novel Direct-Vision 3-D Display Using Luminance-Modulated Two

2-D Images Displayed at Different Depths. *SID Symposium Digest of Technical Papers*, 31(1):1208–1211, 2000.

[Taa+16] Lalitha Saroja Thota, Jawaher alayed, Badryah asiri, Nada al zahrani, and Sara al qarni. Panoramic image stitching using sift and homograph. 2016.

[Tak14] Yasuhiro Takaki. Development of Super Multi-View Displays. *ITE Transactions on Media Technology and Applications*, 2(1):8–14, 2014.

[TCG+18] Adrian R.L. Travis, Liying Chen, Andreas Georgiou, Jiaqi Chu, and Joel Kollin. Wedge guides and pupil steering for mixed reality. *Journal of the Society for Information Display*, 26(9):526–533, 2018.

[THdJ+15] Yu-Hsiang Tsai, Mao-Hsiu Huang, Wei de Jeng, Ting-Wei Huang, Kuo-Lung Lo, and Mang Ou-Yang. Image quality affected by diffraction of aperture structure arrangement in transparent active-matrix organic light-emitting diode displays. *Appl. Opt.*, 54(28):E136–E145, Oct 2015.

[TKC+] F. Teubl, C. S. Kurashima, M. C. Cabral, R. D. Lopes, J. C. Anacleto, M. K. Zuffo, and S. Fels. Spheree: An interactive perspective-corrected spherical 3d display. In *2014 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*.

[TLB04] Adrian Travis, Tim Large, and Mark Buckingham. 20.4: Image quality in flat projection wedges. *SID Symposium Digest of Technical Papers*, 35(1):850–853, 2004.

[TLEB09] Adrian Travis, Tim Large, Neil Emerton, and Steven Bathiche. Collimated light from a waveguide for a display backlight. *Optics Express*, 17(22):19714–19719, 2009.

[TLEB13] Adrian RL Travis, Timothy A Large, Neil Emerton, and Steven N Bathiche. Wedge optics in flat panel displays. *Proceedings of the IEEE*, 101(1):45–60, 2013.

[TP06] E. Trucco and K. Plakas. Video tracking: A concise survey. *IEEE Journal of Oceanic Engineering*, 31(2):520–529, April 2006.

[TPZM00] ARL Travis, F Payne, F Zhong, and JR Moore. Flat panel display using projection within a wedge-shaped waveguide. In *Proceedings of the 20th International Display Research Conference*, volume 2000, pages 292–295. Society for Information Display (SID), 2000.

[Tra03] A.R.L. Travis. Flat-panel display using tapered waveguide, December 18 2003. Patent WO/2003/013151.

[TSN+10] Yutaka Tokuda, Yasuhiro Suzuki, Kunihiro Nishimura, Tomohiro Tanikawa, and Michitaka Hirose. Cloud display. In *Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology*, ACE '10, pages 32–35, New York, NY, USA, 2010. ACM.

[TTN11] Yasuhiro Takaki, Kosuke Tanaka, and Junya Nakamura. Super multi-view display with a lower resolution flat-panel display. *Optics express*, 19(5):4129–4139, 2011.

[TY15] Yasuhiro Takaki and Yuta Yamaguchi. Flat-panel see-through three-dimensional display based on integral imaging. *Opt. Lett.*, 40(8):1873–1876, Apr 2015.

[Und97] John S Underkoffler. Occlusion processing and smooth surface shading for fully computed synthetic holography. In *Practical Holography XI and Holographic Materials III*, volume 3011, pages 19–30. International Society for Optics and Photonics, 1997.

[URF08] S.W. Utt, P.C. Rubesin, and M.A. Foody. Display system having a three-dimensional convex display surface, 2008. US Patent 7,352,340.

[VB99] Cees Van Berkel. Image preparation for 3D LCD. In *Stereoscopic Displays and Virtual Reality Systems VI*, volume 3639, pages 84–91. International Society for Optics and Photonics, 1999.

[vBPF96] Cees van Berkel, David W. Parker, and Anthony R. Franklin. Multiview 3D LCD. In Mark T. Bolas, Scott S. Fisher, Mark T. Bolas, Scott S. Fisher, and John O. Merritt, editors, *Stereoscopic Displays and Virtual Reality Systems III*, volume 2653, pages 32 – 39. International Society for Optics and Photonics, SPIE, 1996.

[VDDP18] Athanasios Voulodimos, Nikolaos Doulamis, Anastasios Doulamis, and Eftychios Protopapadakis. Deep learning for computer vision: A brief review. *Computational intelligence and neuroscience*, 2018, 2018.

[VGP15] Fagun Vankawala, Amit Ganatra, and Amit Patel. A survey on different image deblurring techniques. *International Journal of Computer Applications*, 116(13):15–18, 2015.

[VIV16] VIVE — discover virtual reality beyond imagination. https://www.vive.com/fr/, 2016. (Accessed on 09/11/2019).

[Viv17] VividQ - Enabling highly realistic 3D holographic displays. https://www.vivid-q.com/, 2017. (Accessed on 08/03/2019).

[VWW10] Craig Villamor, Dan Willis, and Luke Wroblewski. Touch gesture reference guide. *Touch Gesture Reference Guide*, 2010.

[Wal13] Vivian K Walworth. History of polarized image stereoscopic display. In *Stereoscopic Displays and Applications XXIV*, volume 8648, page 864804. International Society for Optics and Photonics, 2013.

[WCH18] Zihao Wang, Ke Chen, and Renke He. Study on thermal comfort of virtual reality headsets. In *International Conference on Applied Human Factors and Ergonomics*, pages 180–186. Springer, 2018.

[Wet19] Gordon Wetzstein. Lecture notes - the graphics pipeline and OpenGL :transformations. https://stanford.edu/class/ee267/lectures/lecture2.pdf, 2019.

[WGP18] Jonathan David Waldern, Alastair John Grant, and Milan Momcilo Popovich. 17-4: DigiLens AR HUD Waveguide Technology. In *SID Symposium Digest of Technical Papers*, volume 49, pages 204–207. Wiley Online Library, 2018.

[Win53] H. J. J. Winter. The optical researches of Ibn Al-Haitham. *Centaurus*, 3(1):190–210, 1953.

[WLF+18] Zi Wang, Guoqiang Lv, Qinbin Feng, Anting Wang, and Hai Ming. Simple and fast calculation algorithm for computer-generated hologram based on integral imaging using look-up table. *Opt. Express*, 26(10):13322–13330, May 2018.

[WLHR11] Gordon Wetzstein, Douglas Lanman, Wolfgang Heidrich, and Ramesh Raskar. Layered 3D: Tomographic image synthesis for attenuation-based light field and high dynamic range displays. *ACM Trans. Graph.*, 30:95, 07 2011.

[WLHR12a] G. Wetzstein, D. Lanman, M. Hirsch, and R. Raskar. Tensor Displays: Compressive Light Field Synthesis using Multilayer Displays with Directional Backlighting. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 31(4):1–11, 2012.

[WLHR12b] Gordon Wetzstein, Douglas Lanman, Matthew Hirsch, and Ramesh Raskar. Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting. *ACM Transactions on Graphics*, 31:1–11, 07 2012.

[WSO+12] Jiantong Weng, Tomoyoshi Shimobaba, Naohisa Okada, Hirotaka Nakayama, Minoru Oikawa, Nobuyuki Masuda, and Tomoyoshi Ito. Generation of real-time large computer generated hologram using wavefront recording method. *Opt. Express*, 20(4):4018–4023, Feb 2012.

[XLL+13] Xinxing Xia, Xu Liu, Haifeng Li, Zhenrong Zheng, Han Wang, Yifan Peng, and Weidong Shen. A 360-degree floating 3d display based on light field regeneration. *Opt. Express*, 21(9):11237–11247, May 2013.

[YCX+18] Yan Yan, Ke Chen, Yu Xie, Yiming Song, and Yonghong Liu. The effects of weight on comfort of virtual reality devices. In *International Conference on Applied Human Factors and Ergonomics*, pages 239–248. Springer, 2018.

[YIMT02] Sumio Yano, Shinji Ide, Tetsuo Mitsuhashi, and Hal Thwaites. A study of visual fatigue and visual comfort for 3D HDTV/HDTV images. *Displays*, 23(4):191 – 201, 2002.

[YJXP10] Junyong You, Gangyi Jiang, Liyuan Xing, and Andrew Perkis. *Quality of Visual Experience for 3D Presentation - Stereoscopic Image*, pages 51–77. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[YSF10] G. Ye, A. State, and H. Fuchs. A practical multi-viewer tabletop autostereoscopic display. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, pages 147–156, Oct 2010.

[YSL+16] Linwei Yue, Huanfeng Shen, Jie Li, Qiangqiang Yuan, Hongyan Zhang, and Liangpei Zhang. Image super-resolution: The techniques, applications, and future. *Signal Processing*, 128:389 – 408, 2016.

[Zha00] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22, 2000.

[ZJ04] Zhiwei Zhu and Qiang Ji. Eye and gaze tracking for interactive graphic display. *Machine Vision and Applications*, 15(3):139–148, Jul 2004.

[ZWM+17] Qian Zhou, Kai Wu, Gregor Miller, Ian Stavness, and Sidney Fels. 3dps: An auto-calibrated three-dimensional perspective-corrected spherical

display. In *2017 IEEE Virtual Reality (VR)*, pages 455–456. IEEE, 2017.