# Study and optimization of an optical see-through near to eye display system for augmented reality

Jianming Yang

UNIVERSITÉ DE STRASBOURG

**ÉCOLE DOCTORALE MSII (ED n°269)**

**LABORATOIRE ICUBE (UMR 7357)**

# THÈSE

Présentée par:


## YANG JIANMING


Soutenue le 29 mars 2018

pour obtenir le grade de: **Docteur de l'Université de Strasbourg**

Discipline/ Spécialité: **Electronique, Microélectronique, Photonique**

**Study and optimization of an optical see-through near to eye display system for augmented reality**

**CONFIDENTIAL**

**THÈSE dirigée par:**

| | |
|---|---|
| **M. Philippe Gérard** | Maître de conférences, INSA de Strasbourg |

**RAPPORTEURS:**

| | |
|---|---|
| **M. Alois Herkommer** | Professeur, Université de Stuttgart |
| **M. Thierry Lepine** | Maître de conférences, Institut d'Optique Graduate School |

**AUTRES MEMBRES DU JURY:**

| | |
|---|---|
| **M. Bruno Serio** | Professeur, Université de Paris Ouest |
| **M. Patrice Twardowski** | Maître de conférences, Université de Strasbourg |
| **M. Joël Fontaine** | Professeur, INSA de Strasbourg |

# Acknowledgements

I would like to take this opportunity to express my appreciation and thanks for all the persons that gave me the encouragement and support, to reach this point in my academic career. I would like to greatly thank my principal advisor: Philippe Gerard for the help of using software and teaching me methodology. I would also like to pay my highest tribute to my co-advisor Dr. Patrice Twardowski for your kind and professional guidance. The correction of the thesis and articles has taken your many efforts. I would like to express my great gratitude to Prof. Joël Fontaine for your recognition, admission and guidance. Without you, I cannot begin a PhD in France.

I would also like to thank all the members of photonics instrumentation and processes team; especially, Dr. Sylvain Lecler for guidance on photonic nano-jet. Dr. Pierre Pfeiffer for measuring surface quality of our samples, Prof. Paul Montgomery for your English corrections, Dr. HALTER Eric for teaching me Français, Dr. Stéphane Roques for polishing help. I would also like to give a warm thank to all the colleagues from the Icube-IPP team. I would like to thank Ms. Marine Auvert from Société d'Accélération du Transfert de Technologies (SATT) who helped us to obtain financial support from SATT.

I thank my PhD committee: Prof. Bruno SERIO, Prof. Alois HERKOMMER, Dr. Thierry LEPINE who provided me with their time and attention to evaluate my work and give me very useful suggestions. I would also like to thank Dr. Bernard Kress for coming to my defense. Your suggestions would have a profound impact to me.

Finally and importantly, I dedicate very special thanks to my parents, whose love and support through my whole life has given me the passion to pursue my education. I would like also to express many thanks to Chinese schoolmate: LIU Chenchen for your many helps, YU Wenhui for ge.zhong.che.dan, YANG Xiucheng, LIU Guixian for ge.zhong.wan.shua, SUN Xiaoguang for ge.zhong.he.zui, YUAN Ye, ZHEN Xiaopo, TANG Shuangqi, WANG Xinyue, HU Ronghai, WANG Jian for shitang chifan and wanshua, TANG Fujiao and DING Yifan for zhaoxiang, LIU Yingdong, WANG Wen for help and many others.

# List of abbreviations

| | |
|---|---|
| **NED** | Near to eye display |
| **HMD** | Head mounted display |
| **HWD** | Head-worn display |
| **AR** | Augmented reality |
| **VR** | Virtual reality |
| **FOV** | Field-of-view |
| **DFOV** | Diagonal field-of-view |
| **HFOV** | Horizontal field-of-view |
| **VFOV** | Vertical field-of-view |
| **DOE** | Diffractive optical element |
| **HUD** | Head-up display |
| **EPE** | Exit pupil expander |
| **TIR** | Total internal reflection |
| **SBG** | Switchable Bragg Grating |
| **FDTD** | Finite-Difference Time-Domain method |
| **FEM** | Finite element method |
| **PML** | Perfect matched layer |
| **ASM** | Angular spectrum method |
| **HFL** | Harmonic Fresnel Lens |
| **RAP** | Right-angle prism |
| **PNJ** | Photonic nano-jet |
| **FWHM** | Full width at half maximum |
| **WGM** | Whispering gallery Mode |

# Résumé

Récemment, les afficheurs placés près de l'œil (*NED: Near to Eye Display*) de type lunette pour la réalité augmentée (RA) sont devenus un sujet en vogue en raison de la sortie de certains produits commerciaux tels que Microsoft HoloLens et Google Glasses. Pour les systèmes de RA, ces dispositifs permettent aux utilisateurs de voir simultanément l'image virtuelle générée par ordinateur (information ajoutée) et la scène du monde réel; l'ensemble formant ce qu'on appelle la réalité augmentée. Les dispositifs de RA peuvent être utilisés dans de nombreux domaines d'application tels que l'aviation militaire, les systèmes de divertissement commerciaux et les systèmes de visualisation pour applications médicales. Plusieurs approches pour la réalisation de ces dispositifs ont été utilisées : des structures planaires semi-réfléchissantes, des systèmes hybrides réfractifs-réflectifs, des prismes à forme libre, une optique adaptative de correction, des systèmes de projection rétinienne et des guides d'ondes géométriques de type lunette (*GWNED: Guided Waves Near to Eye Display*). Les GWNED sont constitués d'un sous-système appelé coupleur d'entrée (*incoupler*) qui permet de coupler l'image générée par un ordinateur dans le ou les guides d'onde et d'un sous-système appelé coupleur de sortie (*outcoupler*) qui couple l'image propagée par le ou les guides d'onde vers l'œil du porteur de lunette. Le dispositif peut être monoculaire ou binoculaire suivant que l'on équipe un œil ou les deux yeux. Les coupleurs de sortie offrent l'avantage d'augmenter la taille de la pupille de sortie du système sans trop augmenter la complexité du coupleur d'entrée. De plus, ces structures sont très compactes. Pour le coupleur de sortie, la société Lumus utilise par exemple un réseau de miroirs partiellement réfléchissants. Cet ensemble est spécialement conçu pour avoir une réflectance adaptée aux différentes longueurs d'onde et angles d'incidence. La société Optinvent quant à elle utilise de petits miroirs en cascade pour le couplage de sortie. Zhao et al. ont introduit un micro-coupleur de sortie à haute visibilité utilisant des microstructures triangulaires. Les conceptions associées à ces systèmes sont principalement basées sur la simulation par tracé de rayons. Un autre type de coupleur de sortie utilise des éléments optiques diffractifs (EOD) ou des éléments optiques holographiques (EOH). Ces éléments se retrouvent dans les produits des sociétés Microsoft, Konica Minolta, Sony et BAE Systems. Pour le coupleur d'entrée, des lentilles de projection sont utilisées pour des système de type GWNED avec de petits champs de vision (*FOV: Fielf-Of-View*). Ces systèmes nécessitent également un expanseur de pupille bidimensionnel à

l'intérieur du guide d'onde. La conception combinant un prisme de forme libre et un guide d'ondes permet d'obtenir une valeur du champ de vision plus grande à l'aide de deux pupilles séparées. Cependant, le champ de vision vertical est aujourd'hui encore limité à environ 30 degrés.

Pour faire un NED efficace et utile, il faut réaliser un compromis entre les performances optiques, le poids et les dimensions. Jusqu'à présent, à notre connaissance, aucune conception ne peut répondre simultanément à toutes ces exigences. De nos jours, de nombreux types de NED existent avec des performances optiques différentes en termes de FOV et de volume de déplacement admissible pour l'œil VDAO (*Eye-box*). A titre d'exemple, pour les guides d'ondes de type NED, l'appareil Lumus OE-31 a un champ de vision avec une diagonale de 40 degrés et un VDAO de 10 x 8 mm.

# 1. Objectifs de la thèse

Les objectifs de la thèse sont les suivants :

- Concevoir de nouveaux systèmes qui peuvent améliorer les performances optiques des NED.
- Étudier les configurations géométriques des systèmes proposés.
- Étudier les propriétés optiques des systèmes proposés.
- Prouver le concept de manière expérimentale pour un système sélectionné.
- Étudier de nouvelles solutions pour des futures améliorations.

# 2. Description des travaux réalisés

## 2.1. Systèmes proposés basés sur des optiques à surfaces à forme libre

Pour atteindre les objectifs spécifiés ci-dessus, nous proposons le système représenté Fig. 1. La figure montre la structure du système et le trajet des rayons à l'intérieur des composants. Deux guides d'ondes sont empilés avec un petit espace d'air entre eux afin que la lumière puisse se propager à l'intérieur des deux guides d'ondes indépendamment. La lumière émise par le micro-afficheur est couplée dans le premier guide d'ondes par le coupleur d'entrée. La lumière se propage à l'intérieur du premier guide d'ondes jusqu'à atteindre un miroir cylindrique à l'extrémité du guide. Dans la zone

proche du miroir, les deux guides d'ondes sont combinés ensemble. Ensuite, la lumière réfléchie par le miroir cylindrique est couplée dans le deuxième guide d'ondes. Ce couplage a été étudié et optimisé. En utilisant des méthodes existantes telles que les hologrammes ou les miroirs en cascade, la lumière est couplée hors du deuxième guide d'ondes vers l'œil du porteur de lunette. Le système de couplage d'entrée se compose d'un prisme avec quatre surfaces optiques à forme libre différentes (de S2 à S5) et d'une lentille constituée d'une surface à forme libre (S1) et d'une surface plane. La pupille de couplage a une forme rectangulaire avec une taille de 1,9 mm (H) × 12 mm (V) qui agit comme un diaphragme d'ouverture et réduit considérablement la lumière parasite. Le rayon principal de chaque champ passe par le centre de la pupille d'entrée du guide d'onde.

Après la réflexion par le miroir cylindrique, le rayon principal de chaque champ converge vers le centre de la pupille de sortie (œil) à travers le deuxième guide d'ondes et le coupleur de sortie. Ensuite, chaque champ peut être vu comme indiqué dans la Fig. 1(c).

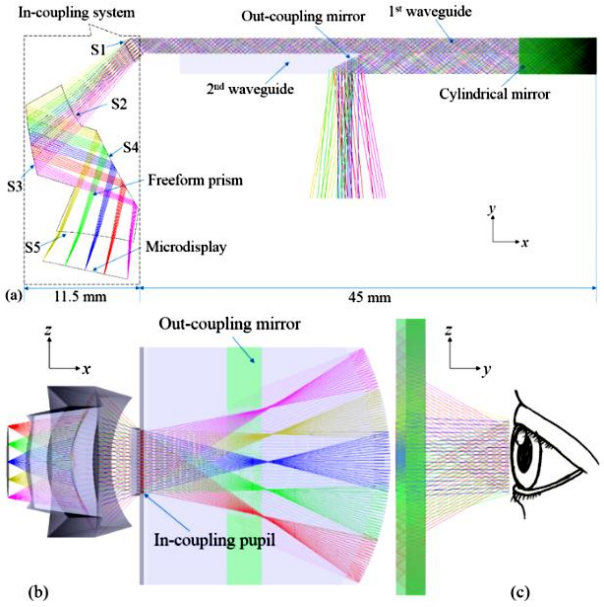



Fig. 1. La structure entière et le chemin optique du système proposé avec différents points de vue : (a) vue yx, (b)vue zx, (c)vue zy.

Nous avons proposé de nouveaux guides d'ondes polychromatiques NED avec un prisme de forme libre et des lentilles de forme libre. Les conceptions ont été réalisées avec le logiciel Zemax. Un micro-afficheur OLED de 0,61 pouce de diagonale avec 852×600 pixels a été utilisé comme source d'image. Le côté des pixels de forme carrée était de 15 μm, ce qui correspond à une fréquence de

Nyquist de 33 npl/mm. En conséquence, les valeurs horizontale et verticale du FOV étaient respectivement de 30° et 60°. Les valeurs de la FTM (Fonction de Transfert de Modulation) des champs objets étaient supérieures à 30% pour 33 npl/mm, ce qui est suffisant pour un système visuel. De plus, une très grande VDAO d'environ 10×8 mm permettant de grands mouvements oculaires a été obtenue.

Cependant, afin de réduire les contraintes sur le processus de fabrication et les tolérances, nous avons conçu et optimisé de nouveaux systèmes avec des lentilles à symétrie de rotation.

## 2.2. Systèmes proposés basés sur des lentilles à symétrie de rotation

Dans les propositions précédentes, comme le miroir cylindrique modifie la symétrie du système, des surfaces optiques à forme libre étaient nécessaires pour compenser l'asymétrie. La fabrication, le test et l'alignement des composants de surface à forme libre sont encore aujourd'hui un défi. C'est pourquoi nous avons conçu d'autres systèmes qui n'utilisent que des lentilles à symétrie de rotation comme coupleurs d'entrée. Pour maintenir les performances optiques, plusieurs prismes à angle droit (PAD) sur une base cylindrique ont été utilisés, au lieu d'un seul miroir cylindrique. Les PAD n'ont pas de puissance optique et peuvent rétroréfléchir la lumière incidente. De cette façon, le système est plus compact et la difficulté de fabrication du système est largement réduite. Une vue artistique est présentée à la Fig. 2. Les propriétés optiques (qualité d'image, efficacité énergétique, lumière parasite et tolérance) du système utilisant les PAD ont été étudiées en détail.

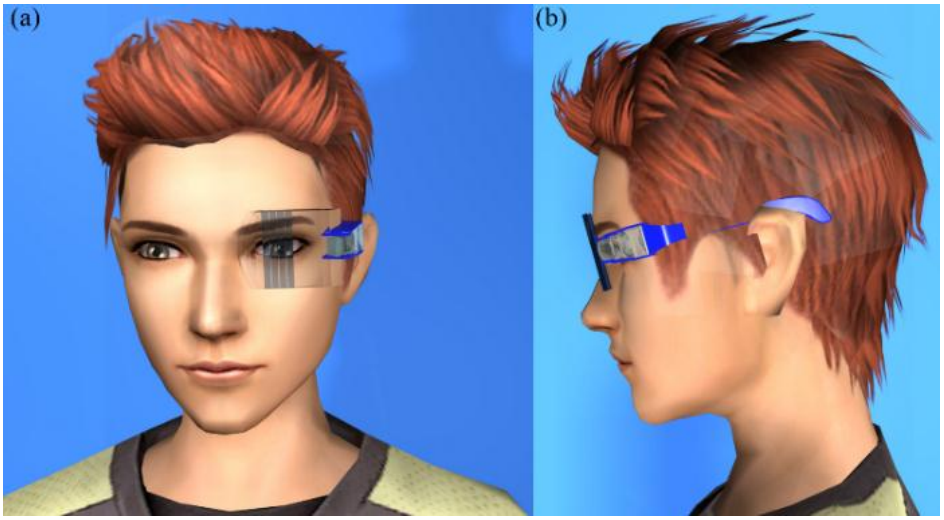



Fig. 2. Vue artistique (a) vue de face, (b) vue de côté.

## 2.3. Résultats expérimentaux pour la structure du guide d'ondes

Nous avons défini et commandé la fabrication des deux guides d'ondes et de la pièce de transition avec des PAD. En raison de contraintes de temps, nous n'avons pas commandé les éléments du coupleur d'entrée, ni réalisé leur assemblage. Cependant, ce sous-système est une sorte d'oculaire qui peut transformer des sources ponctuelles objets du micro-afficheur en ondes planes qui sont couplées à l'intérieur du premier guide d'ondes. Pour cette conception, la partie critique et la plus importante est l'ensemble constitué des deux guides d'ondes empilés et couplés à l'aide des PAD.

Nous avons discuté de la tolérance de fabrication et du processus d'assemblage des trois composants. Ensuite, nous avons présenté les résultats de mesure avec le FOV, l'efficacité du couplage pour les ondes planes d'entrée, leur uniformité en intensité et le coefficient de transmission lié à la vue du monde extérieur. Les résultats ont prouvé la validité du concept de la structure à deux guides d'ondes. Pour protéger la propriété intellectuelle, un brevet international a été déposé. Il a été approuvé et publié en juillet 2017.

Afin de compléter l'étude des NED optiques, nous avons également étudié des conceptions alternatives utilisant des lentilles diffractives, plus précisément les lentilles diffractives harmoniques.

## 2.4. Simulation FDTD d'une lentille de Fresnel harmonique (LFH)

Pour réduire davantage le poids et la taille du système, on peut utiliser un élément optique diffractif (EOD). Par exemple, le coupleur d'entrée peut être remplacé par des EOD dont l'épaisseur peut être négligée. Comme nous sommes intéressés par un système polychromatique, un EOD qui peut fonctionner pour plusieurs longueurs d'onde est nécessaire. Une famille de EOD appelée lentille de Fresnel harmonique (LFH) permet un fonctionnement pour plusieurs longueurs d'onde spécifiques suivant le nombre harmonique choisi. Les publications sur ce sujet sont principalement axées sur les propriétés paraxiales des LFH. Toutefois, lorsqu'une LFH sert de composant dans le coupleur d'entrée, il faut tenir compte de ses propriétés non paraxiales, car la lumière peut avoir un angle d'incidence important. Afin d'obtenir des résultats précis, nous avons utilisé la méthode FDTD (*Finite-Difference Time-Domain*) pour simuler la performance à proximité de l'élément diffractif,

puis nous avons utilisé la projection en champ lointain pour calculer l'intensité du champ électrique dans l'espace libre. Le logiciel commercial Lumerical FDTD Solutions a été utilisé pour les simulations. La figure 3(a) donne un exemple de distribution de l'intensité du champ électrique à proximité d'une LFH et (b) la distribution additionnelle en champ lointain. De cette façon, nous pouvons traiter des géométries relativement grandes et évaluer les efficacités de diffraction.

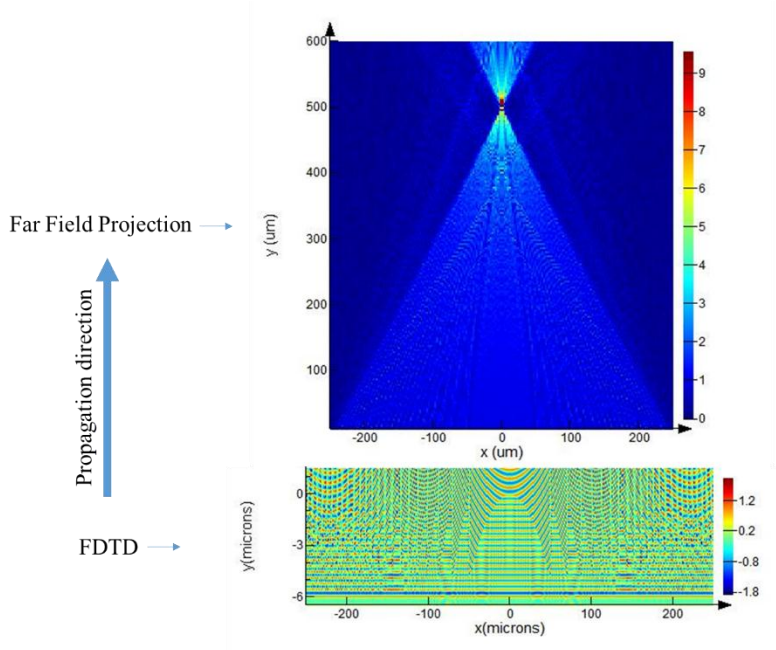



Fig. 3. FDTD combinée à la projection en champ lointain

Cette étude a montré une façon prometteuse de réduire davantage le poids et la taille du coupleur d'entrée. Les LFH présentent l'avantage de bien fonctionner avec plusieurs longueurs d'onde discrètes pour une incidence normale. Nous avons étudié l'efficacité, la distance focale, l'intensité maximale au foyer et les dimensions de la zone focale (*FWHM: Full-Width-Half-Maximum*), largueur à mi-hauteur de l'intensité dans les deux directions, pour plusieurs nombres harmoniques et nombres d'ouverture. Nous pouvons en conclure que :

- Les longueurs d'onde harmoniques ne sont pas sensibles au nombre d'ouverture ;
- L'efficacité de diffraction et la FWHM de la zone focale augmentent au fur et à mesure que le nombre d'ouverture augmente ;
- La polarisation incidente a un effet négligeable sur l'efficacité, la distance focale, l'intensité maximale de la zone focale et la FWHM de la zone focale ;
- La dispersion normale des verres optiques a également un effet négligeable ;
- L'efficacité de diffraction diminue à mesure que l'angle d'incidence oblique augmente en

raison des aberrations ;

- Avec un nombre d'ouverture plus grand, les LFH peuvent accepter des angles d'incidence plus grands.
- Les longueurs d'onde harmoniques sont également inchangées avec des angles d'incidence différents si les aberrations ne sont pas trop élevées.

En raison de ces propriétés, les LFH sont très prometteuses en tant que coupleurs dans les systèmes d'affichage de type lunette afin de les rendre plus compacts. Des études supplémentaires dédiées à l'utilisation de ces composants méritent d'être menées. Cependant, pour faciliter le travail de conception, certains modèles de surface spécifique adaptés aux LFH pour différentes longueurs d'onde devraient être développés pour les logiciels de conception optique commerciaux comme par exemple Zemax ou Code V.

Avant la simulation des LFH par FDTD, nous nous sommes formés sur différents systèmes optiques, notamment un système composé d'un cuboïde diélectrique noyé dans un cylindre diélectrique afin de générer un jet photonique ultrafin. Ce travail a fait l'objet d'une publication dans Optics Express .

# 3. Conclusion

Dans le cadre de cette recherche, nous avons proposé de nouvelles conceptions d'afficheurs proches de l'œil de type lunette pour la réalité augmentée. Nos conceptions peuvent augmenter le champ de vision vertical jusqu'à 60°, valeur presque deux fois plus grande que dans les systèmes de type guide d'ondes existants. Nous avons proposé des systèmes améliorés dans lesquels seules des lentilles à symétrie de rotation sont nécessaires. Les difficultés de fabrication et d'assemblage sont alors largement réduites. Les deux guides d'ondes et la pièce de transition avec des primes à angles droits ont été définis puis fabriqués. Nous avons caractérisé leurs propriétés géométriques et optiques. En mesurant expérimentalement le champ de vision, l'efficacité de couplage et l'uniformité en intensité, le concept proposé pour la structure à deux guides d'ondes a pu être validé.

Pour réduire davantage le poids, nous avons étudié ensuite les lentilles de Fresnel harmoniques, qui peuvent fonctionner avec précision pour plusieurs longueurs d'onde spécifiques, en utilisant la

méthode FDTD et la projection en champ lointain. Les propriétés optiques non paraxiales ont été étudiées. Pour gérer efficacement et finement les conceptions électromagnétiques, nous avons utilisé FEM de COMSOL. Nous avons aussi simulé un nano-jet photonique (PNJ) comme cas d'essai. La simulation d'un cuboïde diélectrique inséré dans un cylindre diélectrique a montré divers cas intéressants : un PNJ se focalisant à l'extérieur de la surface externe du cuboïde, un long jet photonique, la génération de modes de galerie et un effet de type jet photonique courbe (*photonic hook*).

Pour concevoir un sous-système innovant comme coupleur de sortie, les métasurfaces semblent une voie très prometteuse. Leur conception devrait compléter nos résultats et mener à diverses solutions attrayantes pour les futurs afficheurs de réalité augmentée de type lunette.

# Contents

# List of Figures

# List of Tables

# 1. Introduction

## 1.1. The study scopes

The scope of the thesis is near to eye displays (NEDs) for augmented reality (AR). NEDs also known as Head Mounted Displays (HMDs) or Head-Worn Displays (HWDs). They are devices that can be worn by human and display information in front of the eyes. Some of them are for virtual reality (VR) by which only the virtual image is displayed, and the real environment is blocked. The others are for AR by which the real environment and virtual image are optically superimposed by see-through capability.

Recently, as the emergence of some commercial product for instance Google Glass, HoloLens and Oculus Rift. NEDs has become a very hot topic. Many companies are involved in this domain like Microsoft, Google, Facebook, Samsung, Lumus, Magic Leap, Optinvent, etc. Some of them are mainly focused on VR while others are on AR. However, there are still great challenges to overcome. High optical performance (large field-of-view, eye box, high resolution) in nowadays design implies increasing dimension and weight.

The element used inside a NED can be refractive, reflective and/or diffractive. New technologies such as metasurfaces are very promising because the existing architectures still have big limitations. However, the associated manufacture technologies are limited and costly. In this thesis, we have mainly focused our work on traditional optical elements which can be modeled by geometrical and scalar wave optics, with their associated tools: raytracing and MTF (Modulation Transfer Function) analysis. In order to evaluate new path in the design of in-coupler for future NEDs, in the final chapter, we have also investigated diffractive optical elements using a complete electromagnetic approach.

## 1.2. Motivation

The theme of the thesis for the author was originally motivated by chance (Choosing from several available research topics)! However, the continuity of the thesis is supported by the history and background of our lab as this is also a topic of a previous thesis. From a scientific point of view, the necessity of the research is due to the nowadays apparent contradiction between the optical performance and the compactness of the NED system. The existing NEDs are always cumbersome. The purpose of the thesis is to explore some new ways that may overcome the contradiction.

## 1.3. Organization of thesis

The content of the thesis will be organized into five chapters and a conclusion including perspectives for future works. In chapter 2, we firstly present an introduction about human eyes. Then we present a review about the nowadays architectures on NEDs. We classify all the NEDs into three families: NEDs with free space combiner, freeform prism and waveguide. In chapter 3, we present our several designs. All the designs involve a two-layer waveguide with a cylindrical edge. The main differences lie in the design of the in-coupler. Considering a cylindrical mirror as edge element, we describe three different in-coupler designs using respectively: a freeform prism, freeform mirrors and freeform lenses. In chapter 4, we introduce an improved design which also involves a two-layer waveguide but a different edge. The edge is composed of many right-angle prism (RAP) placed along a cylindrical base. The benefit is that we can have a rotationally symmetric in-coupler. All the proposed designs have been protected by an international patent (WO2017121967A3). In chapter 5, we present the preliminary encouraging experimental results related to the design described in chapter 4. In chapter 6, we consider a kind of diffractive optical element called Harmonic Fresnel Lens (HFL). The purpose of this study is the possibility of using HFL as a component inside the in-coupler to further reduce the dimension and weight. We use Finite-Difference Time-Domain (FDTD) plus Angular Spectrum method (AS) to

calculate the response of HFL accurately. Finally, we present the general conclusion and the future perspectives of this study.

# 1.4. Definition of terms

**Field-Of-View (FOV):** The FOV for a NEDs is the apparent angular size of the virtual image seen by the viewer through the NEDs.

**Eye box:** In pupil-forming designs, exit pupil is more used than eye box, while in non-pupil-forming system, eye box means the common area inside which we can see the whole FOV.

**Eye relief:** The shortest distance between the eye and the first component of the NEDs.

# 2. Literature review

This chapter introduces the fundamental concepts of see-through NEDs (Near to Eye Displays) from the perspective of optical design, image quality and field of view. Each NED design is application and user specific. After a short presentation of augmented reality and see-through displays, we begin by a summary of the human visual system and the key factors guidelines for the NED design and the user acceptance. Then we give an overview of the different approaches used in the optical design of NEDs. Finally, the quality of the display imagery presented to the user's eye is discussed.

## 2.1. Augmented reality and see-through displays

The paradigm of **Augmented Reality** (AR) aims at helping the interaction of the user with the real world (i.e. his physical environment, composed of physical objects) and the virtual world (i.e. data and processing means at his disposal) [1]. Milgram and his colleagues interpret augmented reality as being included in a linear continuum going from real to virtual [2]. They define the term of *mixed reality* as the spacing between real and virtual. This mixed reality contents augmented reality, but also augmented virtuality, which consists in integrating virtual elements into the real world. Dubois and his colleagues propose their own taxonomy [3] based on the task object and the type of augmentation (see **Fig. 2.1**).

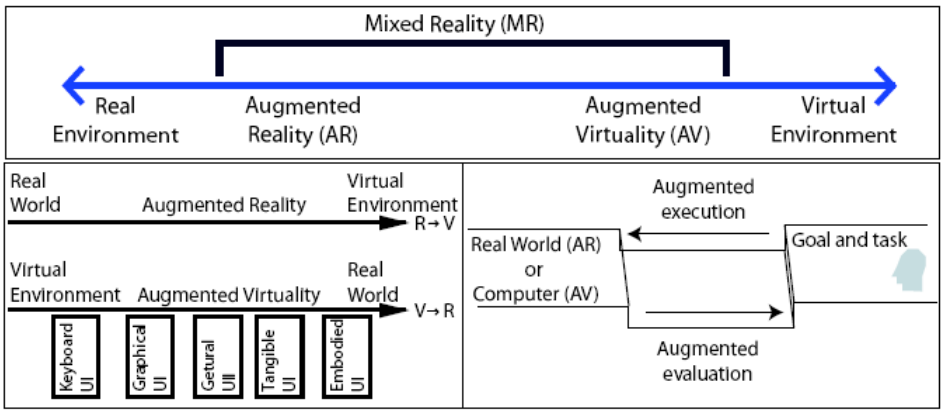



**Fig. 2.1.** Taxonomy [1, 2].

Here we consider only **see-though** Augmented Reality displays where symbiology or imagery is optically superimposed into a see-through forward view of the user. The distinction between AR and virtual reality systems is very important because they are very distinctly different applications and then have very distinctly different performance requirements.

Since the 1960s, see-through AR displays known as HUDs (**Head-Up Displays**) have been developed for military applications especially military aircraft. For this class of displays, a collimator is placed in a fixed location in front of the pilot in order to display virtual information in the far field [4]. The transparent collimator does not obstruct the pilot's view and is situated typically at 200-400 mm from the pilot's eye. This distance, as we will show after, limits the field-of-view (FOV) of the system. Now HUDs have become standard equipment on many civilian airliners around the world [5]. They have been also introduced to automotive industry since the end of the 80s [6, 7]. This increasing use and success were leading researchers to consider see-through NEDs (**Near-to-Eye Displays**) also known as HMDs (**Head Mounted Displays**) or HWDs (**Head-Worn Displays**). They are also collimated virtual displays, but they are placed directly in front of one or both eyes of the user. They can be mounted on some type of helmet or glasses [8, 9]. Another proposed approach for HWDs has been the retinal scanning display. In this system, the displayed image is scanned directly onto the viewer's retina with low-power blue, green and red lasers. The laser beams are modulated by acousto-optic modulators and then are coupled into fiber optics which

carry light to the helmet. Then the light reaches in free space propagation a pair of oscillating scanning mirrors which scan the viewer's retina. [10]. The extreme case for a NED can be viewed as a contact lens with built-in information display [11, 12]. In following text, we do not consider the retinal scanning display and contact lens with information display because the optical design of these systems is based on completely different approaches. As reference books on HMDs we can cite the book of Melzer [13] more dedicated to NED development for fixed-wing aircraft, Reference [14-16] focused on the use of NEDs in U.S. Army rotary-wing aircraft. Specific chapters have been dedicated to NEDs [15, 17, 18]. A paper has reviewed the NEDs for the Next Generation Air Transportation System [19]. The history of NEDs can be found in ref [20].

## 2.2. Different classes of see-through AR NEDs

According to the **ocularity** parameter, NEDs can be classified in monocular, biocular or binocular. In a monocular NED, a single image source is viewed by the sighting dominant eye (dominant monocular NED) or by the nondominant eye (nondominant monocular NED). Our dominant eye is the one we unconsciously choose when performing a monocular task. This effect has been recognized for centuries, but it remains poorly understood. Roughly 70% of the population shows a right-eye preference. This classification is based on the assumption that sighting dominance is a rigid preference. But it seems that it varies with different depth planes and with the position of objects relative to the eyes [21]. In a biocular NED, the same image source is viewed by both eyes. In a binocular NED, each eye views a different image source and with the right parameters, the viewer can see a 3D object [22]. The 3D object is plane for a fixed display distance but can be a volumetric 3D object with variable display planes.

Then another way to class NEDs is according to the display distance. The virtual

image can be displayed at a fixed distance in the far field of the viewer, at a finite distance useful for some manipulation e.g. typically 500mm (distance from the eye to the end of the hand), or the virtual image is displayed in various planes at different distances [23].

# 2.3. Key human factors considerations for NEDs

In this section we present shortly the optics of the eye and the three physiological processes: the accommodation, the vergence and the binocular parallax. Then the eyes movements and the mode of vision are described.

## 2.3.1. Optics of the eye

The instantaneous field-of-view of the human eye is roughly oval and typically measures 130° (V: Vertically) by 160° (H: horizontally) as shown in **Fig. 2.2**. The overall binocular FOV measures about 120°(V) by 200°(H). The eye is not a centered optical system. We can define rather two axes: the pupillary axis and the line of sight which can be considered approximately as the optical axis and the visual axis of the eye, respectively. The pupillary axis is the line perpendicular to the cornea that intersects the center of the entrance pupil. The line of sight connects the fixated object point to the center of the entrance pupil and the center of the fovea (see **Fig. 2.3**) [24]. The fovea is the retina area which provides the highest resolution. For instance, the normal resolution in the fovea is 1 minute of arc. As the incidence light angle increases the off-axis aberrations increase and the ability of peripheral retina to discriminate small objects decreases. And for example the resolutions at 10°, 20° or 30° of eccentricity are 2.5, 5, and 10 minutes of arc, respectively [25]. The average eye has an angle kappa of around 4° horizontally in the temporal direction and is centered vertically as shown in **Fig. 2.3**, where $\kappa$ is the angular distance (in the object space) between the line of sight and the pupillary axis.

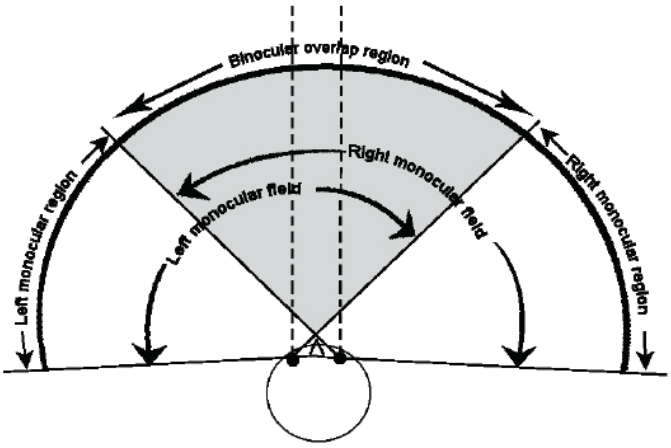



**Fig. 2.2.** Human visual system's binocular field-of-view [26].

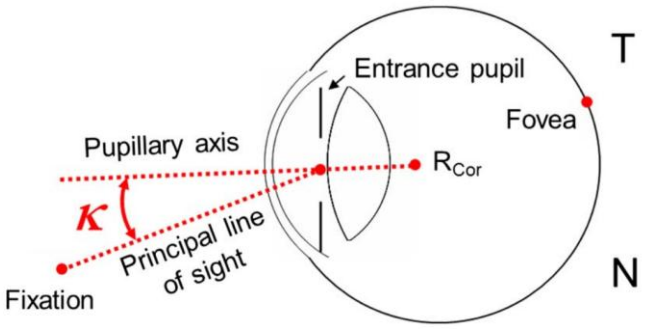



**Fig. 2.3.** The schematic representation of the main axes of the human's eye with the definition of the kappa angle – T: temporal and N: nasal [24].

In order to recognize 3D shapes, the eye and the brain use three physiological processes: the accommodation, the vergence and the binocular parallax.

Accommodation and vergence are cross-coupled in the neural control system and they interact dynamically. The accommodation is the ability of the eyes to focus on objects that are at different distances. The ciliary muscles change the refractive power of the crystalline lens by adjusting its shape in order to give from an object nearer than infinity, a sharp retinal image. For an object at the well-focused distance, the retinal image is the sharpest with minimum blur. As the depth of the object from the focused distance increases, the retinal image is increasingly blurred. In real world, due to the blur effects, accommodation and defocus blur are effective informative depth cue for relatively near objects typically less than two meters away. In vergence process, the two eyes rotate in opposite directions to maintain binocular fixation on objects at different distances. Inaccurate vergence leads to diploplia (double images). In stereoscopic NED

designs, the virtual image is focused at a fixed distance from the eyes while the depth of the virtual objects, and hence the binocular disparity, varies with the content, which results in conflicting information within the vergence-accommodation feedback loops which is called vergence-accommodation conflict (VAC) as shown in **Fig. 2.4** [27]. Virtual display plane, or focal plane, is located at a fixed distance. The virtual objects can be located either in front or, if it is not at infinity, behind it. Thus, the disparity cue drives the eyes to verge at one distance, while the light rays coming from the virtual plane produce retinal blur, which drives the eyes to accommodate to another distance, giving rise to the conflict between these depth cues. **Fig. 2.5** summaries for one eye, the different geometrical and optical quantities as the eye relief distance and field-of-view. Subscripts *e*, *i*, *l*, and *s* represent "eye", "(virtual) image", "lens", and "screen" respectively, so terms such as $d_{il}$ explicitly denote "distance from image to lens", $w_l$ denotes "width of lens"; $f$ is focal length; $t$ is relevant for spatially-multiplexed MFP designs and represents the thickness of the display stack; $M$ is the magnification factor from the physical to the virtual image. To allow viewing of the entire image, FOV must fit within the angular viewing range constrained by $w_l$ and lateral offset of the pupil, which dictates the width of the eye box ($w_e$).

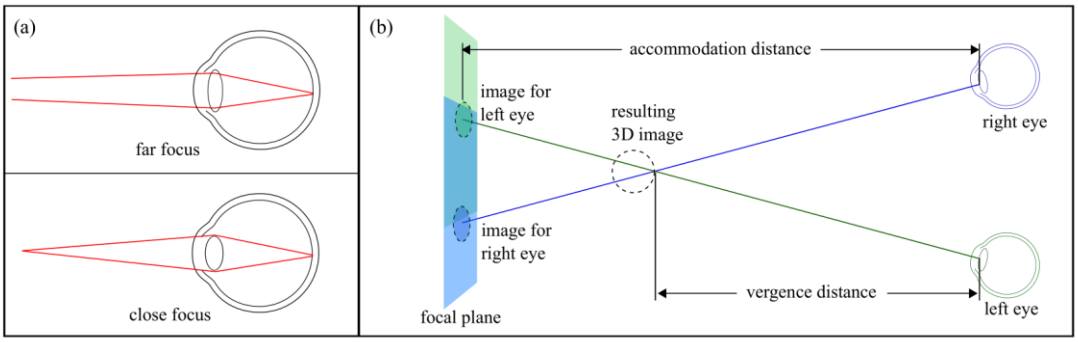



**Fig. 2.4.** Principle of the stereoscopic view and the vergence accommodation conflict. (a) Conceptual representation of accommodation within the same eye. (b) Conceptual representation of the VAC [27].

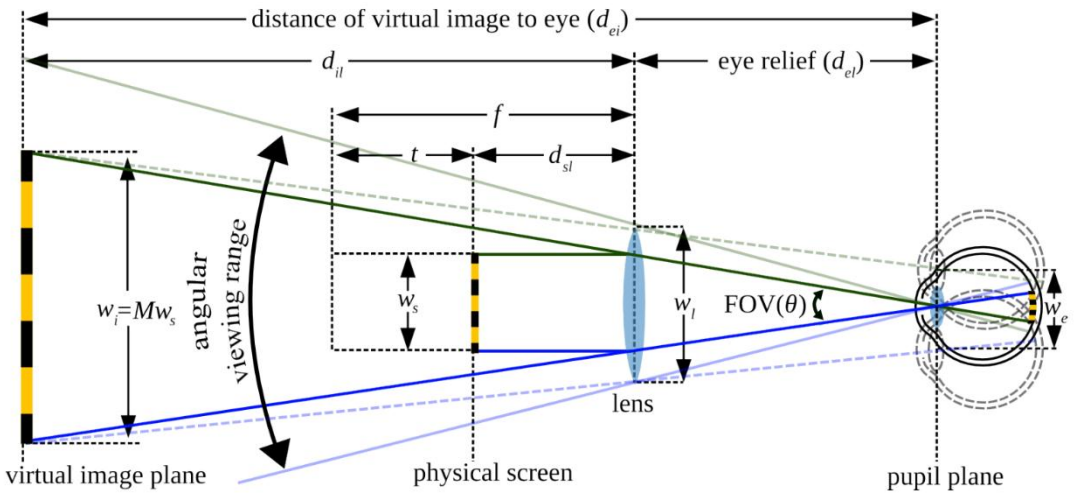



**Fig. 2.5.** Principal geometrical and optical quantities for one eye [26].

The Field-Of-View (FOV) for a HMD is the apparent angular size of the virtual image seen by the viewer through the HMD. A monocular FOV is what is viewed by one eye and a binocular FOV is the total angular size seen by the both eyes. In a NED, FOV and display resolution are linked by the focal length of the collimating optics. If we consider a given image source with a given number of pixels, a larger FOV will magnify the pixels more and then reduce the resolution that can be expressed in pixel per degree. There are three ways to specify FOV: (a) Horizontal FOV × Vertical FOV, (b) Diagonal FOV (DFOV) and (c) Total binocular FOV as shown in Fig. 2.6 [28].

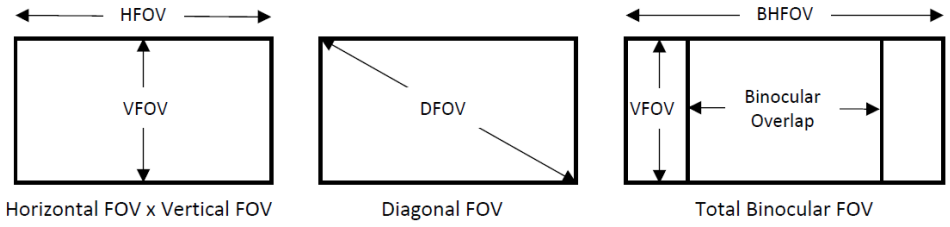



Fig. 2.6. Three ways of specifying FOV [28].

## 2.3.2. Eyes movements

Human moves his eyes about three times each second via unconscious rapid eye movements (saccades) to reorient the fovea through the scene. 90% of viewing time is spent with the eyes apparently stationary, in fixations. These fixations have a duration of about 150-600 ms [29]. Human gaze is a combination of body position, plus head

and eye orientation. Total oculomotor range of eye movement in combination with head movement is ±55°. Without head movement, 90% of all the saccadic eye movements are within ±15°. Then a user who is forced to exceed the ±15° range may suffer eye strain and possibly miss a key piece of information [30].

Human balance ability is based on a combination of inputs from the proprioceptive, vestibular and visual systems. The proprioceptive system gives consciously or not information about the position of organs using the musculature, tendons and skeleton. The vestibular sensory organs are contained in the inner ear. The semicircular canals detect head rotational movements and the otoliths detect linear head acceleration. Then the vestibular system acts as a 6 degree of freedom motion sensor. The vestibulo-ocular reflex permits best vision during head motion by moving the eyes counter to head movement to stabilize the line of sight in space. This reflex allows the vision to lock onto an object even in presence of head motion or vibration. Vision contributes to our sense of balance and spatial orientation.

However, conflicting proprioceptive, vestibular and visual cues can result in spatial disorientation. Nevertheless, visual cues are capable to overriding conflicting inputs from vestibular cues, and wider fields of view contribute to a more stable posture and a better sense of balance. But display latency or other display artifacts can perturb the stabilization obtained by the vestibule-ocular reflex and create sickness (called cybersickness) with physical effects similar to motion sickness.

### 2.3.3. The mode of vision

The human visual system can be divided into two modes: the focal and the ambient modes. The focal mode is activated by the stimulation in the central portion of our vision. It is sensitive to light level, has a high resolution and rather answers to the question of “What?” about our environment. The ambient mode is generally found in the periphery of our vision. It is less sensitive to light level, has a lower resolution, is more susceptible to temporal effects such as movement and flicker and rather answers the question of “Where?” about our environment. Then for tasks requiring high

frequencies information, the focal vision should be activated by displaying NED or HMD imagery in the center of vision. If spatial orientation and balance are required for tasks in the real or virtual world, in that case, the ambient vision should be activated with a wide field of view imagery.

After a short presentation of key human factors that must be considered for NED applications, we detail different designs of NEDs that have been proposed in the scientific literature.

# 2.4. Review of NED optical designs

Morton Leonard Heilig was a pioneer in Head mounted displays(HMDs). He patented the first concept in virtual reality in 1960 [31]. In 1968, Ivan Sutherland created the first HMD system named The Sword of Damocles, with the help of his student Bob Sproull [32]. Since then, many kinds of Near to eye displays (NEDs) have emerged. Recently, Google started selling a prototype of Google Glass called "Glass Explorers" to the public in 2014, which makes the concept of augmented reality(AR) glasses popular. In the same year, Oculus released its pre-production model Oculus DK2, which aroused great interest for the virtual reality (VR) devices. In addition to these two products, many companies have their own designs. In this chapter, we will mainly focus on AR see-through near to eye displays (STNEDs). According to their structures, we separate all kinds of STNEDs into three families: free space combiner systems, freeform prisms and waveguide systems.

## 2.4.1. NEDs with free space combiner

In this section, we will introduce the NEDs with free space combiner. The combiner can be a flat mirror, curved mirror and diffractive optical element. The whole system can be a catadioptric optical system where refraction and reflection are combined in an optical system, usually with lenses (dioptrics) and curved mirrors (catoptrics) or purely reflective.

### 2.4.1.1. 45 degrees Flat combiner

This type is the most straightforward way to achieve see-through ability. The light from image source and the ambient are combined by the semi-reflective mirror as shown in **Fig. 2.7**. The size of the semi-reflective mirror and lenses increase as the FOV increases. The diameter of the first lens is:

$$D = 2l\tan(\frac{FOV}{2}) \quad (2\text{-}1)$$

where $l$ is the distance to from the eye to the lens. Because of this relation, the system becomes very bulky when FOV is large. As an example, with FOV=30° and $l$=50 mm, the diameter of the first lens is 27 mm.

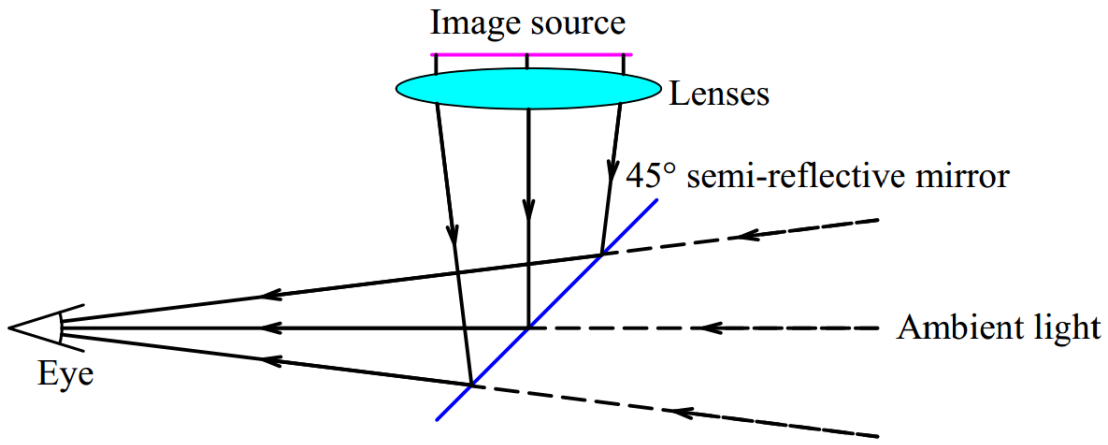



**Fig. 2.7.** The basic structure of 45 degrees flat combiner.

Alternative designs of this type are shown in **Fig. 2.8** (a) from Kessler optics and **Fig. 2.8**(b) which is the optical system of Google glasses [33]. In Kessler's system, a curved mirror is added after the 45° planar mirror. In a sense, it may be classified into curved combiner as we will introduce in the next section. We still classify it into 45° planar combiner family because all the components are on-axis. The curved mirror has two functions. The first is to correct aberrations and the second is to reduce the size of the lenses. Even though the size of the curved mirror is also high, but normally mirrors are lighter than lenses with the same diameter. In the Google's system, a curved mirror is also used, the difference is that the medium between the components is glass instead of air. This increases the weight of the system, the benefit of doing this is an easier

assembly. The drawbacks of adding mirrors is the energy reduction to less than 25% of the original energy. To improve the energy efficiency, a polarized beam splitter (PBS) has been used [34].

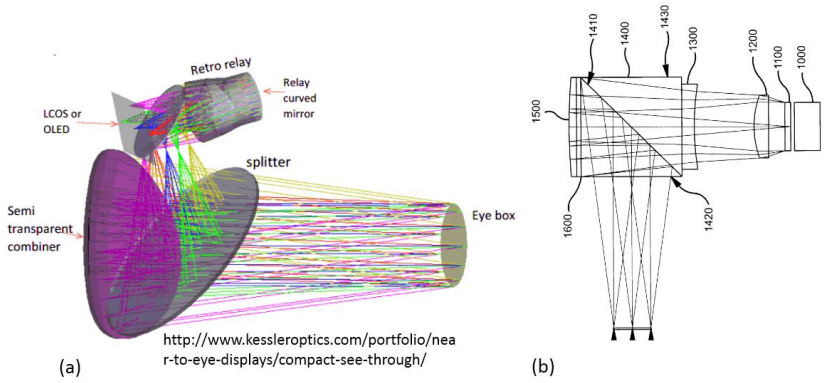



**Fig. 2.8.** Alternative designs of 45 degrees flat combiner. (a) Kessler optics system and (b) Google glass [33].

Polarized head-mounted projection display (p-HMPD) [35] can also be classified into this family. However, they can only be partially see-through. They involve a retro-reflector screen as shown in **Fig. 2.9**. The quarter-wave retarder and PBS are combined to increase the energy efficiency. The advantage of this structure is that a very small projection lens can be applied directly. The DFOV is around 50°-60°. However, the image quality of the retro-reflective screen is unsatisfactory under the current technical conditions.

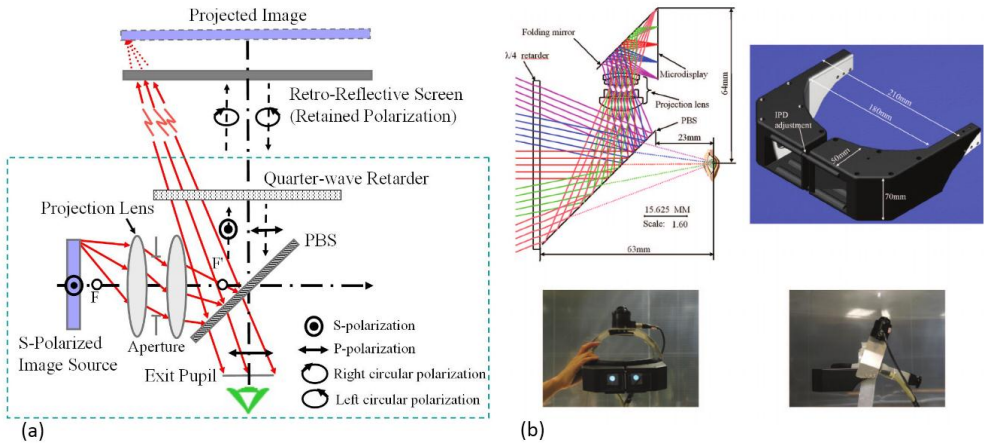



**Fig. 2.9.** (a) Schematic of a monocular head-mounted projection display and (b) Design of a compact p-HMHD prototype [35].

The 45 degrees Flat combiners have the advantage of easy fabrication of both the

combiner itself and the lenses due to its on-axis and rotationally symmetry. However, the size of the lens increases rapidly with the increase of the FOV. Generally, the FOV, eye box and the eye relief are small in this kind of system.

### 2.4.1.2. Curved combiners with catadioptric system

Applying curved combiners can improve the lighting budget of the system, and eliminate the ghost image. This kind of configurations provide the highest theoretical luminance transfer from the display with the highest see-through vision and increased eye clearances for a given FOV [36]. **Fig. 2.10** shows the basic structure of the NEDs using curved combiner [37]. The whole system is an off-axis system. Thus, the main difficulties are to correct the off-axis aberrations. Normally, complex relay lenses are added. Freeform surfaces are suitable in this kind of system. However, as the difficulties and price of fabricating freeform lenses are high, the off-axis aberrations can also be well corrected by tilting and decentering several rotationally symmetric (spherical or aspherical) lenses. This phenomena can be explained by nodal aberration theory [38].

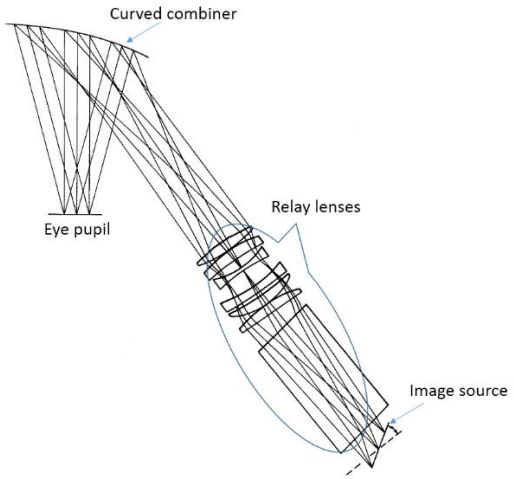



**Fig. 2.10.** Basic configuration of the NEDs using curved combiner [37].

Very good review in this family can be found in ref [9, 39]. Here we mainly focus on the performance and design details of this kind of system. **Fig. 2.11** shows an off-axis design and its image quality and distortion [40]. In this design, the tilt of the combiner is 29.5° from the user's line of sight. Astigmatism is the dominant image-degrading

aberration that is introduced by the off-axis configuration. The combiner in the system has a toroidal shape rather than spherical, to minimize third-order astigmatism. The relay lens contains two prisms to fold the optical system closely around the head of the HMD wearer, moreover, three non-spherical lenses are used: a cylindrical lens located in the second lens after the first prism and two parabolic lenses close to the second prism. To have the maximum image quality, the distortion is unconstrained as shown in **Fig. 2.11**(b). What we can notice is that the tilted combiner also generates large distortion. **Fig. 2.11** (c) and (d) show the MTF for a 10 mm pupil size in different fields.

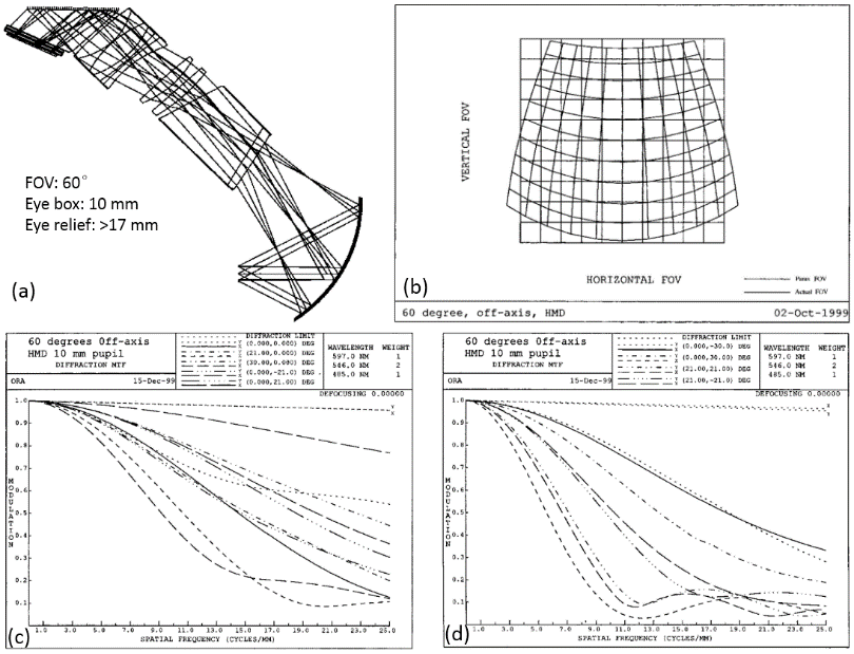



**Fig. 2.11.** (a) Layout of the off-axis, 60-deg FOV HMD, (b) Distortion, (c) and (d) Plots of the polychromatic MTF for a 10-mm pupil size and various points in the FOV [40].

Another example is shown in **Fig. 2.12** [41]. The combiner (labeled 1) is a tilted spherical surface. Two flat mirrors (labeled 2) are used to fold the system snugly around the head of the user. Relay lenses (labeled 3) are used to bring forth the view image for the user and to compensate for the off-axis aberration between the image source and the combiner. The total of six relay lenses of the system consist of a positive meniscus lens, a biconvex lens, a doublet lens, a biconvex lens, and a positive meniscus lens, arranged sequentially in order from the combinerto the image source. The first lens, which is labeled 3′-lens1, has positive power, which can make it possible to reduce the diameter and thickness of the system. Lens1, consisting of an x-y polynomial surface

and a concave spherical surface, was bent to the combiner to reduce spherical aberration. Even order items of x in x-y polynomials were kept reducing the astigmatism aberration and asymmetric field curvature introduced by the tilted spherical combiner. Aspheric lens2 contributes mainly to limit coma. A cemented doublet was applied to the system to provide correction for distortion, axial color, and lateral color aberration. Lens3 with an aspheric surface has a 6.4° tilt to balance the residual astigmatism. Lens4 is a meniscus lens with positive power to limit distortion.

A comparison of MTFs when lens 1 contains aspherical surface or x-y polynomial is shown in **Fig. 2.12** (c) and (d) respectively. With aspherical surface, 15% at a spatial frequency of 20 line pairs/mm across the full field is achieved while 30% at the spatial frequency of 30 line pairs/mm is achieved using the x-y polynomial surface. The maximum distortion is 10% at the top right corner of the field. The shape of the distortion is very like the first example [40].

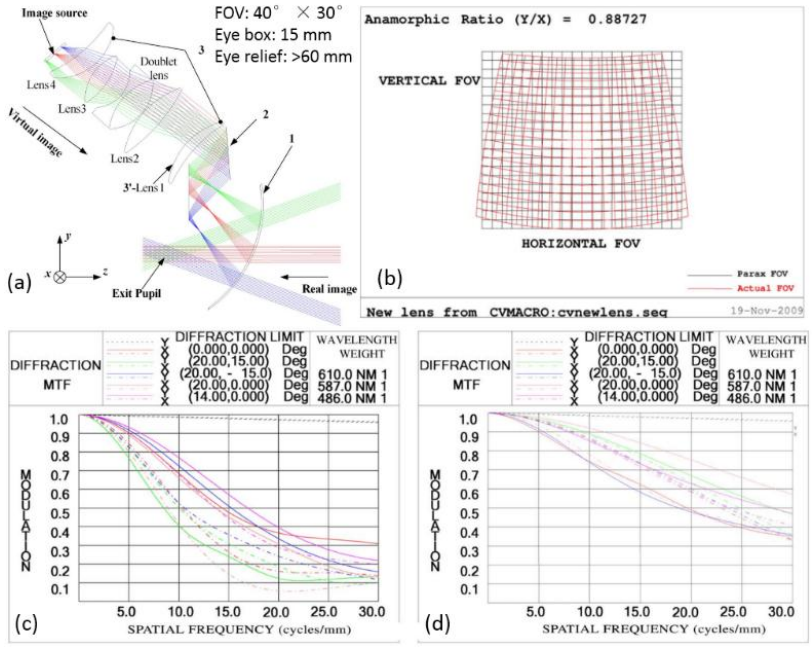



**Fig. 2.12.** (a) Layout of the off-axis see-through HMD optical system, (b) Distortion, (c) Plots of the polychromatic MTF when lens 1 contains aspherical and (d) Plots of the polychromatic MTF when lens 1 contains x-y polynomial surface [41].

The off-axis aberrations are the main difficulties to overcome and distortion is more serious in this kind of systems. The main source of the aberrations and distortion is the tilted combiner. To reduce the influence of the combiner, we can put a compensator

which is similar to the combiner. Yang et. al designed a system according to this concept as shown in **Fig. 2.13** [42].

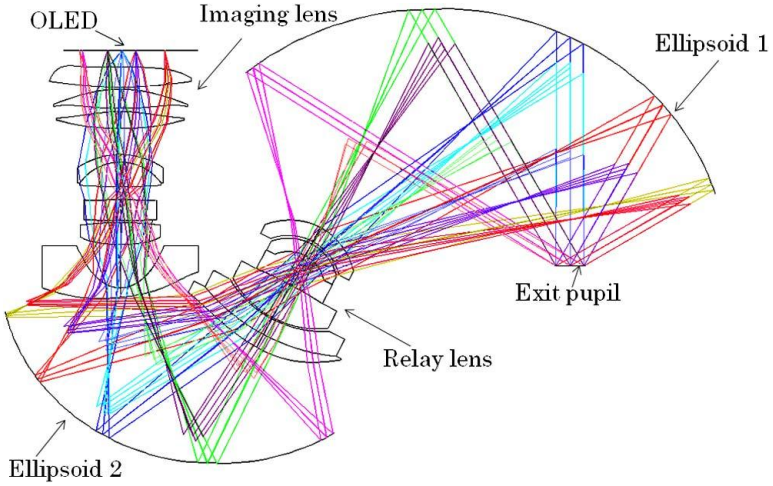



**Fig. 2.13.** Layout of the HMD with two similar ellipsoids [42].

The principle of correcting the distortion is shown in **Fig. 2.14**. **Fig. 2.14** (b) presents the minimum distortion when an angle exists between an image and object planes (known as the Scheimpflug condition) [43]. Such distortion is difficult to correct by simply placing refractive lenses around the second focus. To solve this problem, another ellipsoid with the same conic constant is used to offset the distortion caused by the first ellipsoid [see **Fig. 2.14**(c)]. The right focus of the left ellipsoid overlaps with the left focus of the right ellipsoid. Light emitted from focus A converge to focus B. Rays then converges to focus C because focus B is also a focus of the left ellipsoid. The two ellipsoids have the same conic constant. Thus, $\Delta ADB \approx \Delta CEB$, $CE \parallel DA$. In addition, an arbitrary angle $\alpha$ at focus A is equal to the corresponding angle $\beta$ at focus C (i.e., $\alpha = \beta$). Therefore, the image at FG forms a distortionless image at HL if no other lens exists. The grid distortion is shown in **Fig. 2.14**(d).

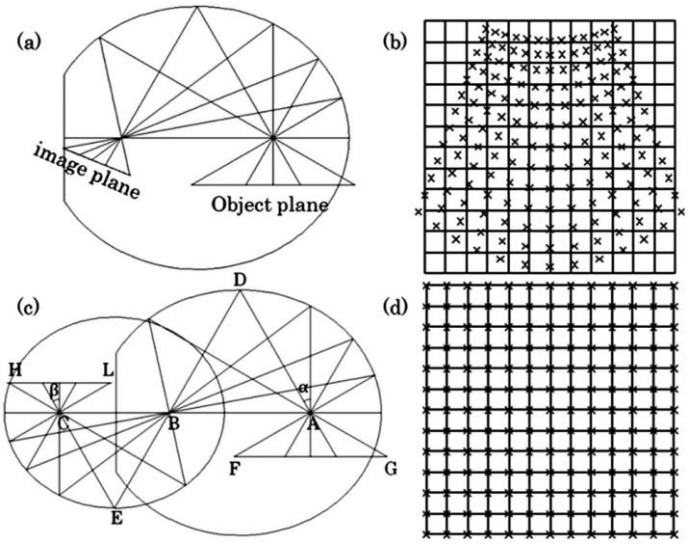



**Fig. 2.14.** (a) Chief rays of just one ellipsoid. (b) Minimum distortion when an angle exists between an image plane and an object plane. (c) Structure of two ellipsoids. (d) Free grid distortion [42].

This system can achieve a single channel field-of-view (FOV) of 120° × 120° with a 6 mm diameter eye box and 49 mm eye relief on the pupil axis. However, the system is complex because of the large FOV. Using the same concept, it can be simplified with reduced FOV [44]. The lay-out of the system is shown in **Fig. 2.15**. The author applies a freeform reflective mirror (FFRM) in the co-focal point of the two ellipsoids. It achieves a FOV of 40°×30°(H×V) and an eye relief of 60 mm. In addition, the average illuminations are higher than 80%.

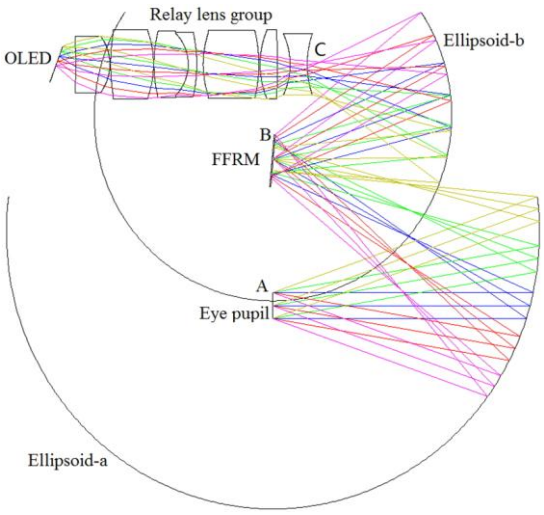



**Fig. 2.15.** Structure of the HMD with freeform reflective mirror(FFRM) based on the two similar ellipsoids. It consists of two ellipsoid surfaces, an FFRM, a 7-pieces co-axis relay lens group, and an OLED [44].

The primary advantage of the Catadioptric system design is that the see-through quality is the best for the moment. It also provides the highest theoretical luminance transfer from the display. The key drawbacks are very complex optical designs, bulky structures, difficulties to assemble and to hold steady the reflective surfaces. In the next section, we will introduce the NEDs using only reflective surfaces.

### 2.4.1.3. Design with reflective surfaces only

Systems built up only with reflective elements present no chromatic aberration. Such a design [45] is shown in **Fig. 2.16**(a). Immy Inc. has used this approach and their schematic diagram and product are shown in **Fig. 2.16** (b) and (c). However, in this kind of systems, off-axis geometrical aberrations are still serious. The FOV is limited when only three mirrors are used. Higher optical performances can be obtained by increasing the number of mirrors. However, this will increase the size of the whole system. At the same time, the optical alignment and assembly become more difficult.

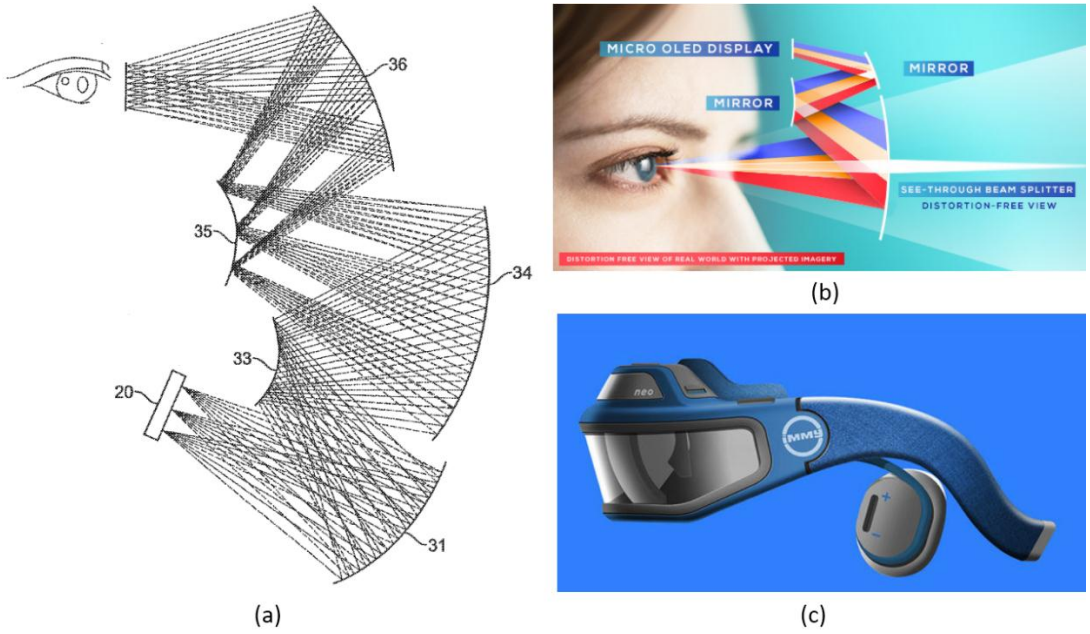



**Fig. 2.16.** (a)Optical layout of the reflective system (b) schematic diagram of Immy Inc. product and (c) a picture of their product [45].

Inside the all reflective system, there is one family that uses only one mirror. This system has been developed recently, in relation with the application of large display

panel.

### 2.4.1.4. One element system

Previously, most of the designs were based on micro displays. Getting larger FOV means higher magnification and difficulties in optical design. Applying large screen such as mobile phone instead opens a new path to NEDs designer. The magnification with the same optical performance is reduced several times when using large screen. This trend is more obvious in VR after Oculus released their products. However, using large screen in AR system is more difficult than in VR because of the off-axis configuration in order to have see-through capability. Anyway, the AR products applying mobile phone have been realized by several companies such as Seer, Seebright and Meta. **Fig. 2.17** shows an interesting comparison (from Seer) between NEDs with micro display and mobile phone panels.

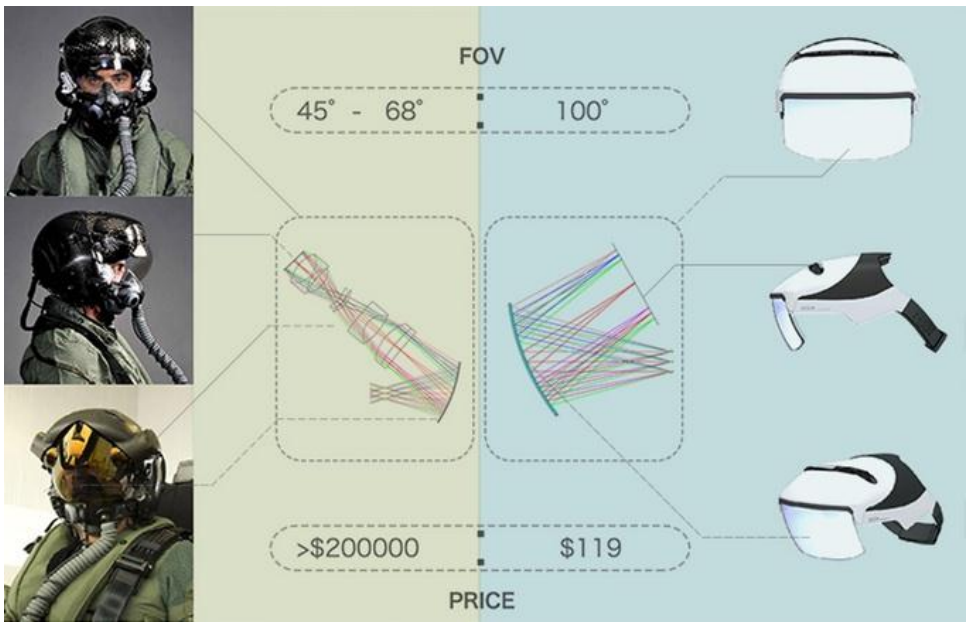



**Fig. 2.17.** Comparison between NEDs with micro display and mobile phone panels [Seer].

The design using large screen reaches a FOV of 100° while the maximum FOV of traditional see-through NEDs are less than 70°. However, the image quality applying large screen is unsatisfactory. It has small eye box and large distortion [46]. **Fig. 2.18** shows an optical design of Meta and the image produced by the system. This design

also belongs to catadioptric system as both sides of the combiner are used. The inner surface which is close to the eye refracts light twice while the outer surface serves as a mirror.

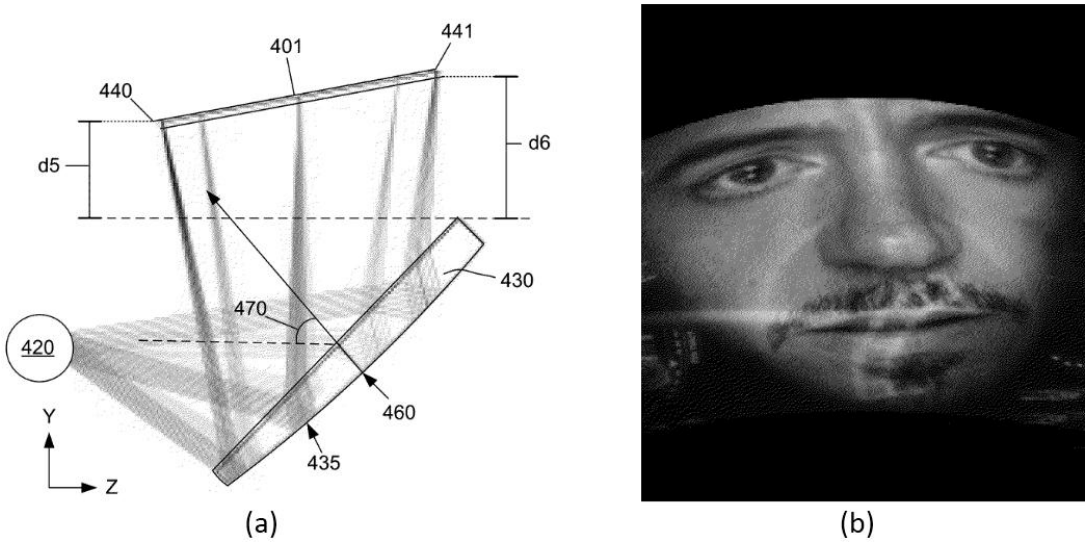



**Fig. 2.18.** (a) Optical layout of design using large screen. (b) An example of images produced by the optical system [46].

Products of Meta, Seebright and Seer companies using mobile phone panel are illustrated in **Fig. 2.19**. Because of the simple structure, the price of the optical kit is largely reduced comparing to conventional off-axis products. However, using large screen reduces the complexity but not the size. Micro image sources are still the right choice towards a compact see-through eyeglass with high optical performances.


**Fig. 2.19.** Products of Meta, Seebright and Seer [Meta, Seebright, Seer].

Design of systems using only one mirror and micro display have been developed [47]. They can achieve a 3-mm diameter pupil size, an eye relief distance greater than 15mm and a 24° diagonal FOV.

All the designs we introduced above use traditional mirrors as combiner. Next, we introduce another type that add diffractive optical element on the combiner.

### 2.4.1.5. Diffractive free space combiner

The aberrations caused by the tilted combiner make it difficult to realize a lightweight and compact system. Diffractive optical elements (DOEs) have been widely developed in optics. Compared to conventional optical elements, DOEs have advantages of thinness (in scale of micrometers) and lightness. Due to these reasons, DOEs have been extensively used in NEDs especially in waveguide NEDs described later in this chapter. Here, we introduce HMDs using DOEs in free space combiner family.

A DOE is a monochromatic device, i.e. its otpical function is well defined for the design wavelength. To provide full color, DOEs must be dynamically reconfigured at a speed faster than the human eye can detect. Thus, switchable holograms enable a full color display. The first example is a design from Digilens (DL40) as show in **Fig. 2.20** [48-50], where a system with specific integrated lenses is placed in the end of the relay lens. These elements are based on a holographic polymer dispersed liquid crystal (H-PDLC) material which offers high efficiency, short switching time and low switching voltage. The combiner is composed of a passive reflection DOE recorded in a DuPont photopolymer material. The system has a FOV of 30° × 40°. The overall weight is only 100 g. The exit pupil is 5 mm and the eye relief is 29.1 mm. The switchable transmission holograms are aberration-corrected re-imaging lenses with efficiency >70% with film thickness of only 5 μm. However, the fabrication can be complex and the frequency response may be not sufficiently high for the moment.

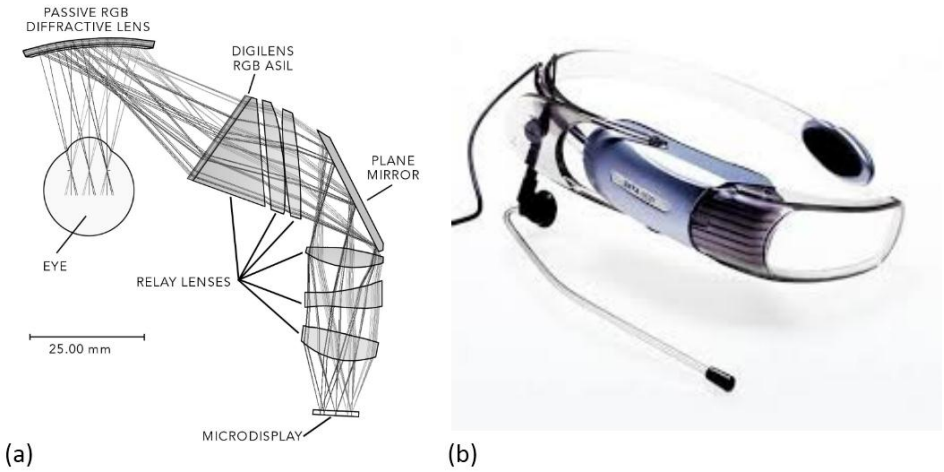



**Fig. 2.20.** System contains DOE. (a) Lay out of the optical system, (c) The products of Digilens (DL40) [48].

Using a dynamic DOE to compensate the aberrations caused by another static DOE point by point is another example of using DOE as free space combiner(**Fig. 2.21**(a)) [51]. The aberrations are corrected point by point over a 10°x15° FOV. An incident monochromatic plane wave $\Phi_{C1}$ of wavelength $\lambda_0$ is diffracted, at the first order of diffraction, and modulated by a dynamic DOE labeled 1 in order to generate a complex wave $\Phi$ D1 (**Fig. 2.21**(b)). This complex wave acts as a reading wave $\Phi_{C2}$ for a static DOE labeled 2. The direction of the optical axis of the system is modified at the surface of DOE 2. Its normal is at an angle $\alpha_1$ to the optical axis before its break and at an angle $\alpha_2$ after its break. The reading wave $\Phi_{C2}$ is backward-diffracted, at the first order of diffraction, by the DOE 2 into a diverging wave $\Phi_{D2}$ until the eye of the user. The user will thus see, through DOE 2, a virtual point I superposed into the real environment. According to the description, DOE 2 is an imaging element. It is a reflective analogical element having a frozen phase transfer function (PTF) $\Phi_{PTF2}$. This PTF has a simple form to ease the fabrication. The position and the aberrations of the image point I, especially coma and astigmatism, depend on the reading wave $\Phi_{C2}$. DOE 1 is a dynamic computer-generated diffractive optical element. It is used to generate at the first order of diffraction a complex wave front that compensates for the aberrations. Its PTF is dynamic. It is optimized for each position of the image point I to minimize aberrations. Because of the limitation of the real-time refresh rate of nowadays display devices, the

system cannot support high resolution image for the moment.

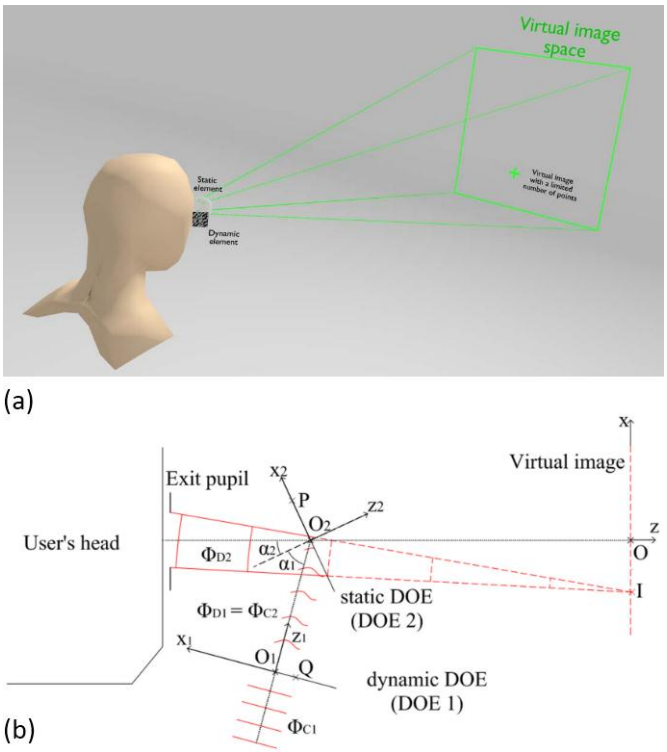



**Fig. 2.21.** (a) An illustration of the setup and (b) Top view of the optical system (not to scale). Q is a point of the plane of DOE 1, P is point of the plane of DOE 2 and I is a point of the plane of the virtual image. $\alpha_1$ is the angle between the normal to DOE 2 and the direction ($O_1O_2$). $\alpha_2$ is the angle between the normal to DOE 2 and the direction ($OO_2$) [51].

In conclusion, NEDs with free space combiner can provide the best see-through capability. It is also a type that has longest history. Various designs can be found in literature. Very large FOV can be achieved in some designs. The key drawbacks are complexity, bulkiness, difficulties to assemble. In the next section, we introduce another type that involves freeform prisms.

## 2.4.2. Freeform prism NEDs

Shoichi Yamazaki invented the freeform prism in 1995 in Canon [52-54]. This element has been successfully implemented into real products, for instance in MREAL HM-A1 (Canon), NED-X2 (NED$^+$ Glass), and Golden-I (Kopin) (see **Fig. 2.22**(a), (b) and (c) respectively.


**Fig. 2.22.** Examples of products with freeform prism (a) MREAL HM-A1 (Canon), (b) NED-X2 (NED$^{+}$ Glass) and (c) Golden-I (Kopin).

The optical principle of the use of freeform prism in NEDs is shown in **Fig. 2.23**. A wedge-shaped freeform prism is cemented with another freeform prism. The first prism serves as the near-eye viewing optics that magnifies the image generated by a micro display, while the cemented prism maintains a non-distorted see-through view. Light is firstly refracted by surface AB and then totally reflected by surface AC. Surface BC is a semi-reflective surface which reflects the rays from AC. Finally, the rays are transmitted through the surface AC again. We should notice that light hits the surface AC two times: one in total internal reflection mode while the other is in transmission mode. The cemented freeform prism is used to counteract the ray shift and distortion caused by the first prism. [55].

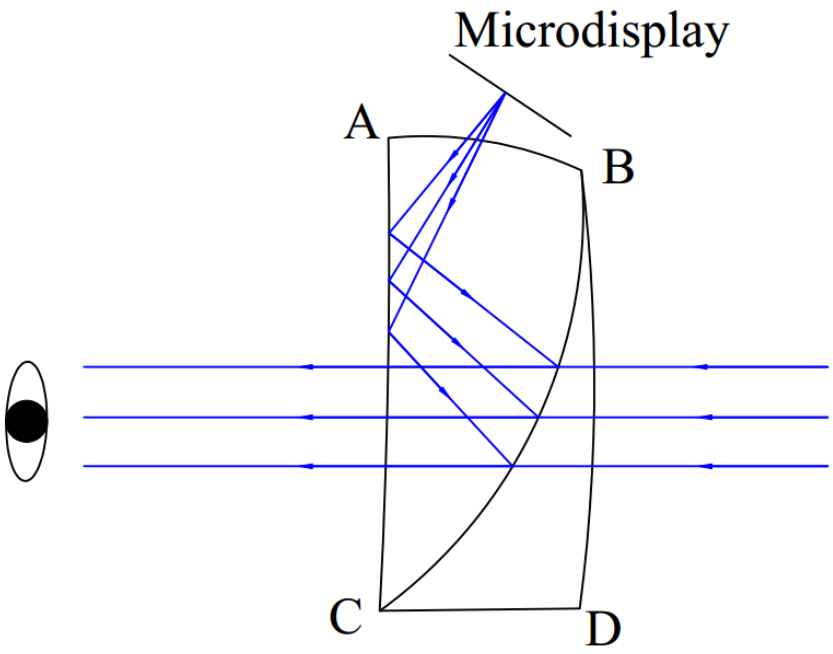



**Fig. 2.23.** Layout of the see-through freeform prism system.

As an example, the ray tracing, prototype, and optical performance of the design in ref [55] are shown in **Fig. 2.24**. The system has a FOV of 45°H × 32°V, 8 mm eye box and >17 mm eye relief.

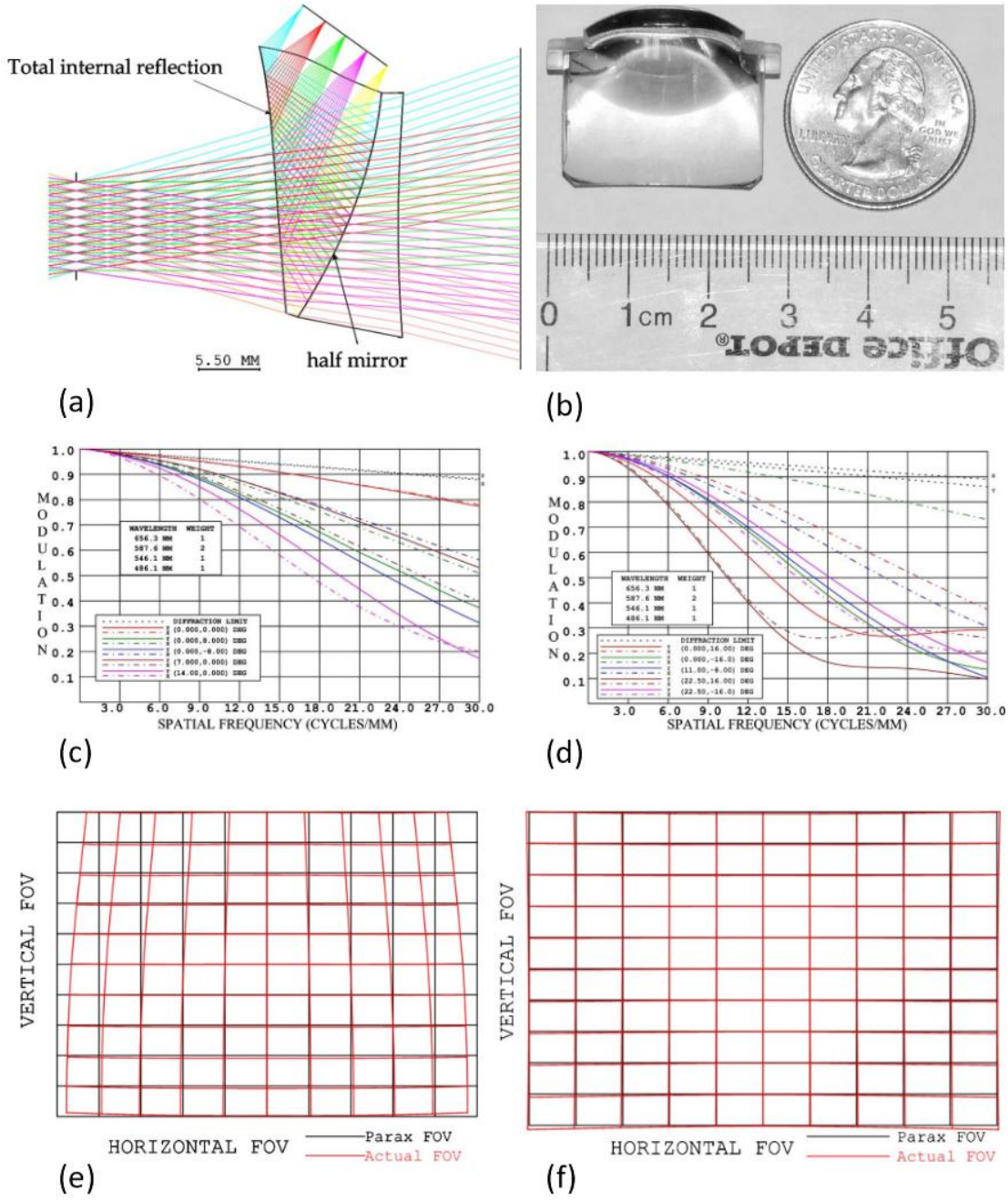



**Fig. 2.24.** (a) Layout of the Freeform prism lens system (b) Module of the Freeform prism, (c) Polychromatic MTF plot of center field of the virtual imaging system, (d) polychromatic MTF plot of marginal fields of the virtual imaging system, (e) Distortion plots of the virtual imaging optical system and (d) Distortion plot of the optical see-through system [55].

Normally, the thickness of the freeform prism with see-through capability is above 10 mm. Alternative designs that may reduce the thickness are shown in **Fig. 2.25** [56,

57]. The associated concept is to use a waveguide prism with multiple reflections. The DFOV in this kind of design is around 30°. Although this kind of design is more ergonomic, the optical performance is hardly improved compared to the ordinary freeform prism. To increase FOV, example of tiling several freeform prisms can be found in ref [58].

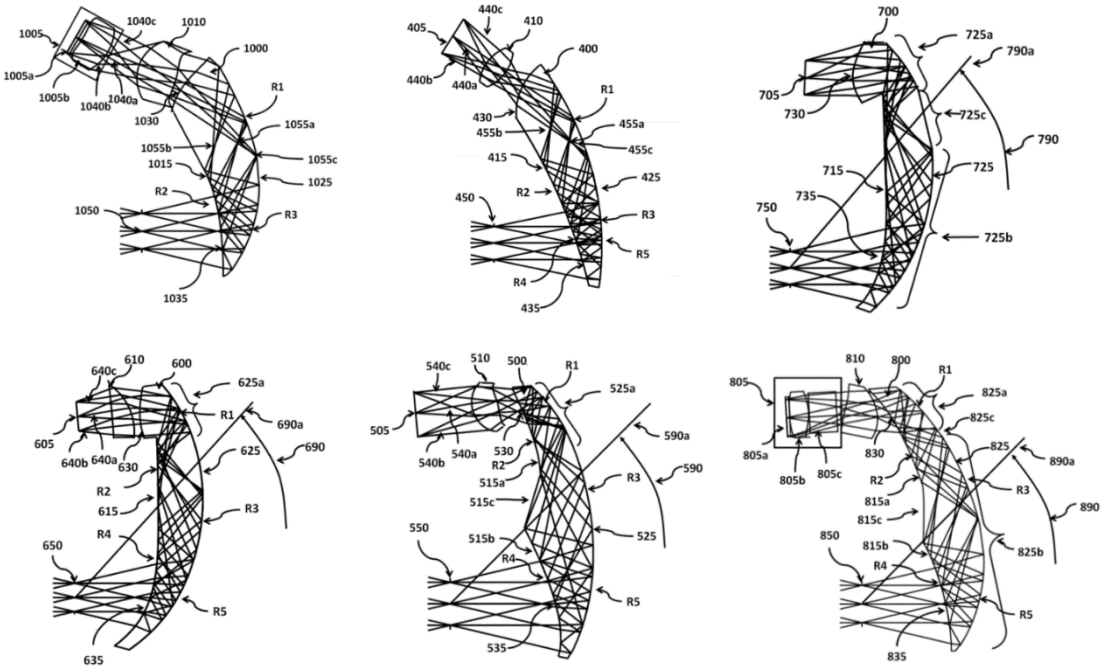



**Fig. 2.25.** Ergonomic designs using multiple reflection [56].

Freeform prisms do provide a compact design for NEDs. However, until now, the performance is limited to a DFOV about 50° and an eye box less than 8 mm. To have see-through capability, a compensator is needed. The final thickness is around 10 mm, which gives cumbersome devices. Both the NEDs using free space combiner or freeform prism have no function of expanding pupil size. In the next section, we introduce a type that can expand pupil size—waveguide NEDs.

## 2.4.3. Waveguide NEDs

The concept of using waveguide for NEDs was firstly proposed by Y. Amitai [59]. More previously, substrate-guided method was proposed in head-up displays (HUDs) [60, 61]. The principal optical advantage for waveguide is its exit pupil expanding

capability. Due to the expanding, large eye box can be obtained without adding additional difficulties to imaging systems. Furthermore, comparing to free space combiner designs, off-axis aberrations can be avoided due to the total internal reflection (TIR) mode. If the 45 degrees flat combiner is inside a substrate, it is very like a waveguide type. The TIR of the waveguide makes the substrate thinner than the 45 degrees flat combiner design size. Therefore, this type of NEDs is very promising. The basic optical principle is shown in **Fig. 2.26**. Generally speaking, the system contains four parts: the projection subsystem, the in-coupling subsystem (called also sometimes in-coupler), the waveguide and the out-coupling subsystem (called also out-coupler). The projection system is composed of the image source and lenses. Light is firstly emitted from the source and then collimated by the projection lenses. The collimated beam is coupled inside the waveguide by the in-coupler. Light propagates inside the waveguide by TIR until it reaches the out-coupler where the light is redirected to the viewer's eye out of the waveguide. The projection system can be designed in the right position according to the eyes by adjusting the length of the waveguide due to the TIR inside the plate waveguide which cannot change the property of light. Hence, waveguide type of NEDs can be very ergonomic.

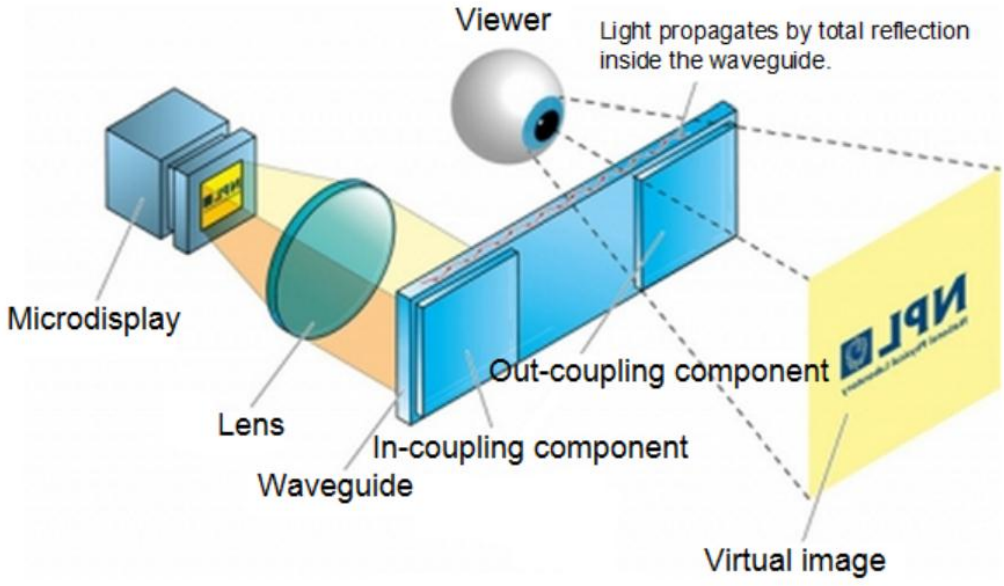



**Fig. 2.26.** Principle of waveguide type designs for NEDs. (from National Physical Laboratory)

The most challenging part of the waveguide NEDs is the out-coupler. The in-coupler may have the same structure as the out-coupler or simply can be a prism or a mirror. According to the type of chosen out-coupler, waveguide NEDs can be classified into four types: diffractive waveguide, holographic waveguide, selective thin film array waveguide and reflective waveguide [62]. The exit pupil expander (EPE) can be one dimensional or two dimensional. Next, we present separately the different existing types.

### 2.4.3.1. Diffractive waveguide

The diffractive waveguide was developed from a laser-beam expander previously proposed in [63]. One-dimensional EPE has been patented in 2001 [64] as well as a two-dimensional EPE in 2003 [65]. Normally, diffractive waveguide refers to systems that use surface relief gratings as in-coupler and out-coupler. If the periods of the out-coupling and in-coupling gratings are equal and their lines are parallel, the input and output beams are parallel. This principle serves as a base of using diffractive EPE. To ensure that the light emerging from the output DOE is uniformly distributed, the localized diffraction efficiencies along the expanded direction must vary. It can be estimated from the formula:

$$\eta(x) = \frac{\eta_0}{1 - a\eta_0 x / S} \qquad (2\text{-}2)$$

where $S$ is the interval between the TIR and $\eta_0$ is the initial efficiency of the reflective mode. The constant a is approximately $a = 1 + \eta_t/\eta_r$, where $\eta_t$ and $\eta_r$ are the diffraction efficiencies of transmissive and reflective modes, respectively. According to Eq. (2-2), local diffraction efficiency, residual energy and diffracted output energy with respect to the propagation distance are shown in **Fig. 2.27** [65].

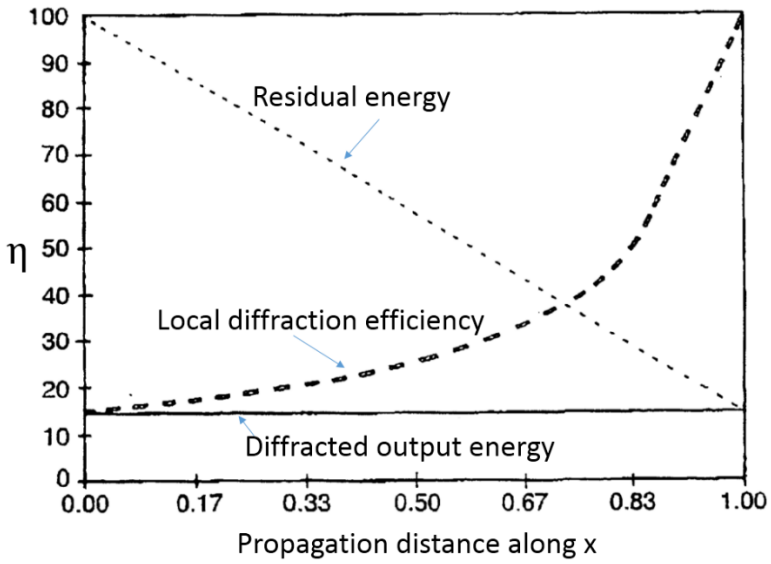



**Fig. 2.27.** Diffractive waveguide: local diffraction efficiency, residual energy and diffracted output energy with respect to the propagation distance [65].

Two-dimensional EPE is more interesting in waveguide NEDs as it can reduce the size of the imaging optics and support a larger vertical FOV. There are at least two different methods that can realize the 2-D expansion [66]: expansion based on odd number of first-order diffractions and expansion based on an even number of first-order diffractions as shown in **Fig. 2.28**. The angles between each grating line in these two cases are different. In the first configuration (**Fig. 2.28**(a)), if the rotation angle of the grating lines of DOE2 with respect to the grating lines of DOE1 is ρ, the angle between DOE3 and DOE1 must be 2ρ. The grating period of DOE2 is:

$$d' = d/(2\cos\rho) \tag{2-3}$$

where d is the grating period of the DOE1 and DOE3. Suppose the grating functions of the DOE1, DOE2 and DOE3 are $\Phi 1$, $\Phi 2$ and $\Phi 3$ respectively. To maintain the direction of the light from the image source, the following equation must be fulfilled [65, 67]:

$$\Phi 1 + \Phi 2 + \Phi 3 = 0 \tag{2-4}$$

If a beam of light meets the same grating two times in a waveguide, the beam direction is maintained. This property can be applied in the second two-dimensional EPE (**Fig. 2.28**(b)). It is preferable that the intermediate grating has the same period and orientation as the first case to have higher efficiency. However, the system is

accurate even if the period deviates 10% from Eq. (2-3). The orientation of DOE2 is not critical either. DOE1 and DOE3 must have the same orientation. It is very similar to the one-dimensional EPE.

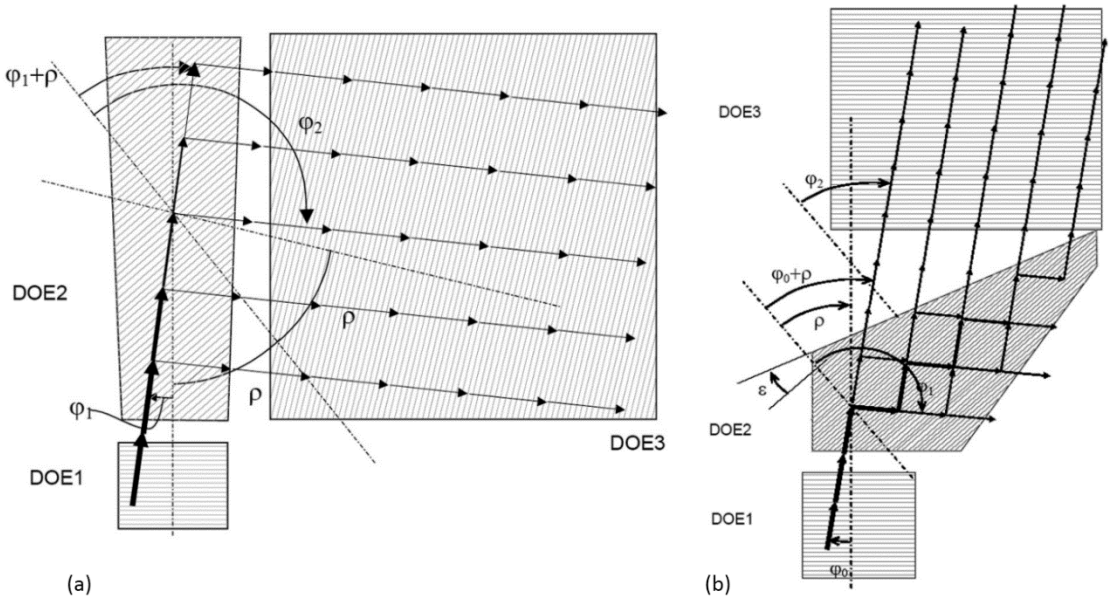



**Fig. 2.28.** (a) Layout of a two-dimensional EPE based on odd number of first-order diffractions and (b) Two-dimensional expansion using an even number of first-order diffractions [66].

Another type of patented two-dimensional EPE uses two overlapped DOE or photonic crystal as out-coupler as shown in **Fig. 2.29** [68]. Light can be expanded in two directions at the same time. It has the advantage of being extremely compact.

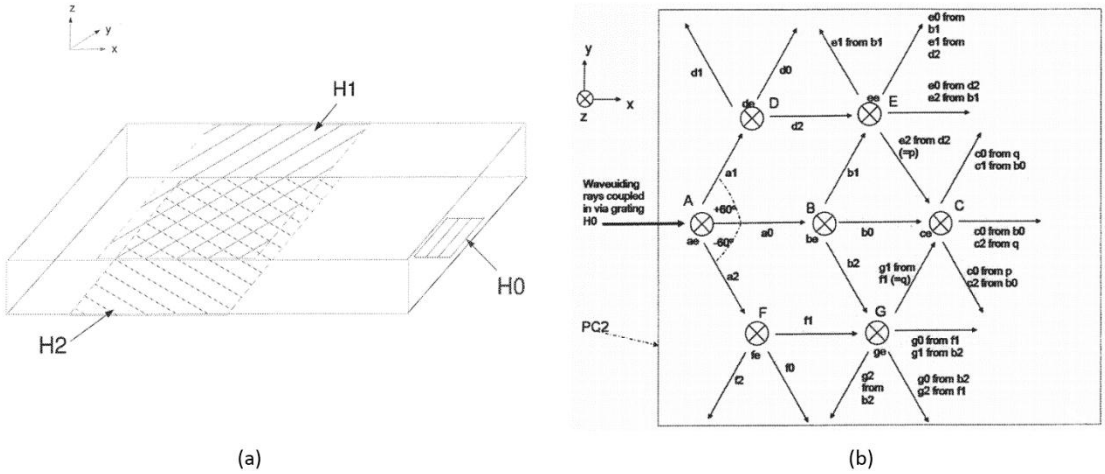



**Fig. 2.29.** (a) Perspective view of the waveguide using two overlapped DOEs as out-coupler. H1 and H2 are overlapped in xy plane and separated by the thickness of the waveguide in z direction. (b) Optical path through this system [68].

If the DOE uses sub-wavelength deep slanted diffraction gratings as in-coupler and out-coupler as shown in **Fig. 2.30** [66, 69-71], light can be diffracted into the desired direction in the first order and the other diffraction orders can be suppressed, which permit a higher overall efficiency. If the asymmetry is strong enough the whole optical system can be split into two mechanically separated parts. Each half of the split EPE is optically isolated from each other. If the incoming left and right half beams are well separated there is no crosstalk between left and right fields. As shown in **Fig. 2.30**(b), the input ports of the EPE can be equipped with fast (LC) shutters so that when the left beam is on, the right shutter is closed and vice versa, which can further improve the image quality.

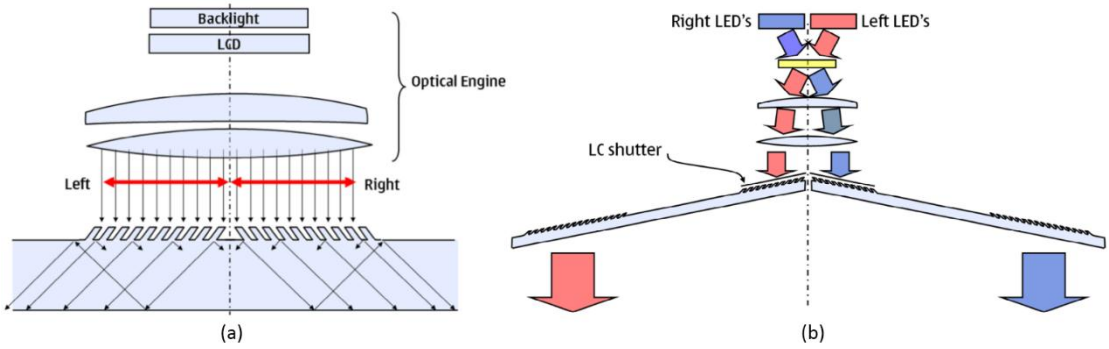



**Fig. 2.30.** (a)The asymmetric in-coupling grating divides the system in two mirrored parts. The OE is usually also symmetric around the center line and if the illumination is done properly sequentially formed stereoscopic images can be shown. (b) Schematic presentation of the light path from illumination to the exit of EPE in a sequential stereoscopic NED [69].

The slanted structures can be etched in special plastics using proper anti-adhesive treatment, molding temperature and the groove shape. 35 degrees of slanting angle using thermal or UV polymerizations have been replicated exactly [70]. Although the direction of light for all the wavelengths out of the out-coupler is the same as the incident direction, the directions are different inside the waveguide after passing only one DOE, which is like a “rainbow effect”. This adds a limitation to the FOV. The problem can be solved using multi-layer structure by propagating different wavelength in different waveguide increasing the number of elements

DOEs serving as in-coupler and out-coupler has great potential to minimize the optical structure volume and make the NEDs more ergonomic. The current developing

axis is to increase the energy efficiency and FOV. The diffractive waveguide technique has been successfully used in the products such as HoloLens of Microsoft and AR3000 of VUZIX as shown in **Fig. 2.31**.


**Fig. 2.31.** Appearance of HoloLens and VUZUX's AR3000.

As we have introduced, this technique has been integrated into commercial product. A DFOV of about 35° is achieved at present. Also, many layers are used for polychromatic system. Next, we introduce a similar system using holograms as in-coupler and out-coupler.

### 2.4.3.2. Holographic waveguide

The holographic technique is similar to the system using DOEs. Reflective volume holograms are used in this kind of system as shown in **Fig. 2.32** [72]. The fringe patterns of the in-coupling and out-coupling holograms are exactly the same and mirrored symmetrically positioned. Reflective volume holograms instead of transmission holograms are used because the diffraction bandwidths of the reflective holograms are much smaller than those of transmission ones. Hence, several holograms for different wavelengths can stack together with small interference between each other, which could potentially enlarge the FOV. As an example, with certain configuration, only an 8° horizontal FOV could be achieved with transmission volume holograms whereas a 24° horizontal FOV could be possible with the use of reflection volume holograms.

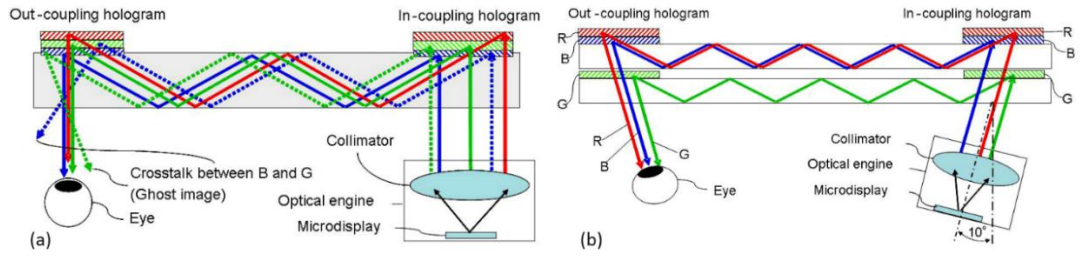



**Fig. 2.32.** (a) A basic structure of a holographic waveguide with 3 reflective holograms stacked together. Crosstalk within the holograms is shown. and (b) Two waveguides are used to minimize the crosstalk and oblique between the waveguide and the optical engine can reduce the Bragg λ disparity [72].

The diffraction efficiency characteristics of the in-coupling hologram layers are shown in **Fig. 2.33**(a). The spectrum range of each hologram layers is very narrow (within 50 nm). The sections inside the rectangles indicate the input R, G, and B LED spectra from the collimator. The black quadrangles demonstrate the spectral ranges coupled into the waveguide. The diffraction efficiency characteristics of the out-coupling hologram are shown in **Fig. 2.33**(b).

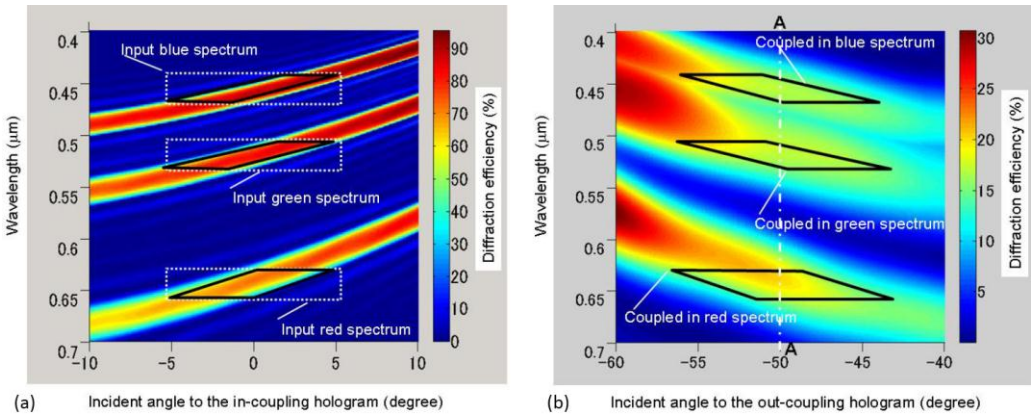



**Fig. 2.33.** (a) Diffraction efficiencies characteristics of R, G, and B hologram layers of the in-coupling hologram with respect to the incident angle. The thicknesses and index modulations of each hologram layer were approximately 5 μm and 0.04 respectively. (b) Diffraction efficiencies characteristics of R, G, and B hologram layers of the out-coupling hologram with respect to incident angle. The thicknesses and index modulations of each hologram layer were approximately 2 μm and 0.035 respectively [72].

Even though the diffraction spectrum of each hologram layer is very narrow, there are still crosstalk between each other. The calculated diffraction efficiencies of the out-coupling B and R hologram layers are shown in **Fig. 2.34**. The diffraction efficiencies of B, and R layer in green spectrum are about 1.5%, which causes stray light and ghost

image as shown in **Fig. 2.32**(a). Using two waveguides one for G and the other for B and R (**Fig. 2.32**(b)) can reduce the crosstalk to less than 0.2%.

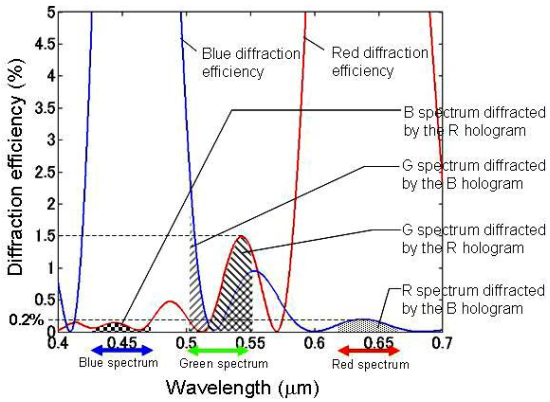



**Fig. 2.34.** Diffraction efficiencies of the out-coupling B and R hologram layers at an incident angle of 50° [72].

In conclusion, holography provides a way to build a NED. The see-through capacity is very good due to its narrow spectrum range. However, the maximum FOV is limited to about 30°. The color non-uniformity is also a problem and efficiency are not sufficient. All these problems need further consideration. Next, we describe systems using selective reflective coatings.

### 2.4.3.3. Selective thin film array waveguide

A system developed by Lumus uses several multilayer coatings inside glass substrate to couple light out of the waveguide [73, 74]. Before discussing the technique in detail, we firstly show the products of this company in **Fig. 2.35**. The diagonal FOV of 40 series is 23° while the 50 series is 40°. The thickness of the waveguides is less than 2 mm. The see-through quality is very good.


**Fig. 2.35.** Optical engine and development kits of Lumus [Lumus].

The main technique in this system is the multilayer coatings inside the substrate. The special property of the coating is its angular selectivity. In **Fig. 2.36**(a), the ray tracing of two incident rays with relative angles with the coatings denoted by $\beta_{ref}$ and $\beta'_{ref}$ are shown. The coating on a plane surface serves as a partially reflective planar mirror. Only in one case the incident rays should be reflected and the other should be transmitted to prevent undesired reflections and ghost images. It is possible to provide a coating with very low reflectance at high (low) incident angles and a high reflectance for low (high) incident angles. Hence, the desired function of the coatings can be achieved if one angle is significantly smaller than the other one.

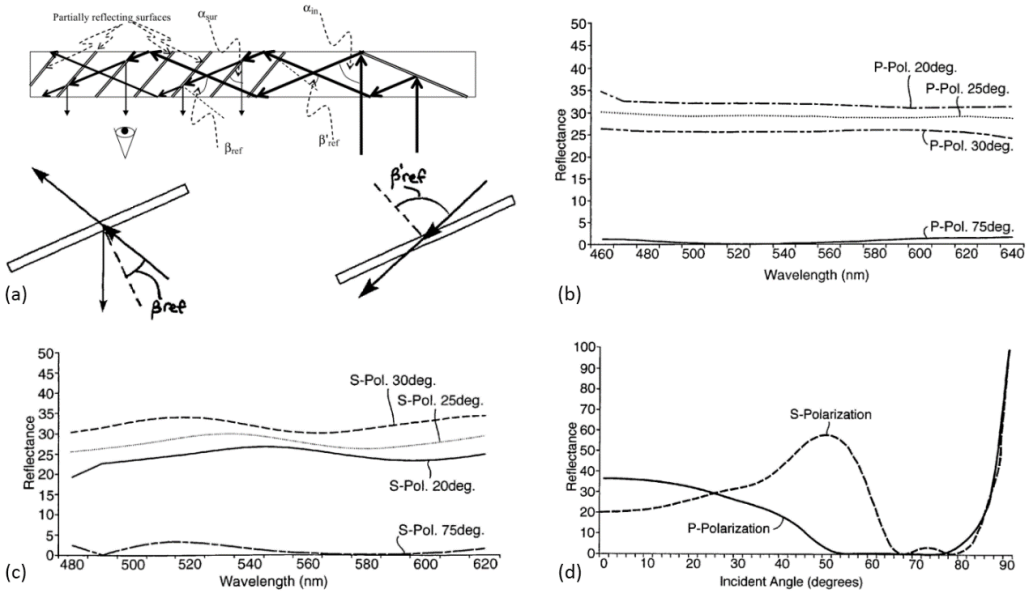



**Fig. 2.36.** (a) (up) A side view of the waveguide having an array of selective reflection coatings. (down) Alternative angular selective coatings. (b) The reflectance curves as a function of wavelength of an exemplary dichroic coating for P-polarization with incident angles of 20°, 25°, 30° and 75°. (c) The reflectance curves for S- polarization. (d) The reflectance curves as a function of incident angle for an exemplary dichroic coating [73].

**Fig. 2.36**(b) and (c) show the reflectance curves of a dichroic coating example with respect to wavelength for four different incident angles: 20°, 25°, 30° and 75° with P-polarized and S-polarized light respectively. The reflectance of the high-angle (75°) ray is negligible over the entire relevant spectrum while the reflectance is around 25% for low-angle (20°, 25°, 30°) rays. **Fig. 2.36** illustrates the reflectance curves at wavelength of 550 nm for the same coating above, as a function of the incident angle for both

polarizations. There are two different regions in this graph: between 15° and 40° where the reflectance is around 25% and between 65° and 80° where the reflectance is very low. For a given FOV, one can ensure only one reflection case and a ghost-free image if $\beta_{ref}$ is between 15° and 40° and $\beta'_{ref}$ is between 65° and 80°.

The low reflectance region in high angle for S-polarization is much narrower than P-polarization light. Hence, it is much more difficult to achieve a constant reflectance for a given angle for the S-polarized light than for the P-polarized light. It is therefore preferable to use only P-polarized light.

This technique is very basic and promising. The difficulty at present is to fabricate at a low cost. In the following section, we introduce another similar technique using small mirror stripe array.

### 2.4.3.4. Small mirror stripe array

Instead of using reflective coatings, metallic films can also be used [62, 75]. This technique was firstly proposed by Optinvent. One of their products ORA-2 is illustrated in **Fig. 2.37**(c). There are two types of configurations that can implement this technique. One is using reflective stripes directly on the waveguide as shown in **Fig. 2.37**(a). Between each stripe, light from the environment can be seen through directly, but the light hitting the mirror strips cannot be seen-through. However as the mirror stripes are very close to the eye and thus are out of the focus of the eye, the see-through capability is obtained. Another configuration is using a monolithic light guide plus a "cover plate" to protect the structure and to maintain the optical see-through function as shown in **Fig. 2.37**(b). The mirror stripes are coated with standard silver coatings and the Reflection/Transmission ratio can be predetermined. The R/T ratio can be adjusted very flexibly. For example, the transmission can vary from 30% to 70% while the reflectance varies from 70% to 30%, which gives the freedom to adjust the image brightness. The environment light hitting the mirror stripes can also be seen through these silver coatings.

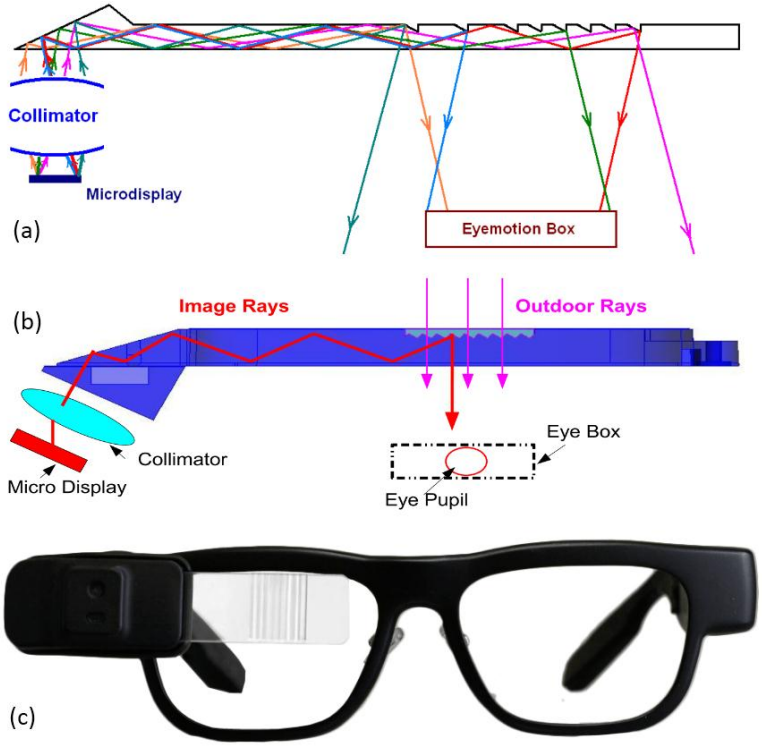



**Fig. 2.37.** System using small mirror stripes. (a) monolithic light guide. (b) monolithic light guide plus a "cover plate. (c) Optinvent ORA-2 [Optinvent].

The advantage of this system is that it can be manufactured at a low price. However, the pupil is not really expanded. A gap without light is generated after each reflection of the mirror stripe. The gap not covered by light can cause missing part in the image.

As we have already mentioned in free space combiner section, switchable Bragg gratings (SBG) can be integrated on the combiner. In the next section, we describe system that use SBG inside a waveguide NEDs.

### 2.4.3.5. Switchable Bragg Grating (SBG) out-coupling

We have described previously the system using SBG as a diffractive free space combiner. It is also possible to implement it into waveguide type. The key reason of using SBG is that it adds a dynamic aspect to the system. By switching on and off, it can avoid crosstalk between each other provided that the SBG is the same as a glass plate in off-state. FOV and colors can be tiled in different time if the switching speed is rapid enough [76]. As shown in **Fig. 2.38**, A, B and C represent different FOVs. Many different SBGs that correspond to different FOV are embedded inside a waveguide plate.

By switching on different SBGs in different time, crosstalk can be eliminated, and the system can have very large FOV if the SBGs have a sufficient switching speed.

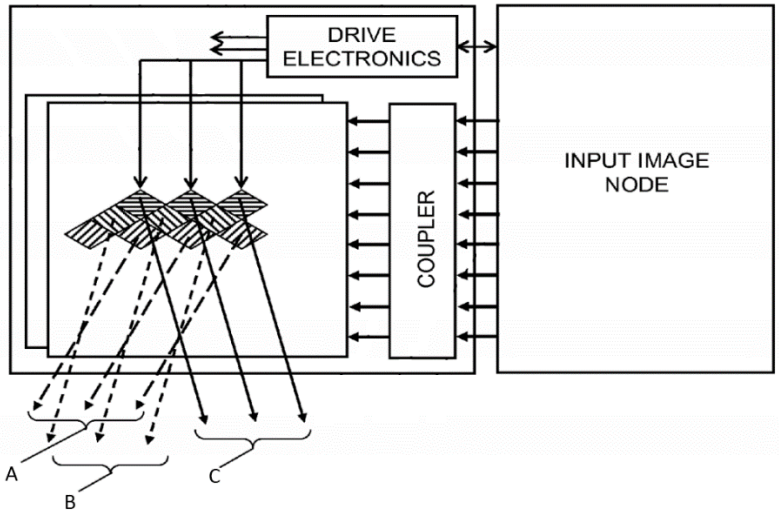



**Fig. 2.38.** An example of waveguide system using Switchable Bragg Gratings (SBGs) [76].

Waveguide NEDs is the most ergonomic structure among all the type of designs. Due to its pupil expansion capability, large eye box can be obtained with small projection system. The maximum DFOV of this kind is around 40°. Uniformity of color and diffraction efficiency over a large FOV is challenging by diffractive and holographic waveguides. With coating or mirrors, the main limitation for the FOV is the vertical FOV.

# 2.5. Conclusion

We have introduced most of existing NED designs and technologies that can be used. To summarize, we classified the existing NEDs into three families: NED with free space combiner, freeform prism NEDs and waveguide NEDs.

The primary advantage of the free space combiner design is that the see-through quality is the best for the moment. It also provides the highest theoretical luminance transfer from the display. Very large FOV (>100°) and eye box (~15 mm) can be achieved in some designs. However, high optical performance is always accompanied with a bulky structure. The key drawbacks are complexity, bulkyness and difficulties in assembling.

Freeform prisms do provide a compact design for NEDs. Some products with this design have been released. However, until now, the performance is limited to a DFOV about 50° and an eye box less than 8 mm. To have see-through capability, a compensator is needed. The final thickness is around 10 mm, which is still cumbersome.

Waveguide NEDs are the most ergonomic structures among all the types of designs. Due to its pupil expansion capability, large eye box can be obtained with small projection system. The maximum DFOV of this kind is around 40°. Uniformity of color and diffraction efficiency over a large FOV is still a challenge with diffractive and holographic waveguides. In systems using selective coatings or mirrors, the main limitation is the vertical FOV. The purpose of this thesis is to look for solutions to enlarge the VFOV of the waveguide NEDs.

## Reference


1. M. Billinghurst, A. Clark, and G. Lee, "A survey of augmented reality," Foundations and Trends® in Human–Computer Interaction **8**, 73-272 (2015).
2. P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," IEICE TRANSACTIONS on Information and Systems **77**, 1321-1329 (1994).
3. E. Dubois12, L. Nigay, J. Troccaz, O. Chavanon, and L. Carrat, "Classification Space for Augmented Surgery. an Augmented Reality Case Study," in *Humancomputer interaction, INTERACT' 99: IFIP TC. 13 International Conference on Human-Computer Interaction, 30th August-3rd September 1999, Edinburgh, UK*, 1999), 353.
4. R. B. Wood, "Holographic head-up displays," Electro-optical displays, 337-415 (1992).
5. J. Crawford and A. Neal, "A review of the perceptual and cognitive issues associated with the use of head-up displays in commercial aviation," The International Journal of Aviation Psychology **16**, 1-19 (2006).
6. J. A. Betancur, G. Osorio, and A. Mejía, "Integration of Head-Up Display system in automotive industry: a generalized application," in *Proc. of SPIE Vol*, 2013), 87360F-87361.
7. A. Hattori, K. Makita, and S. Okabayashi, "Development of HUD combiner for automotive windshield application," in *33rd Annual Techincal Symposium*, (International Society for Optics and Photonics, 1989), 272-282.
8. J. Rolland, "Head-Mounted Display Systems, Encyclopedia of Optical Engineering 2005," DOI **10**, 1-14.
9. O. Cakmakci and J. Rolland, "Head-worn displays: a review," Journal of display technology **2**, 199-216 (2006).
10. M. M. Bayer, "Retinal scanning display: a novel HMD approach for army aviation," in *Proc. SPIE*, 2002), 202.
11. J. Pandey, Y.-T. Liao, A. Lingley, R. Mirjalili, B. Parviz, and B. P. Otis, "A fully integrated RF-powered contact lens with a single element display," IEEE transactions on biomedical

circuits and systems **4**, 454-461 (2010).
12. H. Smet, J. Semt, D. Cuypers, C. Weng, and P. Joshi, "A Contact Lens With Built-In Display: Science Fiction or Not?," EuroDisplay, 8-11 (2013).
13. J. E. Melzer, *Head-mounted displays: designing for the user* (McGraw-Hill Professional, 1997).
14. M. Velger, "Helmet-Mounted Displays and Sights. Artech House," Inc, Boston, London (1998).
15. C. E. Rash, M. M. Bayer, T. H. Harding, and W. E. McLean, "Visual helmet-mounted displays," Chapter **4**, 109-174 (2009).
16. N. R. Council, *Tactical display for soldiers: Human factors considerations* (National Academies Press, 1997).
17. J. E. Melzer, F. T. Brozoski, T. R. Letowski, T. H. Harding, and C. E. Rash, "Guidelines for HMD design," helmet-mounted displays: sensation, perception and cognition issues, 805-848 (2009).
18. D. Lanman, H. Fuchs, M. Mine, I. McDowall, and M. Abrash, "Put on your 3D glasses now: the past, present, and future of virtual and augmented reality," in *ACM Special Interest Group on Computer Graphics and Interactive Techniques Conference, SIGGRAPH 2014*, (Association for Computing Machinery, 2014),
19. J. T. J. Arthur, R. E. Bailey, S. P. Williams, L. J. Prinzel, K. J. Shelton, D. R. Jones, and V. E. Houston, "Review of head-worn displays for the Next Generation Air Transportation System," Optical Engineering **56**, 051405-051405 (2017).
20. D. Desjardins, J. Metzler, D. Blakesley, C. Rister, A.-R. Nuhu, M. Edwards, D. Gains, A. Toet, E. L. Groen, and M. Oosterbeek, "Display Technologies and Applications for Defense, Security, and Avionics II," (2008).
21. D. P. Carey and C. V. Hutchinson, "Looking at eye dominance from a different angle: is sighting strength related to hand preference?," Cortex **49**, 2542-2552 (2013).
22. J. Geng, "Three-dimensional display technologies," Advances in optics and photonics **5**, 456-535 (2013).
23. S. Liu, Y. Li, P. Zhou, X. Li, N. Rong, S. Huang, W. Lu, and Y. Su, "A multi-plane optical see-through head mounted display design for augmented reality applications," Journal of the Society for Information Display **24**, 246-251 (2016).
24. P. Artal, "Optics of the eye and its impact in vision: a tutorial," Advances in Optics and Photonics **6**, 340-367 (2014).
25. P. Artal, "Image formation in the living human eye," Annual Review of Vision Science **1**, 1-17 (2015).
26. C. E. Rash, *Helmet mounted displays: Design issues for rotary-wing aircraft* (SPIE Press, 1999), Vol. 93.
27. G. Kramida, "Resolving the vergence-accommodation conflict in head-mounted displays," IEEE transactions on visualization and computer graphics **22**, 1912-1931 (2016).
28. J. Melzer, "How much is enough? the human factors of field of view in head-mounted displays," in *Degraded Environments: Sensing, Processing, and Display 2017*, (International Society for Optics and Photonics, 2017), 101970P.
29. J. M. Henderson, "Human gaze control during real-world scene perception," Trends in cognitive sciences **7**, 498-504 (2003).

30. M. Dorr, E. Barth, and T. Martinetz, "Effects of gaze-contingent stimuli on eye movements," Unpublished Diploma Thesis, Universität zu Lübeck (2004).
31. "U.S. Patent 2955156," (Oct. 4, 1960.).
32. I. E. Sutherland, "A head-mounted three dimensional display," in *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, (ACM, 1968), 757-764.
33. " U.S. Patent 6747611," (Jun. 8, 2004).
34. "U.S. Patent 20130033756A1," (Feb. 7, 2013).
35. H. Hua and C. Gao, "Design of a bright polarized head-mounted projection display," Applied optics **46**, 2600-2610 (2007).
36. H. Li, X. Zhang, G. Shi, H. Qu, Y. Wu, and J. Zhang, "Review and analysis of avionic helmet-mounted displays," Optical Engineering **52**, 110901 (2013).
37. "U.S. Patnet 3940204A," (Feb. 24, 1976).
38. R. V. Shack and K. Thompson, "Influence of alignment errors of a telescope system on its aberration field," in *24th Annual Technical Symposium*, (International Society for Optics and Photonics, 1980), 146-153.
39. J. Rolland and H. Hua, "Head-mounted display systems," Encyclopedia of optical engineering, 1-13 (2005).
40. J. P. Rolland, "Wide-angle, off-axis, see-through head-mounted display," Optical engineering **39**, 1760-1767 (2000).
41. Z. Zheng, X. Liu, H. Li, and L. Xu, "Design and fabrication of an off-axis see-through head-mounted display with an x–y polynomial surface," Applied optics **49**, 3661-3668 (2010).
42. J. Yang, W. Liu, W. Lv, D. Zhang, F. He, Z. Wei, and Y. Kang, "Method of achieving a wide field-of-view head-mounted display with small distortion," Opt Lett **38**, 2035-2037 (2013).
43. J. M. Sasian, "Image plane tilt in optical systems," Optical Engineering **31**, 527-532 (1992).
44. J. Wang, Y. Liang, and M. Xu, "Design of a Low Distortion Head-Mounted Display with Freeform Reflective Mirror Based on Two Ellipsoids Structure," Journal of the Optical Society of Korea **20**, 234-238 (2016).
45. "U.S. Patent 20140266986A1," (Sep. 18, 2014).
46. "U.S. Patent 20160139413A1," (May 19, 2016).
47. O. Cakmakci, B. Moore, H. Foroosh, and J. P. Rolland, "Optimal local shape description for rotationally non-symmetric optical surface design and analysis," Optics express **16**, 1583-1589 (2008).
48. "U.S. 20040108971A1," (Jun. 10, 2004).
49. M. Popovich and S. Sagan, "45.1: Application Specific Integrated Lenses for Displays," in *SID Symposium Digest of Technical Papers*, (Wiley Online Library, 2000), 1060-1063.
50. R. L. Sutherland, L. V. Natarajan, V. P. Tondiglia, and T. J. Bunning, "Holographic polymer-dispersed liquid crystals in display applications," in *Proc. SPIE*, 2003), 35-43.
51. M. Beuret, P. Twardowski, and J. Fontaine, "Design of an off-axis see-through display based on a dynamic phase correction approach," Optics express **19**, 19688-19701 (2011).
52. S. Yamazaki, A. Okuyama, T. Ishino, A. Fujiwara, and Y. Tamekuni, "Development of super compact hmd with sight line input," in *Proceedings of 3D Image Conference*, 1995), 70-75.
53. S. Yamazaki, K. Inoguchi, Y. Saito, H. Morishima, and N. Taniguchi, "Thin wide-field-of-view HMD with free-form-surface prism and applications," in *Electronic Imaging'99*,

(International Society for Optics and Photonics, 1999), 453-462.
54. H. Hoshi, N. Taniguchi, H. Morishima, T. Akiyama, S. Yamazaki, and A. Okuyama, "Off-axial HMD optical system consisting of aspherical surfaces without rotational symmetry," in *Electronic Imaging: Science & Technology*, (International Society for Optics and Photonics, 1996), 234-242.
55. D. Cheng, Y. Wang, H. Hua, and M. Talha, "Design of an optical see-through head-mounted display with a low f-number and large field of view using a freeform prism," Applied optics **48**, 2655-2668 (2009).
56. "U.S. Patent 20120162549A1," (Jun. 28, 2012).
57. Z. Zhuang, Q. Cheng, P. Surman, Y. Zheng, and X. W. Sun, "A compact and lightweight off-axis lightguide prism in near to eye display," Optics Communications **393**, 143-151 (2017).
58. D. Cheng, Y. Wang, H. Hua, and J. Sasian, "Design of a wide-angle, lightweight head-mounted display using free-form optics tiling," Optics letters **36**, 2098-2100 (2011).
59. Y. Amitai, S. Reinhorn, and A. Friesem, "Visor-display design based on planar holographic optics," Applied optics **34**, 1352-1356 (1995).
60. "U.S. Patent 4099841," (Jul. 11, 1978).
61. "U.S. Patent 4711512," (Dec. 8, 1987).
62. K. Sarayeddine and K. Mirza, "Key challenges to affordable see-through wearable displays: the missing link for mobile AR mass deployment," Proceedings spie **8720**, 87200D (2013).
63. I. Shariv, Y. Amitai, and A. Friesem, "Compact holographic beam expander," Optics letters **18**, 1268-1270 (1993).
64. "U.S. Patent 6169613," (Jan. 2,2001).
65. "U.S. Patent 6580529," (Jun. 17, 2003).
66. T. Levola, "Diffractive optics for virtual reality displays," Journal of the Society for Information Display **14**, 467-475 (2006).
67. L. Eisen, M. Meyklyar, M. Golub, A. A. Friesem, I. Gurwich, and V. Weiss, "Planar configuration for image projection," Applied optics **45**, 4005-4011 (2006).
68. "C.A. Patent 2956872A1," (Feb. 11, 2016).
69. T. Levola, "28.2: Stereoscopic Near to Eye Display using a Single Microdisplay," in *SID Symposium Digest of Technical Papers*, (Wiley Online Library, 2007), 1158-1159.
70. T. Levola, "7.1: Invited paper: novel diffractive optical components for near to eye displays," in *SID Symposium Digest of Technical Papers*, (Wiley Online Library, 2006), 64-67.
71. B. Kress, V. Raulot, and M. Grossman, "Exit pupil expander for wearable see-through displays," in *Proc. SPIE*, 2012), 83680D.
72. H. Mukawa, K. Akutsu, I. Matsumura, S. Nakano, T. Yoshida, M. Kuwahara, and K. Aiki, "A full-color eyewear display using planar waveguides with reflection volume holograms," Journal of the Society for Information Display **17**, 185 (2009).
73. "U.S. Patent 76720551," (Mar. 2, 2010).
74. Y. Amitai, "P-21: Extremely Compact High-Performance HMDs Based on Substrate-Guided Optical Element," in *SID Symposium Digest of Technical Papers*, (Wiley Online Library, 2004), 310-313.
75. G. Dubroca, P. Benoit, K. Sarayeddine, and X. Hugel, "Monolithic Low Cost Plastic Guide

for Full Colour See-Through Personal Video Glasses," in *Proceedings of International Display Workshops (IDW)*, 2010), 1355.

76. "U.S. Patent 2014104665A1," (Apr. 17, 2014).

# 3. Design of an in-coupling subsystem with freeform optics and a specific propagation subsystem

In this chapter we propose to investigate the design of a new waveguide near to eye display (WGNED). In the chapter dedicated to the state of art, we have mentioned the main advantage of these systems: the exit pupil expansion (EPE). After a short presentation of the background on WGNEDs, we will present a new design based on a stack of two waveguides. Firstly, we give a detailed description of the used principle. Then we present the possible in-coupling techniques and study in depth the coupling of the upper guide into the lower guide. Finally, we propose three in-coupling subsystems based on a freeform prism, freeform mirrors and freeform lenses. For each subsystem, the optical performances are provided.

## 3.1. Background

Let us consider the configuration where one dimensional EPE is used as shown in **Fig. 3.1**.

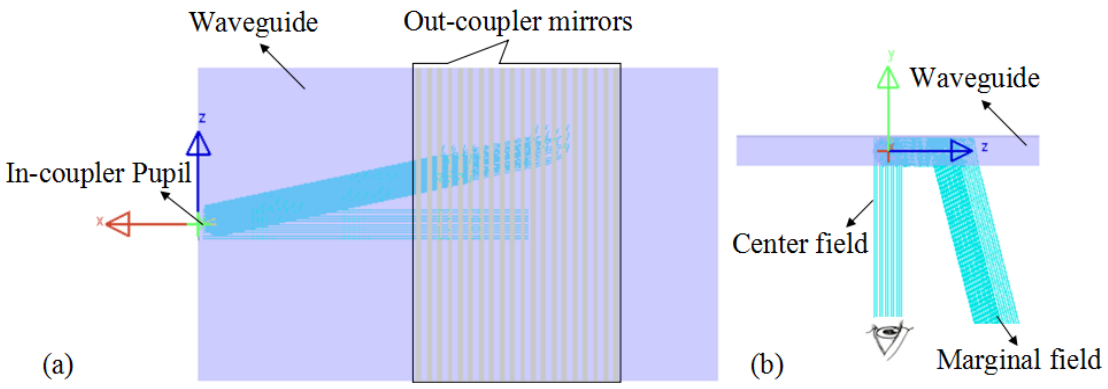



**Fig. 3.1.** (a) WGNED with one dimensional EPE. (a) xz view. (b) yz view.

Light rays of the different fields are passing through the in-coupling pupil. Only

the field around the center can be captured by eye. Other fields propagate far away from eye when coupled out from the waveguide as shown in **Fig. 3.1**(b). The maximum vertical FOV can be approximately calculated by [1]:

$$VFOV_{\max} = \frac{D - d_{eye}}{R_{eye} + l/(n\sin\alpha_{in})} \quad (3\text{-}1)$$

where $n$ is the refractive index of the substrate, typically in the range of 1.5-1.6; $l$, is the distance between the coupling-in the coupling out axes, which is between 40 and 60mm, and $D$ is the lateral dimension of the input waves coupled into the substrate (<15 mm). $d_{eye}$ is the desired exit-pupil diameter (typically, in the order of 8 mm), $R_{eye}$ is the distance between the eye of the viewer and the substrate (called also eye-relief typically, about 25-30mm), and $\alpha_{in}$ is the angle between the in-coupling surface and the normal to the substrate plane (≈60°). **Fig. 3.2** shows the various used geometrical parameters. The green color corresponds to the out-coupler. The blue color indicates the light propagative region. According to these values, the maximum VFOV is about 0.2°. Therefore, for a real application this configuration should be improved.

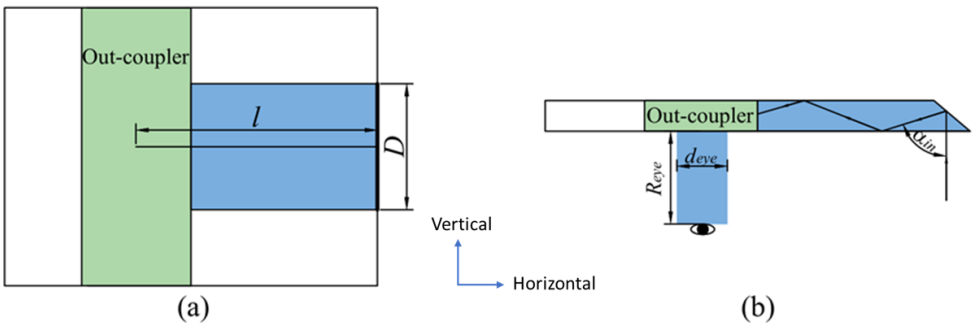



**Fig. 3.2.** Definition of the geometrical parameters, relative to the head. (a) The plane perpendicular to the line of sight of one eye. (b) The top view of the device with the position of the eye.

One possibility is to increase the value $D$, which leads to a configuration of two dimensional EPE. As we have specified previously in the second chapter, the two-dimensional EPE allows WGNEDs to have a very large eye box. This technology has been implemented to some products successfully. However, the diagonal FOV is limited to about 40° and the energy efficiency is very low for the moment.

Another method to enlarge the VFOV is to present two different exit pupils for the in-coupling subsystem using a free-form prism (FFP) [2, 3]. The exit pupil in xz plane

and the exit pupil in yz plane are shown in **Fig. 3.3**. The exit pupil in yz plane is also the in-coupling pupil of the waveguide while the exit pupil in xz plane is the position of the eye. The surface numbers are labeled 1, 1', 2, and 3. Surface number 1 and 1' represent the same surface but with twice ray hitting. Surface number 1 is used as transmission while surface 1' is used as total reflection. By this method, the achieved DFOV is 28°. This method increases the energy efficiency comparing to two dimensional EPEs. However, the FOV is still limited because the size of FFP can increase rapidly as the VFOV increases.

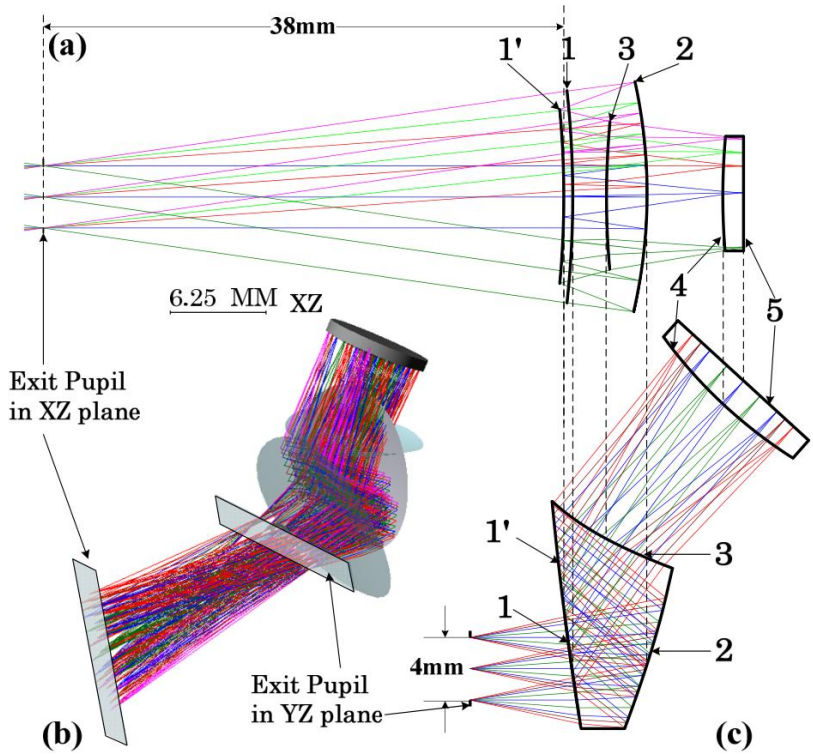



**Fig. 3.3.** Partial optical layout of the in-coupler with freeform eyepiece in (a) sagittal plane, (b) three-dimensional view, (c) tangential plane. The display source is surface number 5, the waveguide is not shown. The YZ plane is the vertical plane and the XZ plane is the horizontal plane [2].

Due to these reasons, we propose a new method to increase the VFOV by using two waveguides and a special edge.

# 3.2. Description of the complete system

## 3.2.1. Basic principle

**Fig. 3.4** shows the structure of the system and the ray path inside. Two waveguides are stacked up with a small air gap so that light can propagate inside the two waveguides independently. The light emitted by the micro-display is coupled into the upper waveguide by the in-coupling subsystem. Light propagates inside this waveguide until reaching an edge cylindrical mirror. In the area near the mirror, the two waveguides are combined together. Then, the light reflected by the cylindrical mirror is coupled into the lower waveguide. Using existing methods such as hologram or cascaded mirrors as out-coupler, the light is then coupled out from the second waveguide.

The center *O* of the in-coupling pupil, shown in **Fig. 3.4**(a), is the center of the cylindrical mirror and the center of the cylindrical edge of the air gap. After the reflection, the chief ray of each field will go backward and converge to a focal point. The associated propagating length is the radius of the cylindrical mirror.

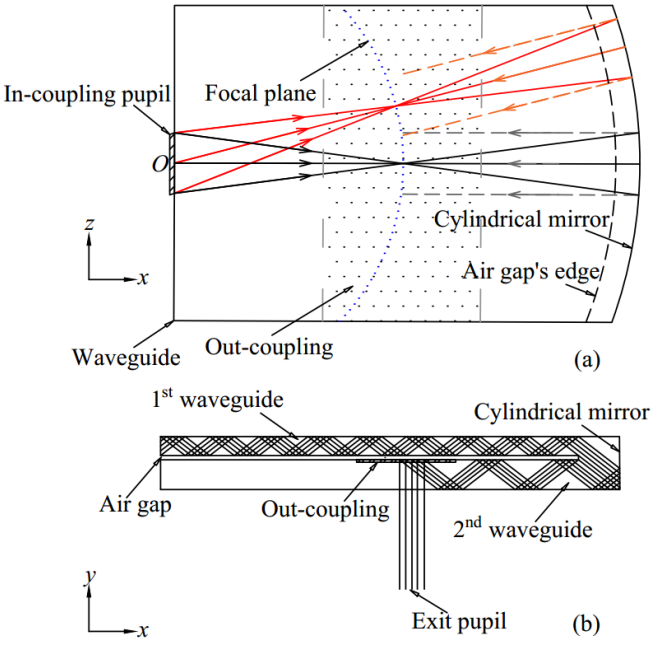



**Fig. 3.4.** Schematic for the designed system. (a) Ray path in zx plane. The dashed rays are the rays that are coupled into the second (lower) waveguide. (b) Ray path in yx plane. The air gap is magnified for clarity.

The focal property of light after the in-coupling pupil is different in zx plane and yx plane because of the cylindrical mirror which is shown in **Fig. 3.4**. We suppose that the radius of the mirror is *R*. In the case of paraxial approximation, rays converge to the focal plane of the cylindrical mirror which is located at *R*/2 from the center in zx plane as shown in **Fig. 3.4**(a). Then reflected rays by the cylindrical mirror become a parallel beam that is coupled into the second (lower) waveguide. However, as the cylindrical mirror has no optical power in yx plane, the light should be a parallel beam after the in-coupling pupil as shown in **Fig. 3.4**(b). To compensate the large optical power caused by the cylindrical mirror, free-form surfaces (FFS) must be used in the in-coupling system. We have proposed three configurations to design the in-coupling subsystem.

## 3.2.2. Possible in-coupling techniques

Normally, there are two ways to connect the in-coupler to the waveguide. The first way is to couple light directly to the waveguide through a beveled plane surface as shown in **Fig. 3.5**(a). The angle between the in-coupling pupil and the waveguide is $\phi$. The width of the light beam that can be coupled into the upper waveguide is *w* and the thickness of the waveguide is *d*. If the angle between the light and the horizontal direction of the waveguide is $\beta$, we have the following relation:

$$w = d\frac{\sin(\beta+\phi)}{\sin(\phi)} \tag{3-2}$$

*w* increases when $\phi$ decreases. However, in case where $\phi < \beta$, some light will be totally reflected by the in-coupling pupil as the red line in the **Fig. 3.5**(a). Those rays form the stray light of the system. The main axis of the in-coupler is perpendicular to the in-coupling pupil. The length of the in-coupler affects the total horizontal size of the NED. This causes one limitation for the design. The second way to couple light is to use a mirror on a beveled plane surface as shown in **Fig. 3.5**(b). The advantage is that the angle between the main axis of the in-coupler and the waveguide can be adjusted according to the angle $\varphi$ between the mirror and the waveguide. There is still an in-coupling pupil AB inside the waveguide the same as the first case if only part AD

of the mirror is used. AB is equivalent to the virtual pupil AC which is the mirrored image of AB. If we extend the mirror to AG, additional light can be coupled inside. The rays with an angle to the horizontal direction of the waveguide larger than $\varphi$ can be reflected again by the mirror. These rays also become stray light as the red lines in **Fig. 3.5**(b). However, if the angle is smaller than $\varphi$, these rays (the green lines) can be coupled inside the upper waveguide normally, which increase the energy efficiency of the system.

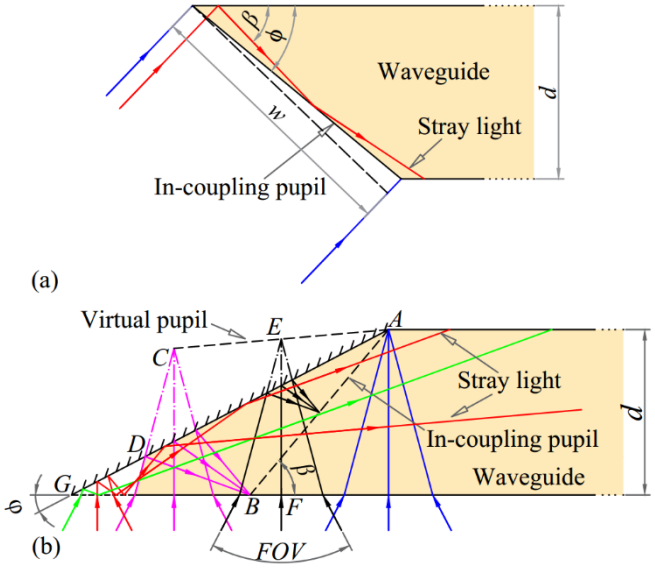



**Fig. 3.5.** Two ways of coupling light into the waveguide. (a) direct coupling, (b) coupling using a mirror.

## 3.2.3. Coupling light from the upper waveguide to the lower one

When light is coupled from the upper waveguide to the lower one, the amount of light $S$ that can be successfully coupled is varying with different incident angles $\theta$. $S$ also depends on $d_1$ and $d_2$ the thickness of the upper and lower waveguide respectively, the length of the transition region $T$, the length of the upper waveguide $L$ and the angle between the in-coupling pupil and the plane of the waveguide $\beta$ as shown in **Fig. 3.6**. The distance L is computed from the beginning of the in-coupling pupil to the beginning of the transition region AB. There is no simple and analytic formula to define the

relation between them. For each angle $\theta$, by analysis the rays hitting AB, we can calculate the amount of light that can be coupled successfully. **Fig. 3.6** shows an example, there are three different regions on AB that are hatched with color. The three colored regions on AB correspond to the colored regions DE, EF and FG at the beginning of the in-coupling pupil. For each $\theta$, there are two cases when rays hit AB: 1. Ray has an upward direction. 2. Ray has a downward direction. We use arrows pointing up and down inside the hatched region to indicate the two directions. The red color means that the rays hitting these areas with certain angle upward or downward cannot be coupled into the second waveguide. They are reflected to the first waveguide and go back to the in-coupling subsystem. The green regions depict rays that can be coupled successfully, i.e. which can hit the entrance of the lower waveguide BC. The length of the green areas on AB or DG may indicate the amount of light energy that can be coupled into the second waveguide.

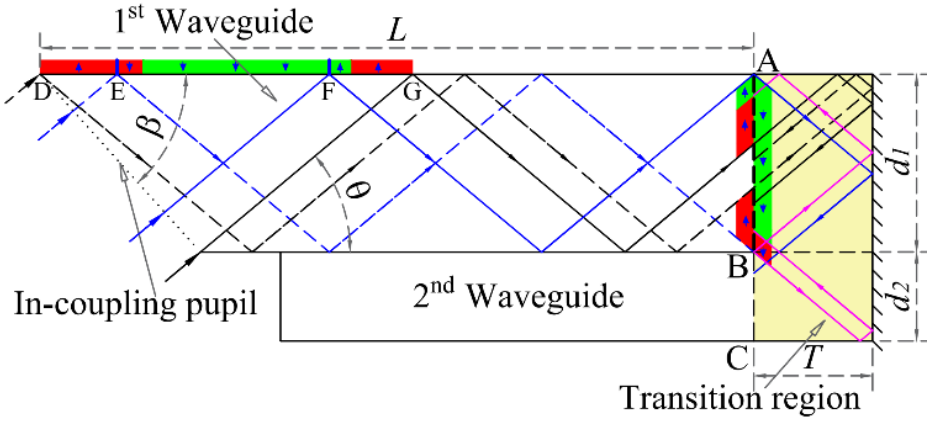



**Fig. 3.6.** The ray tracing from the in-coupling pupil to the transition region. The arrows inside the hatched region mean the ray hitting AB has upwards or downwards direction.

The position of rays that hit AB upward or downward, changes with the value of $\theta$. Even with the same position on AB and direction (upward or downward), different angles $\theta$ have different situations. **Fig. 3.7** shows the eight possible situations for the different angles $\theta$ and directions. The first row shows the rays having upwards direction while the second row shows the rays having downward direction. The green dashed sections on AB represent the useful part that can be coupled while the red dashed sections represent the part that cannot be coupled. The green rays with arrow represent

the rays that can be coupled. The pink rays with arrow show the critical geometry between coupling and non-coupling cases. The value of the four particular crucial angles $\theta_1$, $\theta_2$, $\theta_3$ and $\theta_4$ are given by:

$$\theta_1 = \arctan(\frac{d_1}{T}),$$

$$\theta_2 = \arctan(\frac{d_2}{T}),$$

$$\theta_3 = \arctan(\frac{d_1 + 2d_2}{2T}),$$

$$\theta_4 = \arctan(\frac{d_1}{2T}),$$

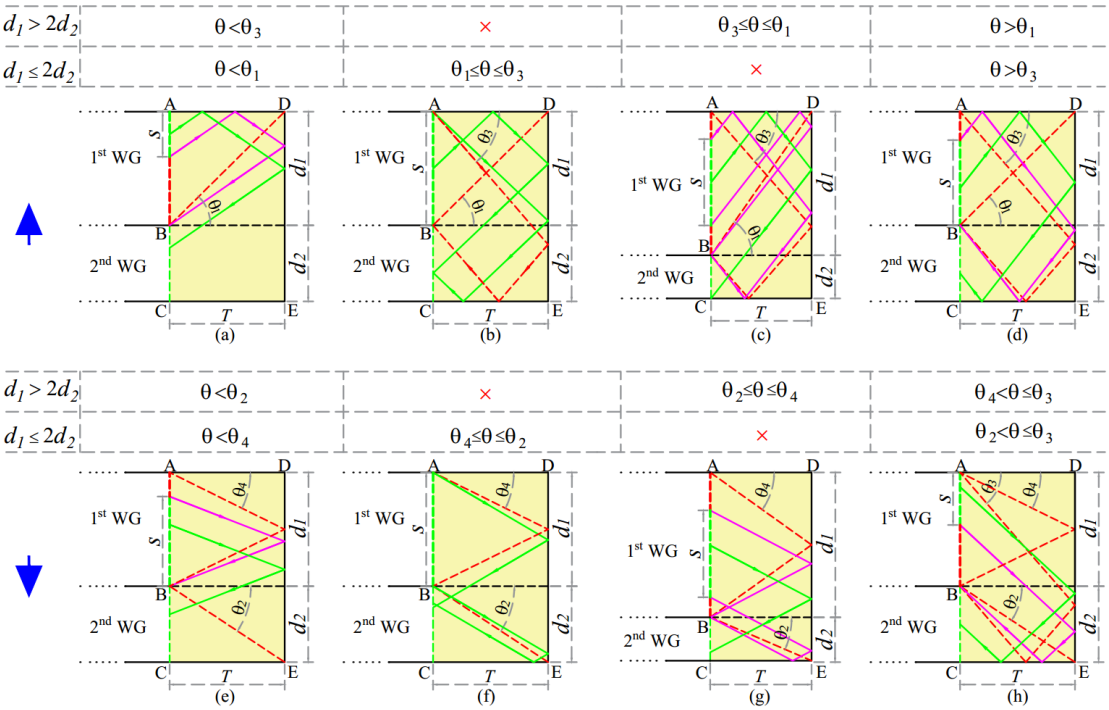



**Fig. 3.7.** The eight possible situations in the end of the two waveguides. The big blue arrow indicates the direction when light hits AB. The green dashed sections represent successfully coupling cases. The red dashed sections represent the non-coupling cases. The pink rays with arrow show the critical geometry between coupling and non-coupling cases.

$\theta_1$ represents the angle between the ray which hits the inner edge B and that is reflected to the up-corner D and the plane of the upper waveguide as shown in **Fig. 3.7**(a). $\theta_2$ represents the angle between the ray that passes through the inner edge B and reaches the down-corner E and the plane of the lower waveguide as shown in **Fig. 3.7**(e). $\theta_3$ represents the angle between the ray that is coming from point A and is reflected by

surface DE and CE and finally reaches point B and the plane of the upper waveguide as shown in **Fig. 3.7**(b). $\theta_3$ represents the angle between the ray that is coming from point A and is reflected by surface DE and reaches point B and the plane of the lower waveguide as shown in **Fig. 3.7**(e). When $d_1$>2 $d_2$, we have $\theta_3 < \theta_1$ and $\theta_2 < \theta_4$. The relationship changes when $d_1$<2 $d_2$, then we have $\theta_3 > \theta_1$ and $\theta_2 > \theta_4$. We indicate this separately in **Fig. 3.7**(b) and **Fig. 3.7**(c) when the rays have upward direction and in **Fig. 3.7**(f) and **Fig. 3.7**(g) when the rays have downward direction.

We use $s$ to represent the length of the ray section on AB that can be coupled. In the case where the rays have upwards direction and $d_1 \leq 2\ d_2$, we have:

$$s(\theta)=\begin{cases}2T\tan(\theta)-d_1 & \theta<\theta_1\\ d_1 & \theta_1\le\theta\le\theta_3\\ 2(d_2+d_1-T\tan(\theta)) & \theta>\theta_3\end{cases} \quad (3\text{-}3)$$

When $d_1$>2 $d_2$, we obtain:

$$s(\theta)=\begin{cases}2T\tan(\theta)-d_1 & \theta<\theta_3\\ 2d_2 & \theta_3\le\theta\le\theta_1\\ 2(d_2+d_1-T\tan(\theta)) & \theta>\theta_1\end{cases} \quad (3\text{-}4)$$

In the case where the rays have downwards direction and $d_1 \leq 2\ d_2$, we have:

$$s(\theta)=\begin{cases}2T\tan(\theta) & \theta<\theta_4\\ d_1 & \theta_4\le\theta\le\theta_2\\ d_1+2d_2-2T\tan(\theta) & \theta_2<\theta\le\theta_3\\ 0 & \theta>\theta_3\end{cases} \quad (3\text{-}5)$$

When $d_1$>2 $d_2$, we obtain:

$$s(\theta)=\begin{cases}2T\tan(\theta) & \theta<\theta_2\\ 2d_2 & \theta_2\le\theta\le\theta_4\\ d_1+2d_2-2T\tan(\theta) & \theta_4<\theta\le\theta_3\\ 0 & \theta>\theta_3\end{cases} \quad (3\text{-}6)$$

According to these relationships and combining **Fig. 3.6**, we can calculate the total green section $S(T, L, d_1, d_2, \beta, \theta)$ on AB with respect to $\theta$ by numerical calculation. Meanwhile, we can calculate the total energy $Ec$ that can be coupled into the second waveguide:

$$Ec(T,L,d_1,d_1,\beta,\theta)=S(T,L,d_1,d_1,\beta,\theta)\cos(\theta) \quad (3\text{-}7)$$

When designing the system, several parameters should be fixed at the beginning

according to other requirement. We set $L$=50 mm according to the size of glasses; $\beta$=50° according to the geometry of in-coupling. The thickness of the second waveguide is set to be $d_2$=1.5 mm. $\theta$ represents the horizontal FOV inside the waveguide. After all the parameters above are given, $Ec$ is then a function of $T$, $d_1$ and $\theta$. We also define another function $Ec_{min}$ which is the minimum value of $Ec$ with respect to the change of $T$ and $d_1$, when $\theta$ is between 25° and 50° as shown in **Fig. 3.8**(a). We choose the values of $T$ and $d_1$ which can lead to a maximum value of $Ec_{min}$ to have a maximum coupling efficiency in the full range of $\theta$ from 25° to 50°. Optimal values are obtained when $d_1$=2.72 mm and $T$=2.72 mm. Applying these parameters, we calculated $Ec$ with respect to $\theta$ (the green curve) as show in **Fig. 3.8**(b). The length of the TDRRs is set to $l$=1mm. The blue and red points correspond to the simulated result in Zemax with $w$=0.707 mm and $w$=5 mm. The rays can be full of one TDRR when $w$=0.707 mm while w=5 mm corresponds to size of in-coupling pupil. The theoretical result fits well the simulated result obtained by ray tracing. From the result, we can conclude that $w$ has little impact on the coupling efficiency. $Ec$ is not uniform with different incident angle $\theta$. The energy uniformity ( defined as $(I_{max}-I_{min})/(I_{max}+I_{min})$ ) is about 50% which should be corrected on the image source.

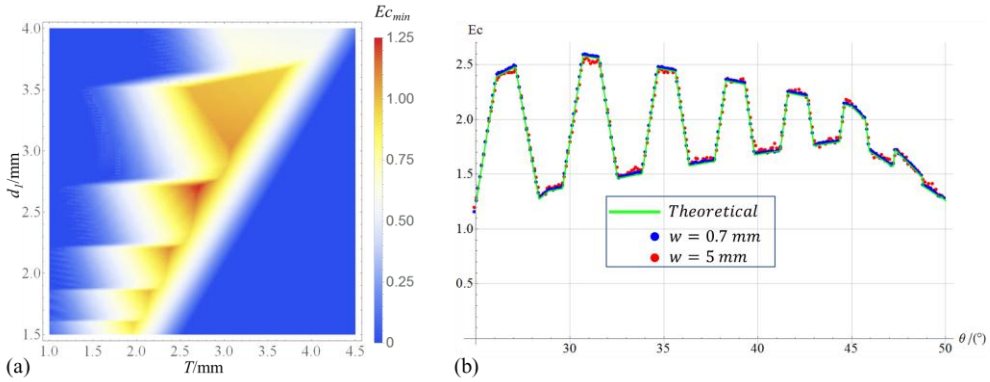



**Fig. 3.8.** Numerical Calculation results of the coupling from the upper waveguide to the lower waveguide. (a) Minimum coupled energy $Ec_{min}$ in all range of $\theta$ with respect to the thickness of the upper waveguide $d_1$ and the length of the transition region $T$, (b) $Ec$ with respect to $\theta$ when $d_1$=2.72 mm and $T$=2.72 mm Three curves in different color represent theoretical value, simulated two cases $w$=0.7 mm and $w$=5 mm.

We have studied the coupling between the two waveguides according to their

dimensions. The numerical calculated results have been verified by simulations. In the next sections, we introduce different ways of designing an in-coupler and their optical performance. The results of coupling between the two waveguides are valid in all cases.

### 3.2.4. In-coupling subsystem with a free form prism

The in-coupling subsystem consists of a FFP with four different optical FFSs (from S2 to S5) and a lens made of a freeform surface (S1) and a plane surface is shown in **Fig. 3.9**. This complex prism gives more freedom to correct aberrations. Moreover, it can be fabricated in one piece. As the FOV in horizontal direction is limited to 30° in air, the total light section is small; as a consequence, S2 and S4 surfaces can be separated. All the surfaces are XY polynomials. The optical power in yx plane is mainly given by the mirror S3 (**Fig. 3.9**(a)). The in-coupling pupil has a rectangular shape with a size of 1.9 mm (H) × 12 mm (V) which acts as a stop aperture and then greatly reduces the stray light. Pupil size is enlarged by adding vignetting along *z* direction as shown in **Fig. 3.9**(b) and (c). The position of the eye should be near the point where all the fields converge, so that each field can be seen as shown in Fig. 3.9(c).

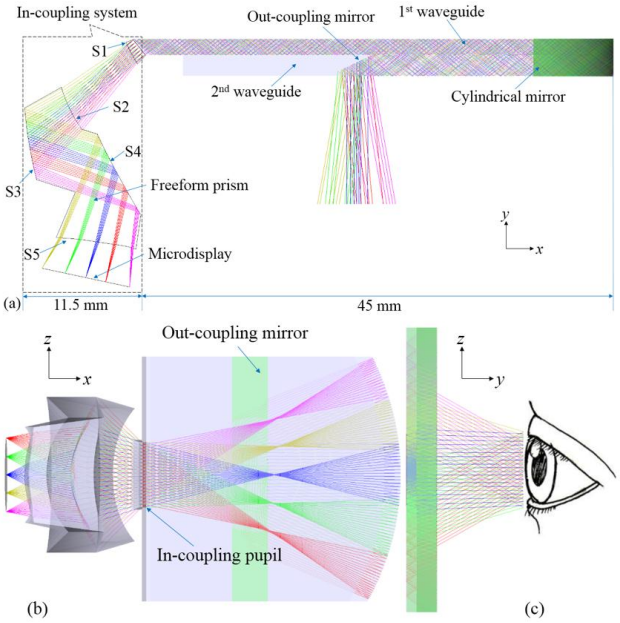



**Fig. 3.9.** The whole structure and optical path of the proposed system from different points of view: (a) yx view, (b)zx view, (c)zy view.

The advantage of this design is a large pupil size (12 mm) in z direction with high image quality. Using the in-coupling system mentioned above, the beam then becomes parallel in both directions after the reflection from the cylindrical mirror. Finally, it is coupled out to the user's eye.

We designed this system using Zemax. A 0.61 *in* OLED micro display with 852×600 pixels is used as image source. The pixel size is 15 μm which corresponds to a Nyquist frequency of 33 lp/mm. The horizontal FOV is set to 30°. The vertical FOV is 60°. The MTF of the 12 object fields are plotted in **Fig. 3.10**. They are evaluated with a full rectangular pupil size of 1.9 mm (H) × 12 mm (V). All the values are above 0.3 at 33 lp/mm, which is sufficient for a visual system. The vertical and horizontal magnifications are different and in a ratio of 1.6. **Fig. 3.11** shows the grid distortion of the system with respect to the central field. The maximum distortion is 4.5% at the corner.

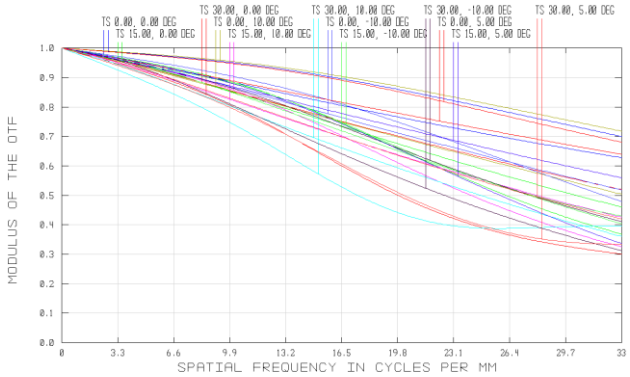



**Fig. 3.10.** MTF curves of the system with the 12 evaluated fields in object angle mode. T and S represent the tangential and sagittal MTFs of each field separately.

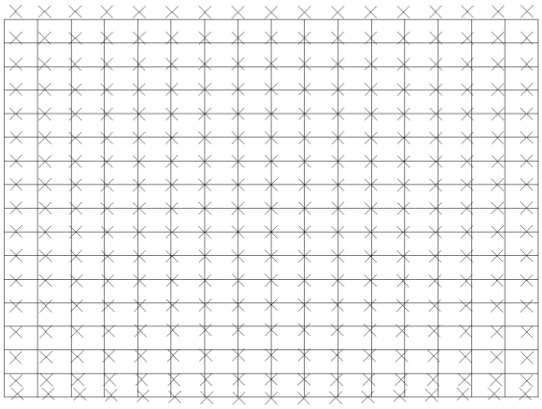


**Fig. 3.11.** The grid distortion of the system with a maximum distortion of 4.5% at the corner.

The performance of the system is summarized in the **Table. 3.1**. To estimate the value of the eye box in the horizontal direction a Lumus waveguide type out-coupling system has been used. The estimated weight in table 1 takes into account all the optical components.

**Table. 3.1.** Performance of the system

| Parameters | Value |
|---|---|
| Field-of-view (H×V) | 30° × 60° |
| In-coupling pupil (H×V mm) | 1.9×12 |
| Eye box (H×V mm) | ≈15×12 |
| System MTF | > 0.3 *at* 33 lp/mm |
| Distortion | < 4.5% |
| Monocular estimated weight (g) | 11 |
| Magnification ratio (V/H) | 1.6 |

**Fig. 3.10** and **Fig. 3.11** show that a good image quality can be expected for the whole FOV. A foreseen difficulty comes from fabrication and testing since the freeform prism has a complex shape. The current way to realize it, is plastic molding injection. A prototype fabrication is still a high cost operation.

## 3.2.5. In-coupling subsystem with free form mirrors

In this section, we will introduce an alternative way of designing an in-coupler: using freeform mirrors instead of prism. The benefit is a reduction of fabrication difficulties and weight. We firstly describe the whole system. Then we introduce the optimization method we have used. In the end, we present the optical performances and tolerance analysis.

### 3.2.5.1. Description of the system

**Fig. 3.12** shows the 3D structure of the system and optical rays for different object fields. Each object field is represented by a different color. We use a two-waveguide structure with a cylindrical mirror. This later changes the symmetry of the in-coupling subsystem, which leads to a big difference on the optical power of the in-coupling subsystem in vertical and horizontal directions. Then, making use of refractive surfaces only, the aberrations correction becomes difficult and the in-coupling pupil size is limited. In this design, the in-coupling subsystem is composed of two freeform mirrors, one freeform plastic lens and one spherical glass lens.

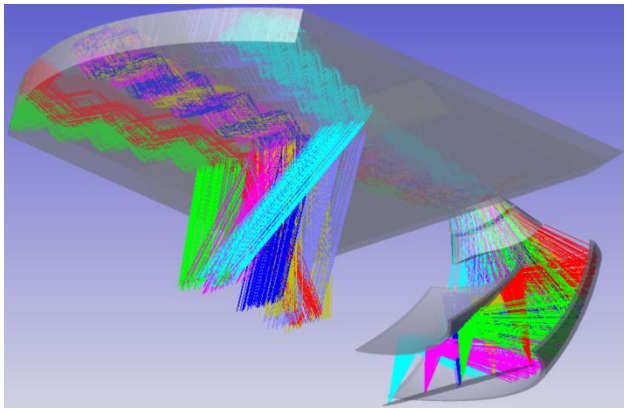


**Fig. 3.12.** 3D structure and optical path of the system

**Fig. 3.13** shows different views of the system. M1 and M2 are freeform mirrors. L1 is a plastic lens with freeform surfaces in both sides. The material is E48R from ZEONEX®. The type of all the freeform surfaces is XY-polynomial. L2 is a spherical glass lens with material of SF56A from SCHOTT®. The glass lens does not add too much weight to the system because the volume is very small (0.05 cm3). Rays emitted from the micro display are firstly reflected by M1 and M2. Then the rays go through the two lenses L1 and L2 and reach the in-coupling pupil. For eyepieces, we need parallel beams as output. Therefore, the rays of each field after the out-coupling mirror should be parallel. As the two-layer waveguide has no optical power in yx plane, the rays entering the in-coupling pupil should be parallel also in yx plane. However, in zx plane, the two-layer waveguide has positive optical power because of the cylindrical mirror. Let R be the radius of the cylindrical mirror, in paraxial approximation, the rays after the in-coupling pupil are focused on a circle of radius R/2 in zx plane as shown in

**Fig. 3.13**(b). With this configuration, the rays of each field become parallel after the reflection on the cylindrical mirror as shown in **Fig. 3.13**(c). The pupil size is designed to be 10 mm by 2 mm. Using cascaded mirrors as out-coupler, the eye box size can be enlarged to 15 mm in horizontal direction and 10 mm in vertical direction.

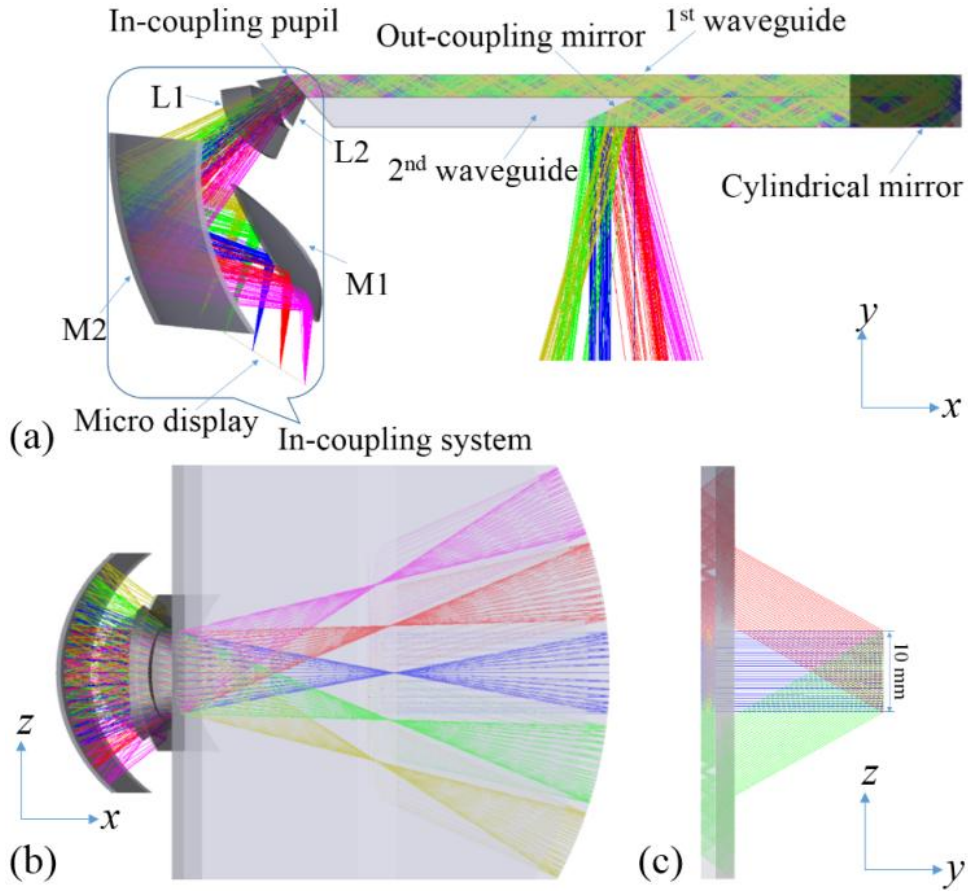



**Fig. 3.13.** Different views of the proposed system. (a) yx view (b) zx view (c) zy view

### 3.2.5.2. Optimization method

In the design process, we can model the waveguides as non-sequential components and build a complete system. As shown in **Fig. 3.14**, the object is the micro display. We used an OLED display with 0.61 *in* (1.5494 cm) diagonal size. An ideal thin lens must be added after the out-coupling waveguide to focus the beams on the image plane so that we can optimize the system. The focal length of the ideal thin lens was set so that the image surface size matches the object micro display size. In this case, the system is considered well corrected when the spot size is comparable to the pixel size of the micro

display. However, we found that inserting non-sequential components greatly increases the computing time. Moreover, the optical design software ray aiming function must be used so that each chief ray from the micro display reaches the center of the in-coupling pupil. The activation of the ray aiming function greatly increases the calculation time when the system is off-axis and when the rays go through several freeform surfaces. Because of these two reasons, the optimization becomes very time consuming. On a computer with two cores and 3.3 GHz CPU frequency, optimization time for one cycle by inserting non-sequential components is about 108 s in Zemax. We built an equivalent system to reduce the computational complexity. By this method, the optimization time for one cycle under the same conditions is only 0.47 s, which is 230 times faster. In fact, as the waveguide has no focal power, it can be replaced by a homogeneous isotropic medium. The position of the cylindrical mirror in the end of the first waveguide should be properly computed to keep the equivalence of the systems.

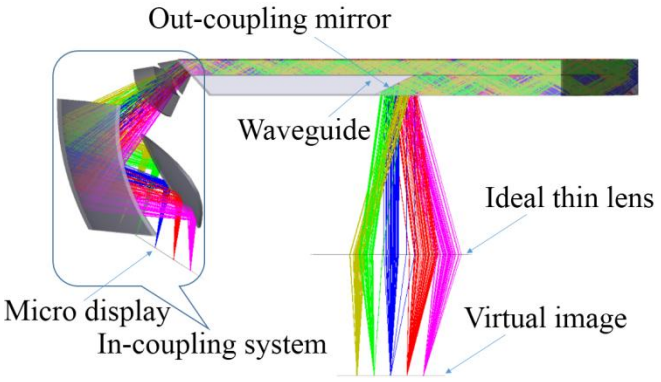



**Fig. 3.14.** Optimization set up by using non-sequential waveguides and adding ideal thin lens

As shown in **Fig. 3.15**, we firstly consider the chief ray of the central field which passes through the center of the in-coupling pupil O. It is totally reflected inside the waveguide several times, with the path OABCD. The edges of the waveguide act as planar mirrors, obviously, we have the following equalities AB'=AB=B'C'=BC=C'D'=CD and D'O'=DE. Let $n$ be the refractive index of the waveguide medium. Therefore, if we consider that the external medium is identical to the waveguide medium, $n\times$OO' represents the total optical path equivalent to the optical path $n\times$OABCDE inside the waveguide. The original cylindrical mirror can be replaced

by an equivalent cylindrical mirror with center O', with the same radius of curvature R, decentered by OO'. The value can be calculated by:

$$OO' = \frac{R}{\cos(\varphi)} \tag{3-8}$$

where φ is the angle between the chief ray of the central field and the Ox axis. The equivalent cylindrical mirror is enlarged enough to reflect all the incident object rays. Let β be the angle between the in-coupling pupil plane and the waveguide axis. Thus, the angle between the equivalent mirror and the in-coupling pupil is the supplementary angle of β. With these parameters, the position between the in-coupling pupil and the equivalent mirror is determined. We add a boundary after the equivalent mirror. It acts as the output plane of the second waveguide. The distance between F, the center of the output plane, and the point O' can be chosen quite roughly in the range from 10 mm to 50 mm without a deterioration of the image. The normal to the boundary plane is parallel to the chief ray of the central field.

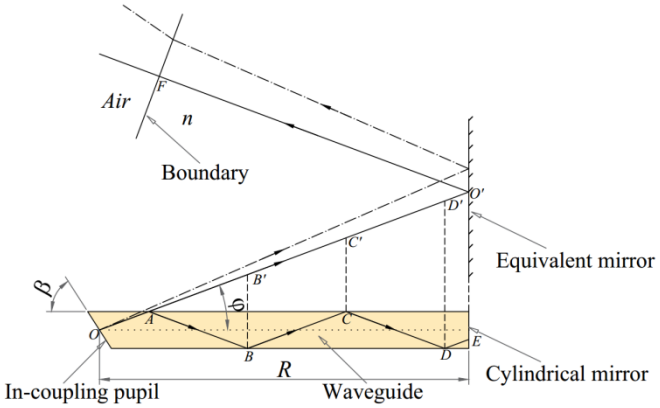



**Fig. 3.15.** Principle of replacing the waveguide by a medium of the same refractive index and an equivalent mirror

We can still design the system starting from the micro display and add an ideal lens to focus the light after the equivalent mirror as shown in **Fig. 3.16**(b). To reduce the computational effort using the ray aiming function, an inverse ray tracing design is best suited. Indeed, starting from the user's eye to the micro display, all the chief rays of each field are aiming to the center of the in-coupling pupil after reflection on the equivalent mirror as shown in **Fig. 3.16**(a). Here, the ray aiming function calculation is done for only one surface: the equivalent mirror.

To verify the equivalence between the non-sequential system and the proposed sequential system, after the optimization, we reversed the system from **Fig. 3.16**(a) to **Fig. 3.16**(b). We use an ideal thin lens after the boundary plane to focus the spots associated to each field on the image plane. Using the same ideal thin lens, we get the spots from the configuration in **Fig. 3.14**. The result is identical, which verifies the equivalence from the ray tracing point of view.

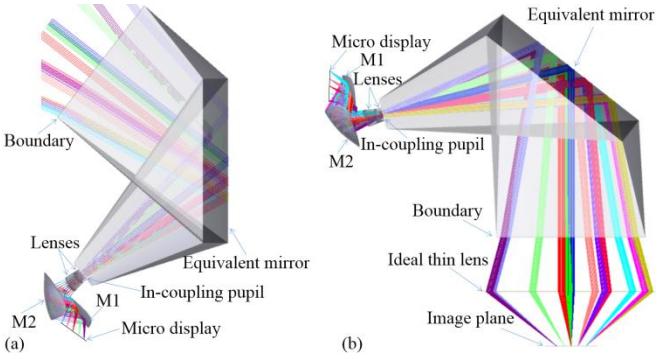



**Fig. 3.16.** Equivalent optimization setup (a) design from the user's eye to the micro display (b) design from the micro display to the eye by adding an ideal thin lens

### 3.2.5.3. Optical performance of the system

Using the setup shown in the **Fig. 3.16**(a), we optimized the system and evaluated the optical performance on the plane of the micro display. The polychromatic MTF of the system is shown in **Fig. 3.17**. All the values are above 0.15 at 30 lp/mm. The average value is 0.256 at 30 lp/mm. This MTF is evaluated through the whole in-coupling pupil which has a rectangular shape of 2 mm × 10 mm. As the vertical size is about 3 times larger than the human eye's pupil that has a diameter of about 3 mm in diurnal vision, the given MTF values are the minimal values.

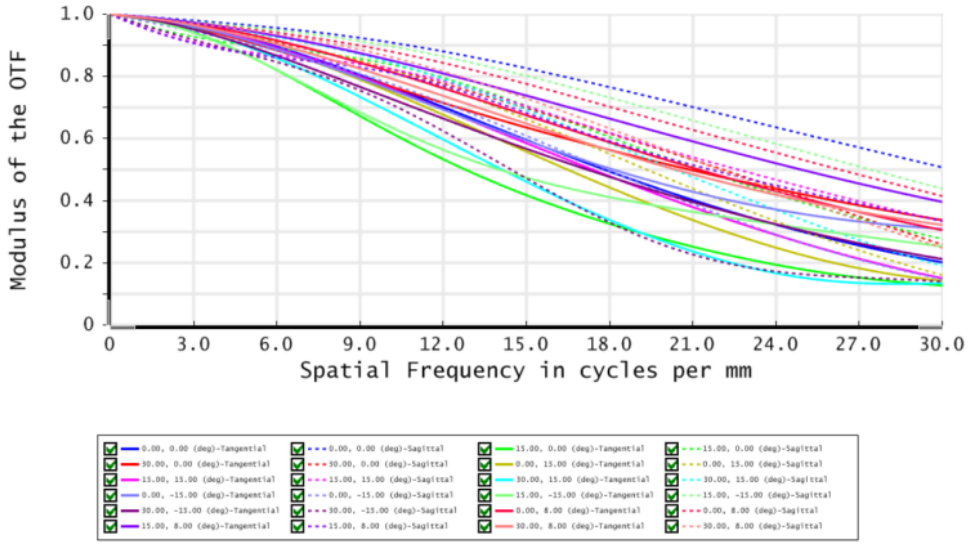



**Fig. 3.17.** Polychromatic MTF of the system.

**Fig. 3.18** shows the Root-Mean-Square (RMS) spot radius field map. The horizontal and vertical axis correspond to the x and y field angle respectively. The maximum value of the RMS spot radius is 11.5 μm which is very like the pixel size of the micro display.

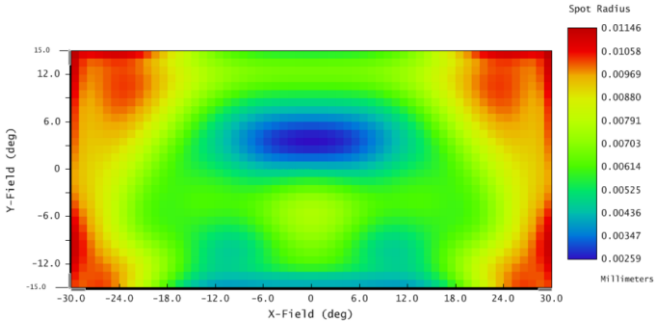



**Fig. 3.18.** RMS spot radius field map.

The grid distortion is shown in **Fig. 3.19**. The barrel distortion has a maximum value of 6.6%. The distortion can greatly affect the image quality of the system. However, distortion correction by purely optical means is limited and usually large distortion is corrected by image processing, which can affect the resolution of the image. Good compromise between the image quality and resolution should be considered according to the specific application.

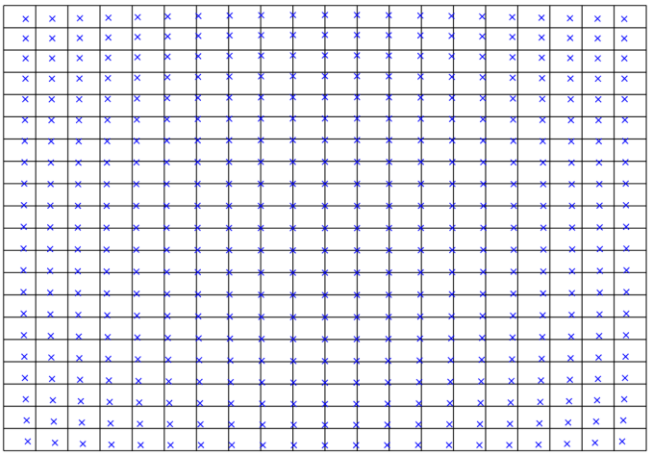


**Fig. 3.19.** Grid distortion of the system

The characteristics of the system are summarized in **Table. 3.2**. The total mass of the optical elements is estimated to 9.3 g which includes the two freeform mirrors, the plastic freeform lens, the spherical glass lens and the two-layer waveguide. The freeform mirrors are supposed to be 1 mm thick and made of Aluminum.

**Table. 3.2.** Performance summary of the system

| Parameters | Value |
|---|---|
| Field-of-View (H×V) | 30° × 60° |
| In-coupling pupil (H×V mm) | 2×10 |
| Eye box (H×V mm) | ≈15×10 |
| System MTF | > 0.15 at 33 lp/mm |
| Distortion | < 6.6% |
| Monocular estimated weight (g) | 9.3 |
| Waveguide thickness (mm) | 3.5 |

### 3.2.5.4. Tolerance of the system

It is hard to evaluate the tolerance of a system with freeform surfaces because there are too many parameters and each parameter has no obvious physical meaning. But we can analyze the regular parameters like the thickness, decenter and tilt for each surface, except the surface parameters of the freeform surface. From these regular parameters,

we can have a rough knowledge of the sensitivity of the system. We apply the precision criterion which is described in Ref [4] as tolerance criteria as shown in **Table. 3.3**(a) and run 1000 cycles of Monte Carlo analysis, the results are shown in **Table. 3.3**(b). The simulation was done at 30 lp/mm for all the fields. Based on the Root-Sum-Square method, the estimated averaged MTF is 0.183 which gives a reduction of 28.5% comparing to the nominal MTF as shown in **Table. 3.3**(c). Per the result, the system is not sensitive. Concerning the fabrication and the test of the freeform mirrors, nowadays these procedures are not standard and then costly. But in the future, we can expect that they become standardized and easy to test.

**Table. 3.3.** Tolerance criterion and result.

| (a) Tolerance criterion | |
|---|---|
| Parameters | Value |
| Wavefront residual | 0.1 RMS, 0.5 P-V |
| Thickness (mm) | 0.01 |
| Radius (%) | 0.1 |
| Index | 0.0001 |
| V-number (%) | 0.1 |
| Decenter(mm) | 0.01 |
| Tilt (arc sec) | 10 |
| (b) Monte Carlo analysis: Number of trials: 1000 | |
| Probability | MTF |
| 98% > | 0.1695 |
| 90% > | 0.1974 |
| 80% > | 0.2087 |
| 50% > | 0.2299 |
| 20% > | 0.2466 |
| 10% > | 0.2529 |

| (c) Estimated performance changes based upon Root-Sum-Square method: | |
|---|---|
| Nominal MTF | 0.2561 |
| Estimated change | -0.0730 |
| Estimated MTF | 0.1831 |

## 3.2.6. In-coupling subsystem with freeform lens

The structure of the system and the ray path inside is shown in **Fig. 3.20**. The system is very similar to the two previous cases we have described except that only refractive FFSs are used in the in-coupling subsystem. However, it becomes very hard to correct the large optical power difference in two directions with only refractive surfaces. The pupil size is limited to very small value (1.9 mm in the design).

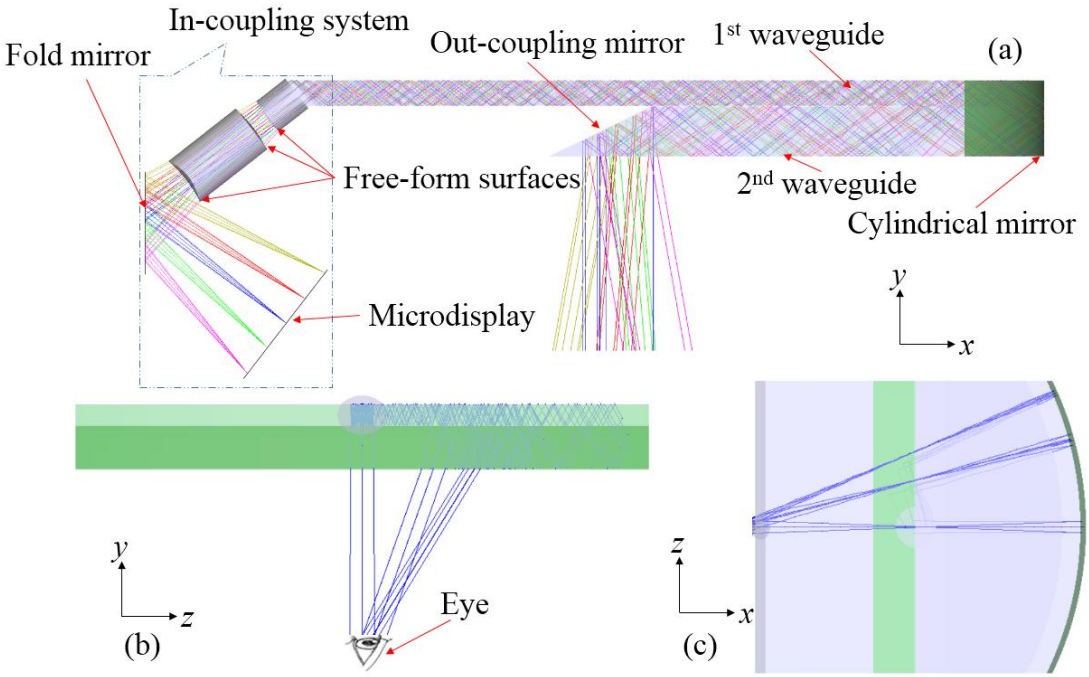



**Fig. 3.20.** The whole structure and optical path of the proposed system in different view. (a) yx view. (b)zx view. (c)yz view.

We have used a 0.61*inch* polychromatic micro-display with 852×600 pixels as the image source. The pixel size is 15 μm which corresponds to a Nyquist frequency of 33 lp/mm. The horizontal FOV is set to 20° and the vertical FOV to 55°. The polychromatic modulation transfer function (MTF) of the 12 object fields are plotted in

**Fig. 3.21**(a). The grid of distortion at the reference wavelength of 587.5 nm is shown in **Fig. 3.21**(b). The maximum distortion is 7%.

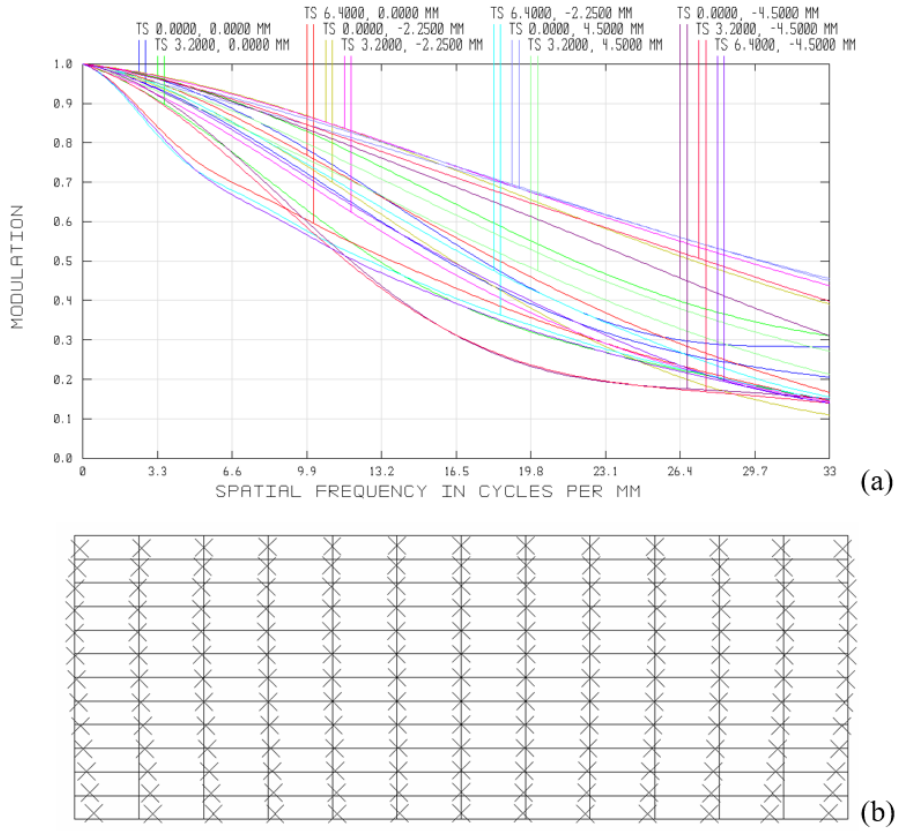



**Fig. 3.21.** (a) The polychromatic MTF of the 12 object fields of the system. (b) Grid distortion at the reference wavelength.

In conclusion, we firstly analyzed the coupling between the two waveguides which is valid in for all the setup of in-coupler. Then we proposed three ways to design an in-coupler: using freeform prism plus freeform lenses, using freeform mirrors + freeform lenses and using freeform lenses only. The optical performance of using freeform prism and mirrors are similar but better than only using freeform lenses. The reason of involving freeform surfaces is because the asymmetry caused by the cylindrical mirror in the end of the waveguides. In the next chapter, by designing the edge of the waveguides, we demonstrate that the asymmetry can be removed, and the in-coupler can be rotationally symmetrical only.

## Reference

1. Y. Amitai, "P-27: A Two-Dimensional Aperture Expander for Ultra-Compact, High-Performance Head-Worn Displays," in *SID Symposium Digest of Technical Papers*, (Wiley Online Library, 2005), 360-363.
2. D. Cheng, Y. Wang, C. Xu, W. Song, and G. Jin, "Design of an ultra-thin near-eye display with geometrical waveguide and freeform optics," Opt Express **22**, 20705-20719 (2014).
3. J. Han, J. Liu, X. Yao, and Y. Wang, "Portable waveguide display system with a large field of view by integrating freeform elements and volume holograms," Optics express **23**, 3534-3549 (2015).
4. R. R. Shannon, *The art and science of optical design* (Cambridge University Press, 1997).

# 4. NEDs with rotationally symmetric in-coupler

As we have described in the previous chapter, freeform surfaces are needed to compensate the asymmetry caused by the cylindrical mirror. In this chapter, for an equivalent large FOV, we propose a new design with a rotationally symmetric in-coupler in order to reduce fabrication and alignment difficulties. We present the geometry of the system first with its simulated optical performances. Then a stray light analysis is led according to the geometrical parameters of the system. Stray light loss is evaluated by calculation and compared with ray tracing simulation.

## 4.1. Geometry of the whole system

We use a set of right-angle prisms (RAPs) placed along a cylindrical base instead of a cylindrical mirror. The RAPs have no optical power and can reflect light back in the same direction.

**Fig. 4.1** details the proposed system. Two waveguides are stacked up with a small air gap so that light can propagate inside the two waveguides-independently. Light emitted from the micro display is collimated by the in-coupling lenses. The collimated beams then enter into the upper waveguide and propagate by total internal reflection until they reach the RAPs as shown in **Fig. 4.1**(c). In zx plane, the RAPs reflect the incoming light back to the same incident direction while in yx plane they serve as planar mirrors which couple partially light into the lower waveguide as shown in **Fig. 4.1**(b). Finally, the out-coupler reflects different portions of the image towards the eyes with a large vertical FOV as shown in **Fig. 4.1**(d) with the green rays. **Fig. 4.1**(a) presents a perspective view of the whole system with ray tracing.

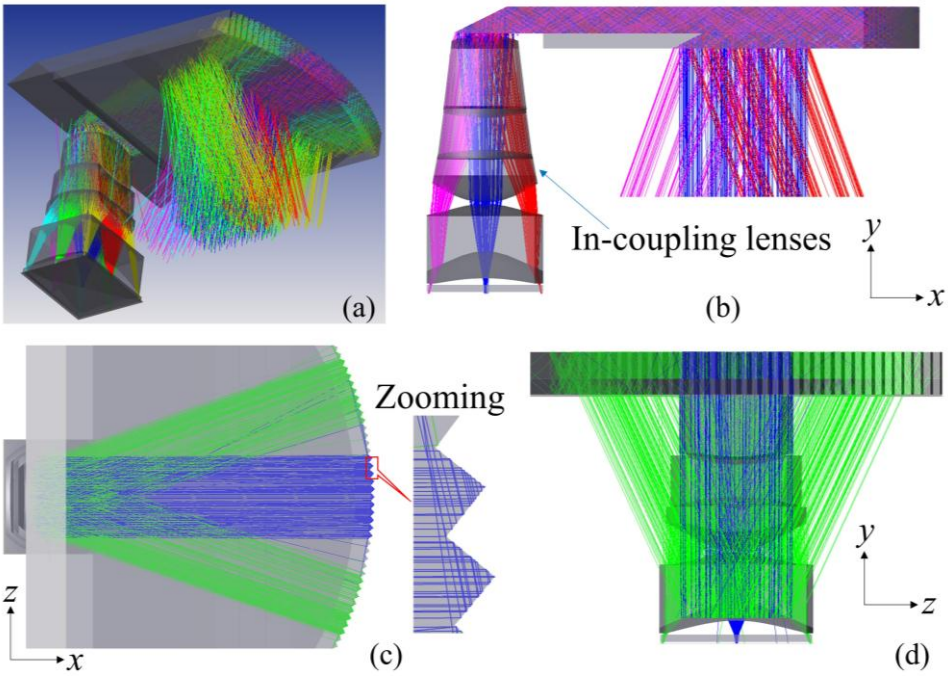



**Fig. 4.1.** Geometry of the system and ray tracing. (a) Perspective view, (b) yx view and the in-coupling lenses, (c) zx view with zoomed right-angle prisms, (d) yz view.

To have an intuitive feeling of the wearing effect, we combine the whole system with a human head model together as shown in **Fig. 4.2**(a) and **Fig. 4.2**(b) for one eye using 3ds Max software. The scale between the head and the NED is kept. For each eye, the in-coupling lenses can be contained inside the sidepiece of the glasses and the whole structure can be very ergonomic.

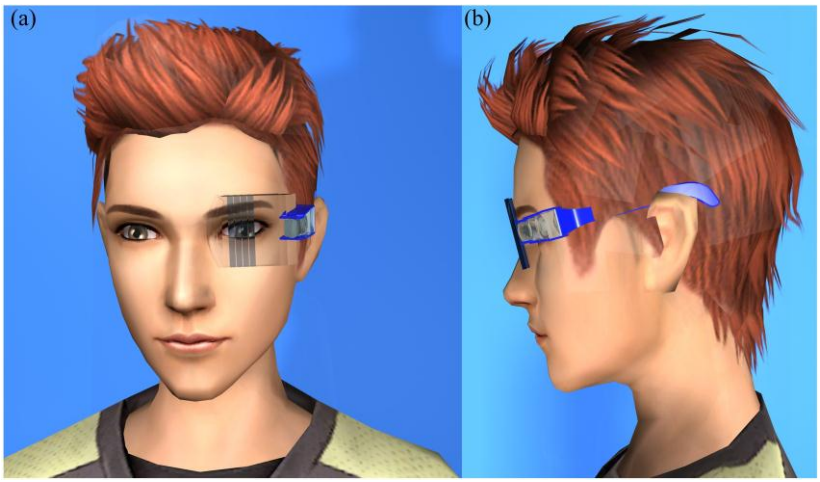



**Fig. 4.2.** Schematic diagram showing wearing effect in different views. (a) front view and (b) side view.

# 4.2. The rotationally symmetric in-coupler

The in-coupling subsystem contains 4 lenses which are labeled L1, L2, L3 and L4 as shown in **Fig. 4.3**(a). L1, L3 and L4 are plastic aspherical lenses made in E48R and L2 is a spherical NSF56 glass lens. All the lenses are rotationally symmetric. This in-coupler is a kind of eyepiece but with small eye relief and lenses with small diameter which mainly depends on pupil size even with large FOV. The pupil size of the in-coupler in z axis is 8 mm, which is the value of vertical eye box, and in x direction is reduced to 4 mm because of the upper waveguide thickness. The unused part of the lenses can be removed to increase compactness and reduce weight as shown in **Fig. 4.1**(a) and (b). Then the pupil size in x direction is expanded by an out-coupler to a larger value, e.g. 15 mm. The total length of the in-coupler is 25 mm and the expected total weight is 4.5 g. The half FOV is 30°. The Root-Mean-Square (RMS) spot radius for three wavelengths and its polychromatic value (unitary weighted average of the monochromatic RMS spot radii) vs field is shown in **Fig. 4.3**(b). All the radii are around 10 μm along the total field. The grid distortion is shown in **Fig. 4.3**(c) with a maximum negative distortion of -7%. Polychromatic MTF is given in **Fig. 4.3**(d), its value is always above 0.3 at 30 lp/mm. According to these simulation results, we can conclude that the in-coupling subsystem can provide a very good image quality. By combining this in-coupler with the two-layer PMMA waveguide, a horizontal FOV of 30° and a vertical FOV of 60° are obtained for the whole system. The vertical size of the eye box along (z axis) is 8 mm while the horizontal size along (x axis) depends on the out-coupler size.

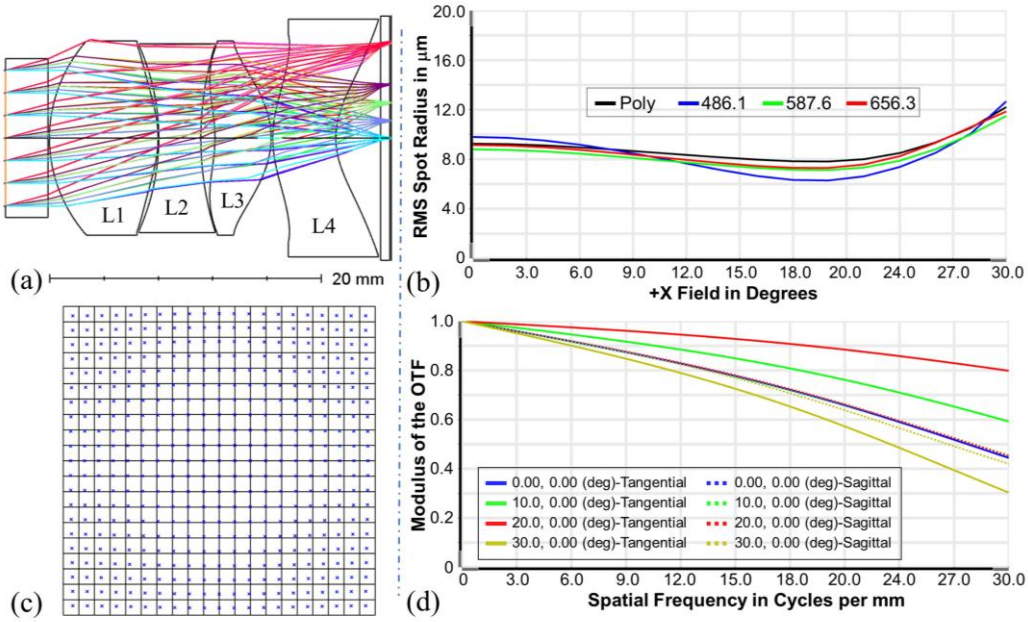



**Fig. 4.3.** Layout and optical performances of the in-coupling lenses. (a) raytracing, (b) plot of Root-Mean-Square (RMS) spot radius vs field, (c) grid distortion and (d) Polychromatic MTF.

# 4.3. Stray light analysis

As the waveguides have no optical power, the optical performance of the whole system is the same as the in-coupler in terms of image quality. However, one factor that can influence the image quality is the stray light caused by the RAPs. In this paragraph we analyze and compute the stray light of the system caused by RAPs. They are used as retroreflectors. As example **Fig. 4.4**(a) shows the ray tracing of a single RAP. According to the geometrical incident conditions, only three surfaces are used to reflect back the incident light. However, some incident rays cannot be reflected by the prism (see red ray in **Fig. 4.4**(a)) and become stray light of the system. The prism can reflect all the light back only in case of normal incidence (the green ray in **Fig. 4.4**(a)). In order to evaluate analytically the loss due to the stray light, we first compute the stray light section of each prism and then compute the ratio of their contribution over the total projected pupil size.

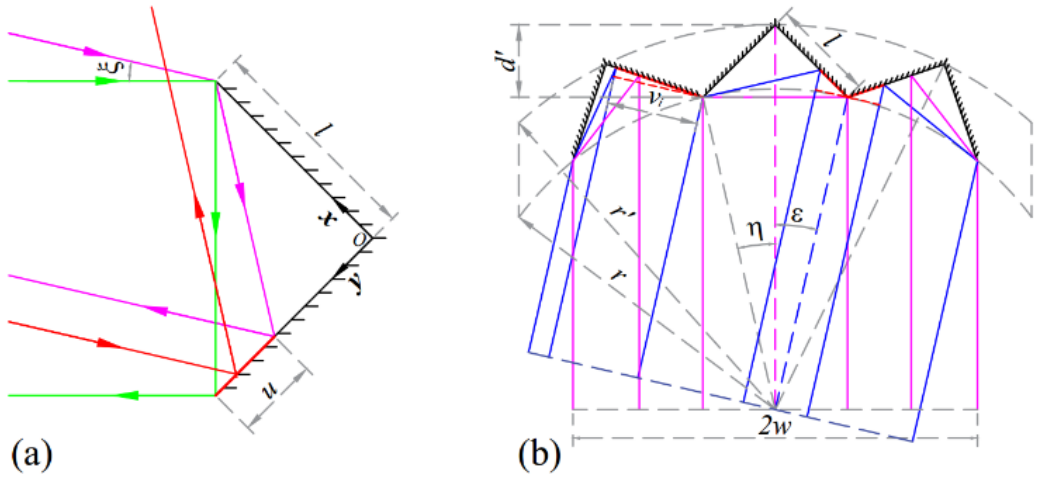



**Fig. 4.4.** Ray tracing through a right-angle prism in a plane. (a) single retroreflector function ray tracing, (b) ray tracing and stray light analysis of the set of RAPs (drawing not at scale).

Let the length of the reflective plane segments be $l$ and the incident angle with respect to the normal direction be $\xi$. The stray light section on the mirror is:

$$u = l\left(1 - \tan\left(\frac{\pi}{4} - \xi\right)\right) \tag{4-1}$$

In the proposed NED system, the chief ray of each field goes through the center of the in-coupling pupil. The set of prism vertices forms an inner and outer circle of radius $r$ and $r'$ respectively, with the same center. As the in-coupling pupil in $z$ direction has a length $2w$ as shown in **Fig. 4.4**(b), the total contribution of the stray light is the sum of all the stray light sections of each prism. Angle $\varepsilon$ indicates the angle of rays diverging (the blue rays in **Fig. 4.4**(b)) from the (0°,0°) central field. For the convenience of analysis, we assume that there is an odd number $N$ of half-prisms within a section size of $w$, with a minimum value of 3. Hence, the half central angle $\eta$ of each prism (see **Fig. 4.4**(b)) is calculated by:

$$h = \frac{\arcsin\left(\frac{w}{r}\right)}{N} \tag{4-2}$$

The length of the reflective mirror is:

$$l = \sqrt{2}r\sin(\eta) \tag{4-3}$$

For NED systems, $r$ is around 40-60 mm, while $w$ is about 5-8 mm. To simplify the calculation, we suppose that an integer number of prisms are used within the pupil. Due

to the circular symmetry, we need only to consider the case where $\varepsilon \leq \eta$. We note $v_i$ the stray light cross section for each prism with one hit as shown in **Fig. 4.4(b)**. To compute it, we consider separately the stray light effect on the left and the right parts of each prism. The ratio of the stray light of each field is then:

$$R_t = \frac{\sum_i v_i}{2w\cos(\varepsilon)} = \frac{\sum_{i=1}^{\frac{N-1}{2}} K_i + \sum_{i=0}^{\frac{N-1}{2}} M_i}{2w\cos(\varepsilon)} \quad (4\text{-}4)$$

in which $K_i = l\left[\left(1-\tan\left(\frac{\pi}{4}-2i\eta+\varepsilon\right)\right)\cos\left(\frac{\pi}{4}-2i\eta+\varepsilon\right)\right]$ is the light cross-section of ith left half prism;

$M_i = l\left[\left(1-\tan\left(\frac{\pi}{4}-2i\eta-\varepsilon\right)\right)\cos\left(\frac{\pi}{4}-2i\eta-\varepsilon\right)\right]$ is its equivalent right side.

The stray light ratio for four different fields: $\eta$, $\eta/2$, $\eta/4$ and 0, with $w$ equals to 5mm and $r$ equals to 40, 50 and 60 mm is given in **Fig. 4.5**. This ratio is mainly depending on the value $r/w$. For different fields it tends to the same value when $N$ increases. A given stray light ratio imposes a minimum number of prisms. For instance, if we suppose r=50mm with a stray light loss of 5%, we should use at least about 4 prisms. Considering these values for our proposed NED system, we can conclude that the stray light loss is limited and not critical.

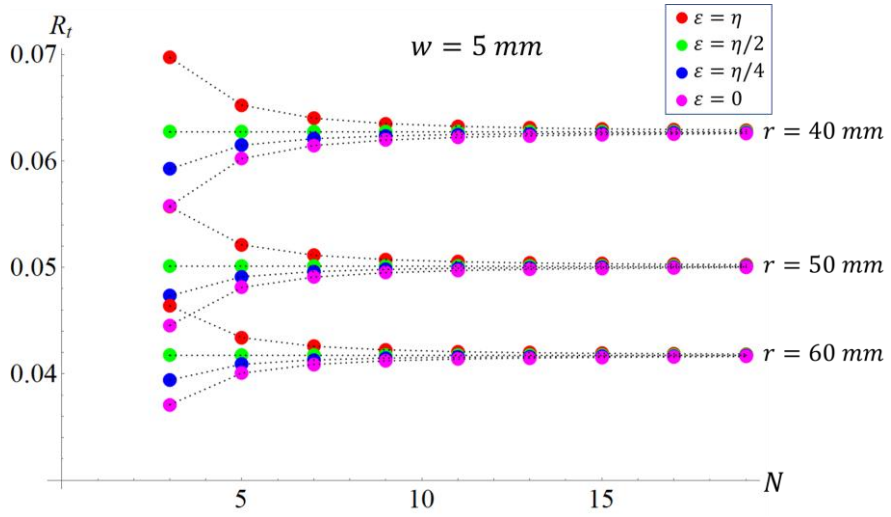



**Fig. 4.5.** Stray light ratio with respect to number of half RAPs for r=40 mm, r=50 mm and r=60 mm when w=5 mm. Different colors represent different $\varepsilon$ as shown in top right corner.

Eq. (4-4) calculates the stray light ratio assuming a perfect coupling between the waveguides. Ray tracing simulation permits to take into account this coupling. We have used Zemax in non-sequential mode to simulate $R_t$ with respect to different incident angles $\theta$ (defined in **Fig. 3.6**) of the object field as shown in **Fig. 4.6**.

To simplify the simulation, we have considered as out-coupler a single plane mirror. A polar detector is placed after the out-coupler to record angular energy distribution. Here $R_t$ is calculated by:

$$R_t = (E_{tot} - E_d)/E_{tot} \quad (4\text{-}5)$$

where $E_{tot}$ is the total energy received by the detector and $E_d$ is the energy in the desired direction. For simulation, we have taken r=52.72 mm, $w$=5 mm, $l$=1 mm, n=7, the length of the transition region (common area between waveguides) is 2.72 mm and the thickness of the upper and lower waveguides are 2.72 mm and 1.5 mm respectively. The calculated value without considering coupling is about 5% as shown in **Fig. 4.5**, which is different from the simulated result. However, they are on the same order of magnitude. The average simulated $R_t$ values for $\varepsilon$=0 and $\varepsilon=\eta$ are around 4%. $R_t$ is oscillating with different incident angles.

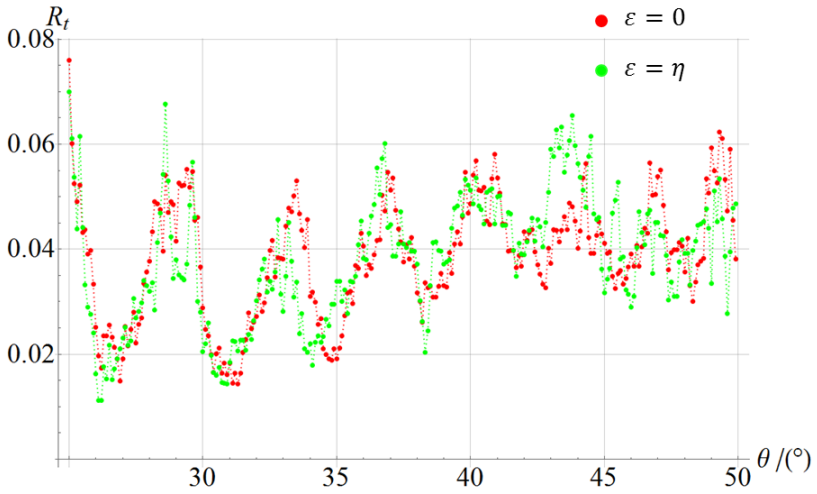



**Fig. 4.6.** Simulated result of stray light ratio with respect to incident angle.

With these geometrical conditions, the calculated stray light loss is limited to 5% where more than 4 prisms are used. Moreover, this value is very close to the simulated value even if the coupling between waveguides is not perfect.

# 4.4. Conclusion

In conclusion, we have presented a new ergonomic and compact see-through NED design. The introduction of a set of right-angle prisms along a cylindrical base has permitted to suppress any optical power in the propagation of object fields inside the two waveguides and the out-coupler. Then the in-coupling subsystem uses only rotationally symmetric lenses. The fabrication and design difficulties are largely reduced. If, for instance, a Lumus type out-coupler is used, a large eye box (H 15mm ×V 8 mm) and FOV (H 30° × V 60°) can be obtained. In this case, the image quality is sufficient with polychromatic MTFs always higher than 0.3 at 30 lp/mm. Moreover, the stray light loss has been also investigated by calculation and simulation. Its value is limited to 5% where more than 4 prisms are used which is not critical for NED devices.

# 5. Experimental results

We have realized the two waveguides and the transition edge with right-angle prisms (RAPs) described previously. This has allowed us to do preliminary tests of the system and evaluate the characteristics of each element. As we have discussed in the previous chapter, the rotationally symmetric system has the advantage of avoiding freeform optical surfaces. We have decided to fabricate the stack of waveguides and the edge element based on our design. Therefore, we have ordered the fabrication of an upper waveguide, a lower waveguide and a transition edge. One end of the lower waveguide was bevelled and coated to function as a simple out-coupler (without pupil expanding function). The in-coupler generated beams were replaced by collimated beams.

In this chapter, we firstly present the fabricated components (two waveguides and a transition edge) and their characteristics. Then the experimental results are presented.

## 5.1. Fabricated components and their characteristics

### 5.1.1. Waveguides

A set of three pairs of the two waveguides were ordered and fabricated by Polymer Optics Ltd (see **Fig. 5.1**). The corresponding technical drawings are shown in **Fig. 5.2**. The unit is in mm.

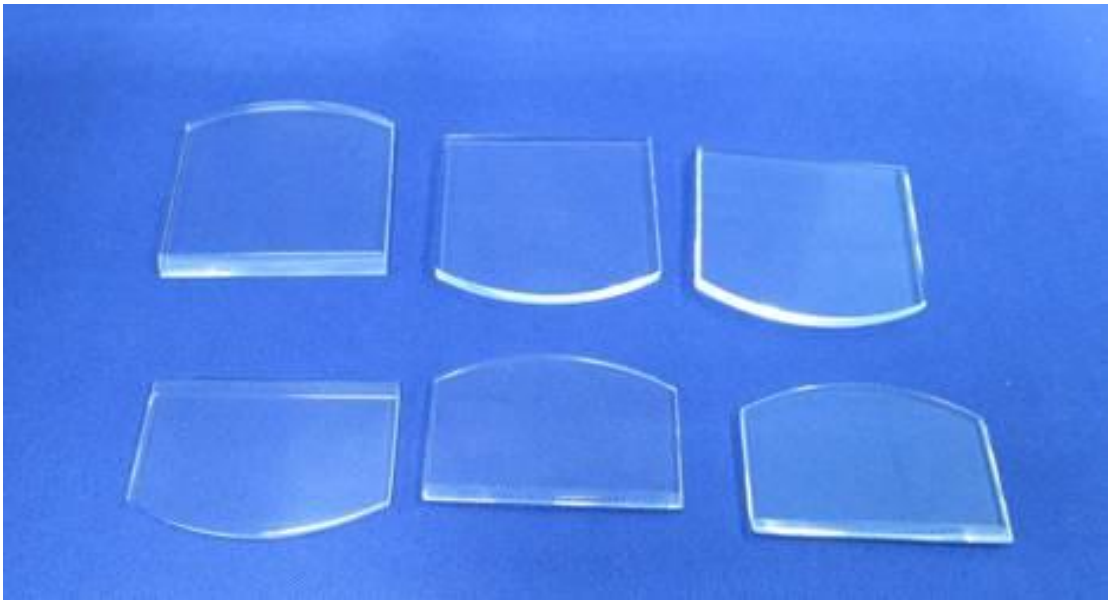


**Fig. 5.1.** The two waveguides have been manufactured in three pairs.

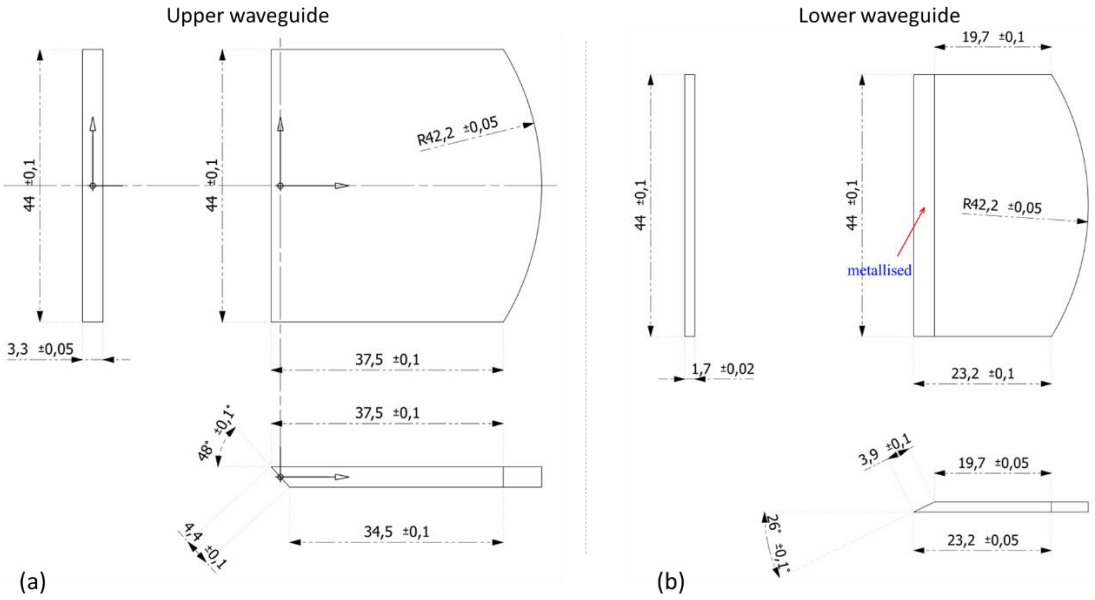



**Fig. 5.2.** Dimensions of the manufactured waveguides (unit in mm).

The measured dimensions are shown in **Table. 5.1** and **Table. 5.2**. The see-through transmittance was measured with an incident light perpendicular to the waveguide. Several wavelengths were used, and the given see-through transmittance is the average value. Moreover, we have noticed no chromatic effect by direct see-through vision.

**Table. 5.1.** Measured dimensions and see-through transmittance of the upper guide.

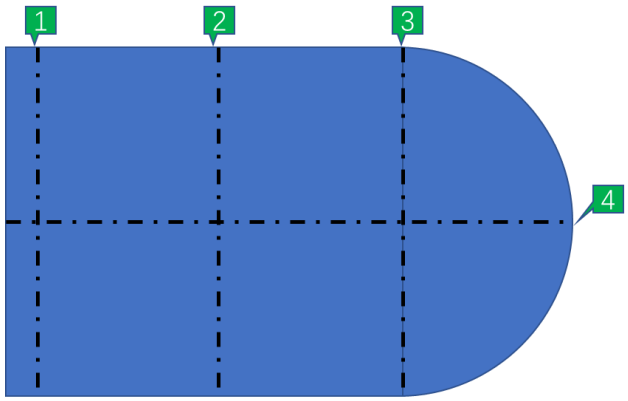
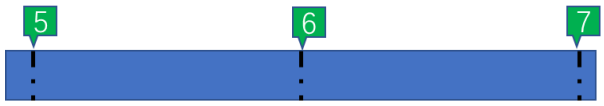
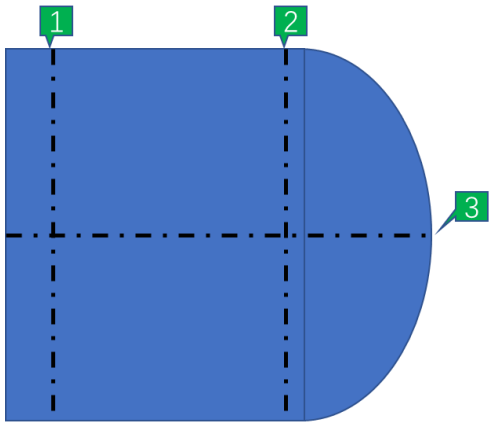
| | |
|---|---|
|  | **Measurements (mm)**<br>**Nominal Width:** 44±0.10<br>**Line 1:** 44.01±0.01<br>**Line 2:** 44.01±0.01<br>**Line 3:** 43.95±0.01<br><br>**Nominal Length:** 43.76±0.10<br>**Line 4:** 43.6±0.3<br><br>**Transmittance:** 91% |
|  | **Measurements (mm)**<br>**Nominal Height:** 3.3±0.5<br>**Line 5:** 3.34±0.02<br>**Line 6:** 3.34±0.02<br>**Line 7:** 3.33±0.02 |

**Table. 5.2.** Measured dimensions and see-through transmittance of the lower guide.

| | |
|---|---|
|  | **Measurements (mm)**<br>**Nominal Width:** 44±0.10<br>**Line 1:** 42.75±0.01<br>**Line 2:** 42.72±0.01<br><br>**Nominal Length:** 27.7±0.10<br>**Line 3:** 28.0±0.3<br><br>**Transmittance:** 86% |
|  | **Measurements (mm)**<br>**Nominal Height:** 1.70±0.02<br>**Line 4:** 1.67±0.02<br>**Line 5:** 1.68±0.02 |

In conclusion, the nominal dimensions of the two waveguides have been mostly respected. The transmittance is around 90% because of the reflection on the two

surfaces of the waveguide and the absorption of the material. For this stage of evaluation, the waveguide was fabricated without any antireflection coating. One set of waveguides was reserved for the final set-up. A set was dedicated to intermediate tests and one set to test the gluing process for the assembling.

The roughness of the surfaces for both waveguides has been measured with the Zygo interferometer (NewView 7200) of the IPP team of the ICube Laboratory (see snapshots **Fig. 5.3**). The roughness average was about Ra=11nm (root mean square value about Rms=15nm) for the surfaces of the two sets put in reserve and higher for the test set used for gluing (Ra=18nm and Rms=26nm). For sake of comparison we have also measured the roughness average of a glass waveguide of Lumus type (see **Fig. 5.4**). The Lumus type waveguide (described in ch. 2) is an out-coupler composed of a set of selective reflective coatings. The measured values were about Ra=2.4nm (Rms=2.9nm) in a region outside the mirror area and about Ra=2.5nm (Rms=3.1nm) in the mirror area. We have noticed a factor 5 in Ra between the two materials.

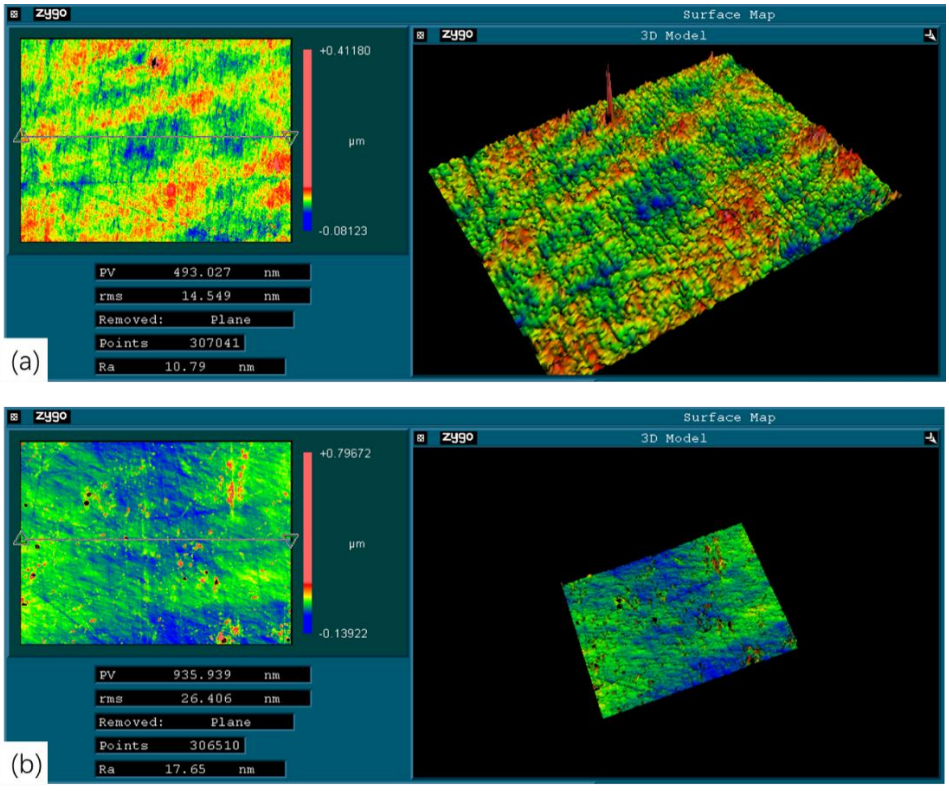



**Fig. 5.3.** Snapshots of the roughness measurements of the PMMA upper waveguide (a) put in reserve and (b) used guides.

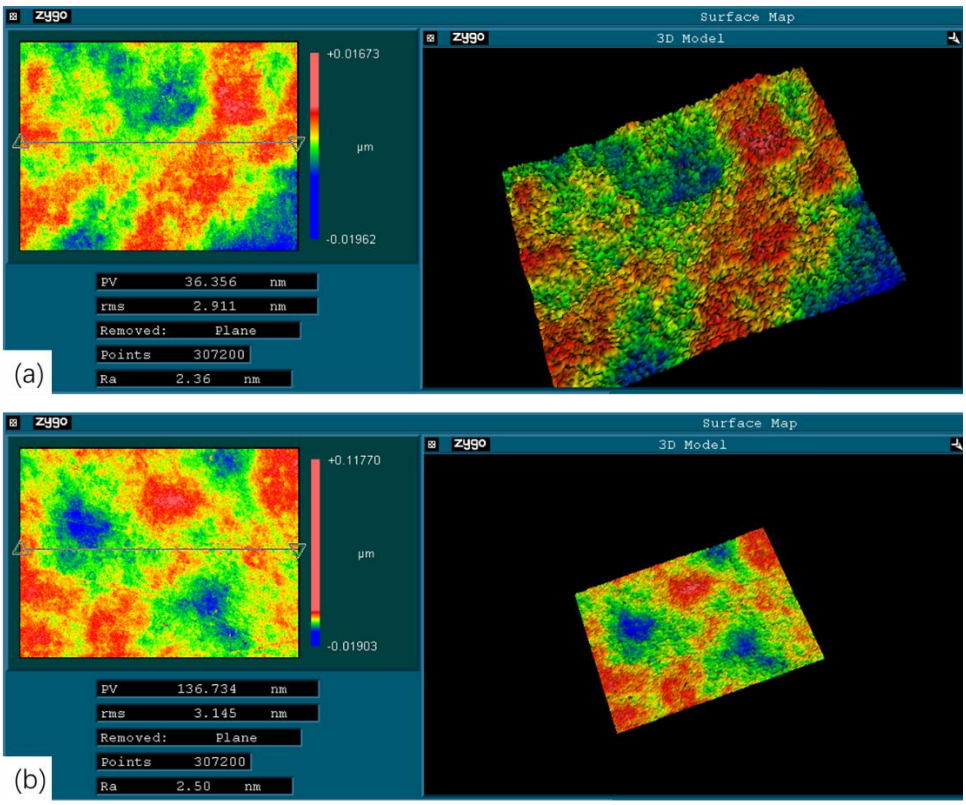



**Fig. 5.4.** Snapshots of the roughness measurements of the Lumus glass waveguide (a) Outside the mirror area and (b) Inside the mirror area.

After having presented the dimensions and surface roughness of the two fabricated waveguides, we present the characteristics of a fabricated transition edge.

## 5.1.2. Transition edge

We have searched for a great number of companies involved in PMMA fabrication. Only a few have answered to our quotation request. In a first try with a Chinese company more specialized in mechanical parts, the roughness average was too big (about 10 microns). Then we have tested another company more involved in optical elements manufacturing (ARRK company). The designed element is described in **Fig. 5.5**.

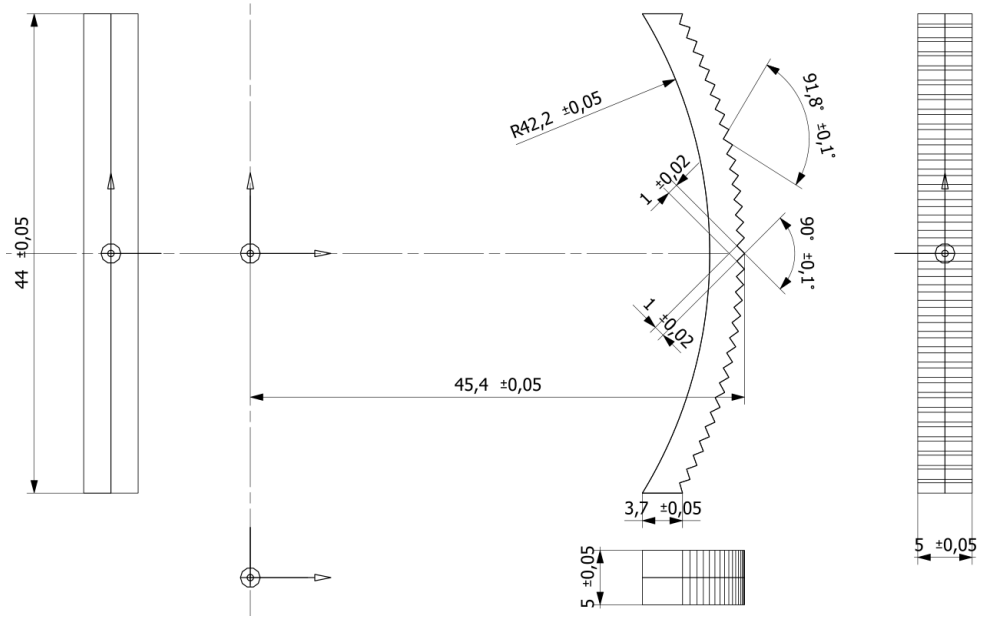



**Fig. 5.5.** Dimensions of the transition edge (unit in mm).

After the manufacturing, we have measured the top angle of some teeth and unfortunately the angle value and its tolerance were not respected (see the **Fig. 5.6** with a top angle around 95.4° and 93.9° instead of 90°).

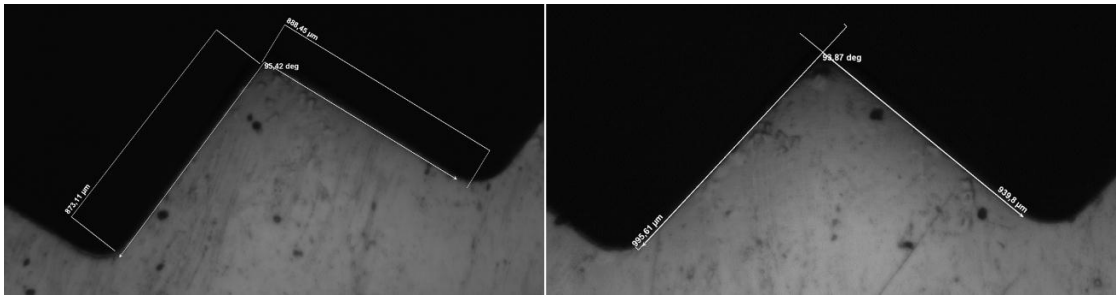



**Fig. 5.6.** Top angle measurement of some teeth.

Moreover, a surface roughness measurement of the transition edge has been realized. We have measured two locations, one is on the top surface outside the teeth and the second one inside a tooth. For the first location, the result was about Ra=11.4nm (Rms=20nm), which is similar to the manufactured waveguides. For the second location, the result was about Ra=184nm (Rms=228nm) as shown in **Fig. 5.7**. This average roughness is about 75 times higher than the glass roughness. This result indicates that the transition edge is the main challenge for the manufacturing. Because of the large roughness on the tooth region, we have detected diffused light when we assembled the whole system, which has prevented us to see a clear spot.

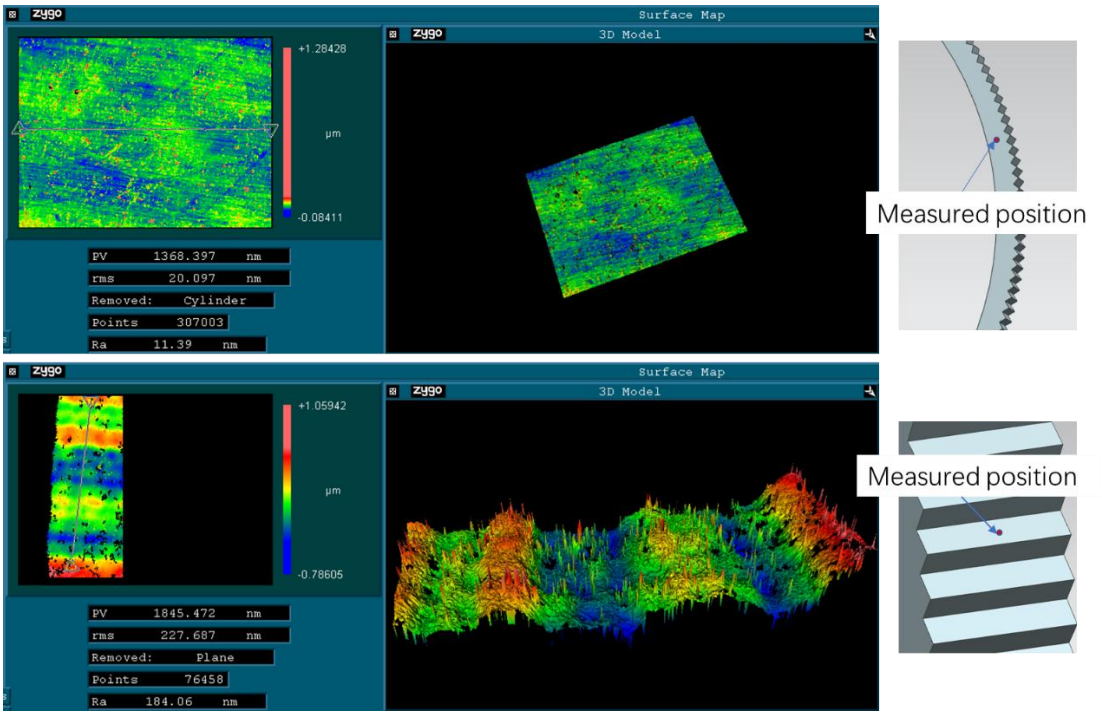



**Fig. 5.7.** Surface roughness measurement of the transition edge in two locations.

With these fabricated components, we carried out optical measurements and the result are presented in the next section.

## 5.2. Optical experiment results

As the surface inside the sawteeth is very rough with Ra=184nm (Rms=228nm), scattering effects have been seen. Therefore, to reduce these effects, we firstly measured the optical response with a single transition edge. The experimental results are shown in **Fig. 5.8**. The transition edge is illuminated from the inner side by a white light plane wave with an angle about 10° as shown in **Fig. 5.8**(a). Then by the reflection of the RAPs, the output light is received by a detector in different distances d≈5 cm, d≈10 cm and d≈15 cm as shown in **Fig. 5.8**(b), (c) and (d). Ideally, the spot size should keep the same size. We notice that the spot size becomes larger with increasing of distance; this is due to scattering by the transition edge surfaces.

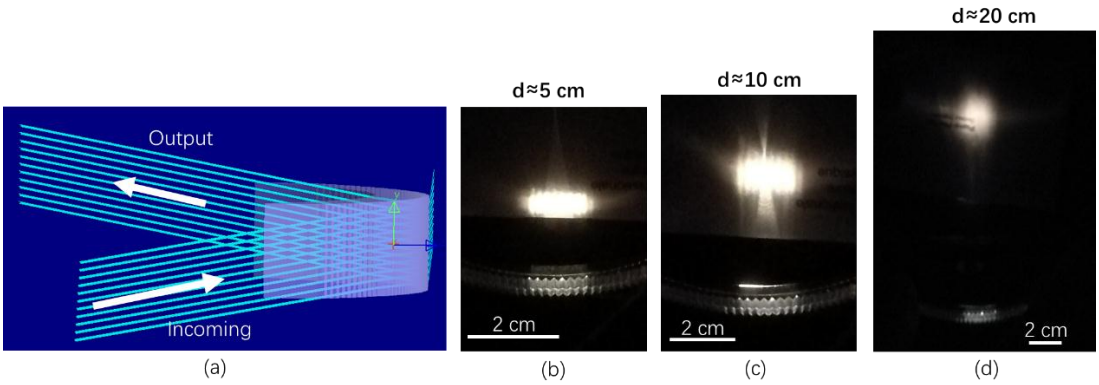



**Fig. 5.8.** Experimental analysis for a single transition region. (a) Optical setup, (b) light output received at a distance about 5 cm. (c) light output received at a distance about 10 cm. (d) light output received at a distance about 15 cm.

To measure the whole system, we have built a setup as shown **Fig. 5.9**. We have used two kinds of light sources: An Ocean Optics HL 2000 HP Tungsten Halogen white light source (see its characteristics in appendix A) and a He-Ne laser. The function of the collimator is to collimate the light from the source to the upper waveguide and the aperture is used to obtain the right beam size. The holder is positioned where the tested system is placed. The camera is used as a receiver to obtain total received energy (see its characteristics in appendix B).

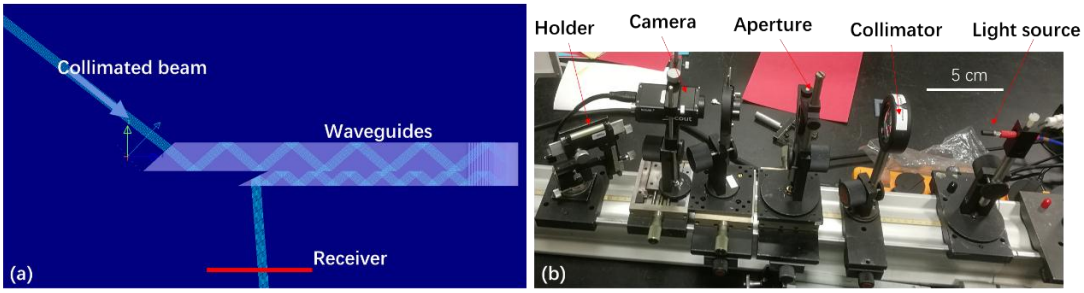



**Fig. 5.9.** Setup for measurement. (a) Schematic diagram for the measurement and (b) Experimental setup for the normal incident white source intensity measurement.

The waveguides and the transition edge have been glued. The scattering effects obtained from the out-coupling mirror can be clearly seen as shown in **Fig. 5.10**. These effects were enhanced by the speckle of the laser (worse case). The surfaces of the waveguide are also scattered but less than the transition edge. Because of the scattering, we cannot get a spot as we expected. However, we can still distinguish an intensity peak.

With different illumination angles, we can receive intensity peaks also in different angles, which respects the raytracing analysis.

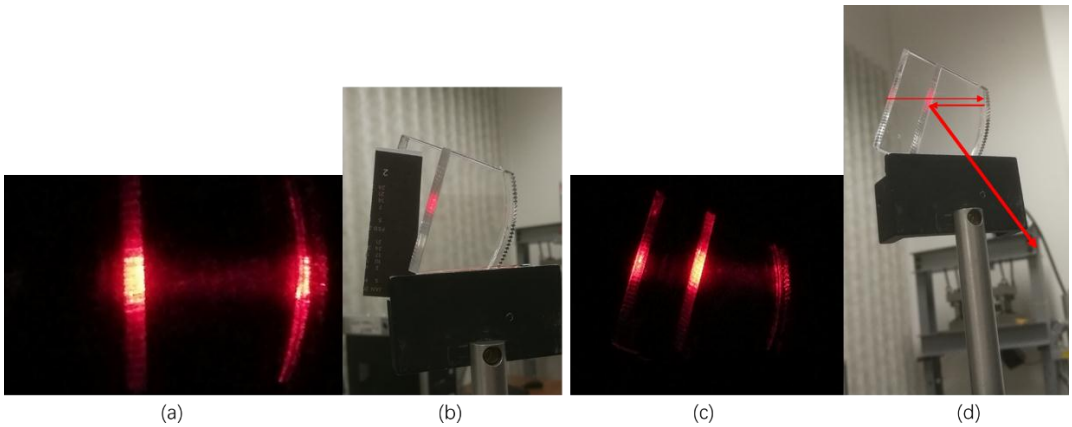



**Fig. 5.10.** Optical response of the whole system. (a) illuminated with small angle without ambient light. (b) illuminated with small angle with ambient light. (c) illuminated with large angle without ambient light. (d) illuminated with large angle with ambient light.

As the transition edge scatters essentially in a vertical plane, we have also tried other ways that may prove the concept. We have used a 5 mm right angle prism directly glued to the end of the two waveguides to serve as the transition edge. However, only three surfaces were well polished as shown in **Fig. 5.11**(a). The top and bottom surfaces were unpolished, and the scattering effect is also visible. Because of the time limitation for the pre-maturation program (ended in 2017), we could not order rapidly new prisms with all the 5 surfaces polished and we have decided to polish by ourselves as shown in **Fig. 5.11**(b) before to receive polished prisms (two month of delay). However, it was difficult to obtain a very flat optical surface, also we have glued a very thin flat glass on the prism and let the glue filled the irregular space.


**Fig. 5.11.** (a) A 5 mm prism with top and bottom surfaces unpolished and (b) hand polish setup and the polished surface.

Finally, we have glued the modified prism and the waveguides together to test the optical response as shown in **Fig. 5.12**. We can see very bright output power when we take picture directly. The intensity is much larger than with a transition edge as in **Fig. 5.10**. In this case, we can detect the light by a screen as shown in **Fig. 5.12**(b). To calculate the efficiency, we have used a monochromatic camera: Basler-Scout-sc640-70fm, to measure the total energy. As a result, we can estimate that the efficiency is between 10% and 20%. There are several possible explanations for the loss of efficiency.

a> All the elements that have no antireflective coatings.

b> The surface roughness of the waveguides is too high. Some light is scattered in each total reflection, but it is not the main source of loss in some applications it could be acceptable as it is.

c> The home-made polished surfaces are not good enough and cause scattering of light.

d> The absorption of the optical material (quite low) diminishes the efficiency.

e> The quality of the gluing is not very good, and we can see bright light in the glued sections.

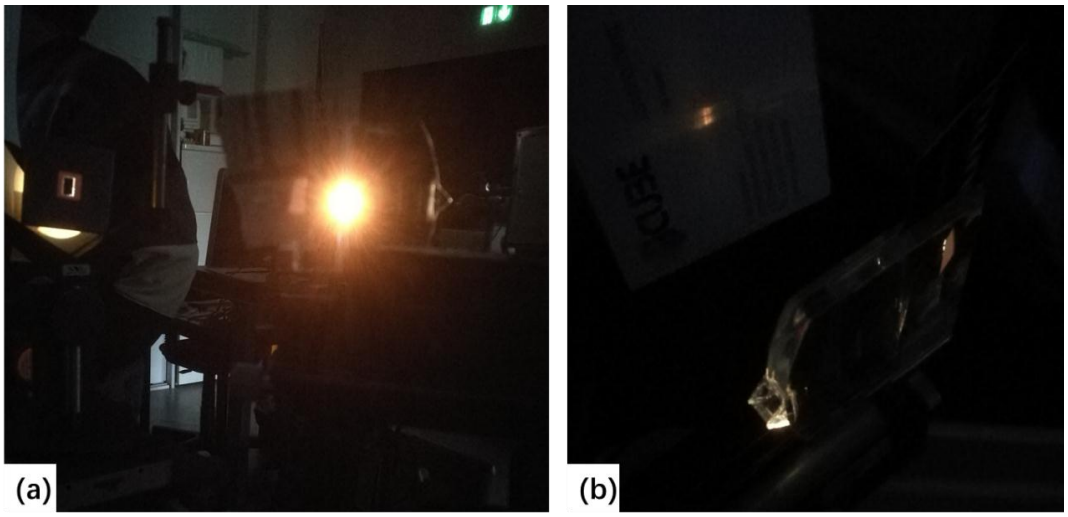



**Fig. 5.12.** Optical response of the system with a 5 mm RAP. (a) direct observation and (b) received by a screen.

# 5.3. Conclusion

In conclusion, the waveguides have been ordered and manufactured with acceptable dimension tolerances. However, a lower surface roughness close to the glass plate roughness is preferable. The main trouble has come from the transition region with sawteeth. The associated roughness was too high for our requirements and until now the manufacture of this element is still a challenge. An experimental setup has been built to measure the optical response of the element alone and the whole system. However, by using a 5 mm RAP, we have observed improved and increased light intensity with the right output angles. However, for the moment, the total efficiency is only between 10 to 20%. The next step will consist in fabricating the transition region with our requirements in angle and roughness. In the same time, the industrial manufacturing of the rotationally symmetric in-coupler, the mechanical mounting and the assembly process will be investigated.

# 6. Study of an in-coupling subsystem using a harmonic lens

The previous sections mainly deal with refractive and reflective optics. However, the conventional lenses are bulky compared to diffractive optical elements (DOEs). Lightweight and compact components are desired in NEDs. In this section we will discuss the possibility of using particular DOEs in our designed system: the harmonic lenses. We begin by presenting the interest of Harmonic Lenses (HLs) and the useful tools for the simulation. Then we detail the principle of these lenses and give a parameter study according to their F-numbers.

## 6.1. Interest of harmonic lenses

As we consider polychromatic object sources, three main solution series can be investigated: stacks of diffractive elements (at least one DOE for each principal color), achromatized diffractive elements (a combination of optical materials with different Abbe numbers), and harmonic lenses. The stacks of diffractive elements (slanted gratings and holographic elements) have been studied in depth [1-9]. They are now used as out-couplers in Hololens device for instance. The achromatization of diffractive elements have been considered in various imaging applications [10-13], but their applications in micro imaging devices with severe constraints are still a challenge. We have rather decided to consider a special family of DOEs called harmonic Fresnel lenses (HFLs) [18]. These elements permit to obtain the same phase transfer function for several specific wavelengths. Specifically, we have investigated the possibility to use HL as a part of the in-coupling subsystem. In literature HLs have been considered with paraxial conditions, which correspond to large F-number systems. But in NEDs, small F-numbers (f/#) are desired, also in the following paragraphs we will consider mainly

these systems. The dimensions of conventional lenses are large compared to the wavelength so that ray tracing is an accurate method to simulate. For DOEs, as the structure size is comparable with wavelength, more advanced simulation methods like Finite-Difference Time-Domain method (FDTD) are needed. To obtain accurate simulation results, we have applied FDTD method to calculate the electromagnetic near field and used the angular spectrum method (ASM) to calculate its far field distribution [14-17]. The advantage of this approach is an efficient computation with accurate results whatever the incident angle is. In all of our simulations, the propagating medium after the lens is supposed to be air. As FDTD is a time-domain technique, when a broadband pulse is used as the source, then the response of the system over a wide range of frequencies can be obtained with a single simulation. This property is useful in our study because we investigate the broadband properties of HLs.

# 6.2. Principle of the harmonic lens [18, 19]

In the case of the paraxial approximation, the maximum effective thickness of an ordinary Fresnel lens is [13, 20, 21]:

$$h_{\max} = \frac{\lambda_0}{n-1} \tag{6-1}$$

where $\lambda_0$ is the design wavelength and $n$ is the refractive index of the lens material. For a Harmonic Fresnel Lens (HFL) with a harmonic number $p_h$, the thickness is:

$$h_{\max} = p_h \frac{\lambda_0}{n-1} \tag{6-2}$$

The harmonic wavelengths are a set of discrete values given by:

$$\lambda p = \frac{p_h \lambda_0}{m} \tag{6-3}$$

where $m$ is a positive integer that equals the diffractive order. In the case of the paraxial approximation, the dependence of the focal length with wavelength $\lambda$ can be expressed as:

$$f(\lambda) = \frac{p_h \lambda_0 f_d}{m\lambda} \tag{6-4}$$

where $f_d$ is the aimed focal length.

The scalar diffraction efficiency of the $m$th diffracted order is given by:

$$\eta_m = \sin c^2\{\frac{p_h\lambda_0[n(\lambda)-1]}{\lambda[n(\lambda_0)-1]} - m\} \quad (6\text{-}5)$$

where $\lambda_0$ is the design wavelength and $\eta_m$ is the diffraction efficiency of the $m$th order, which equals to 1 when $\lambda$ satisfy Eq. (6-3) and assuming that the index $n$ is constant in the visible spectral range.

An ordinary Fresnel lens corresponds to a HFL with $p_h$=1. As $p_h$ increases, the height of the lens becomes higher and each groove larger. **Fig. 6.1** shows this effect. As example, **Table. 6.1** gives numerical values of height and harmonic wavelength for a F/#=1 HFL with different harmonic numbers. The lens material is supposed to be NBK7 glass from Schott, the aimed focal length is 0.5 mm and the design wavelength 500nm. Comment (As $p_h$ increases the number of harmonic wavelengths in the visible domain also increases but with higher diffractive orders.)

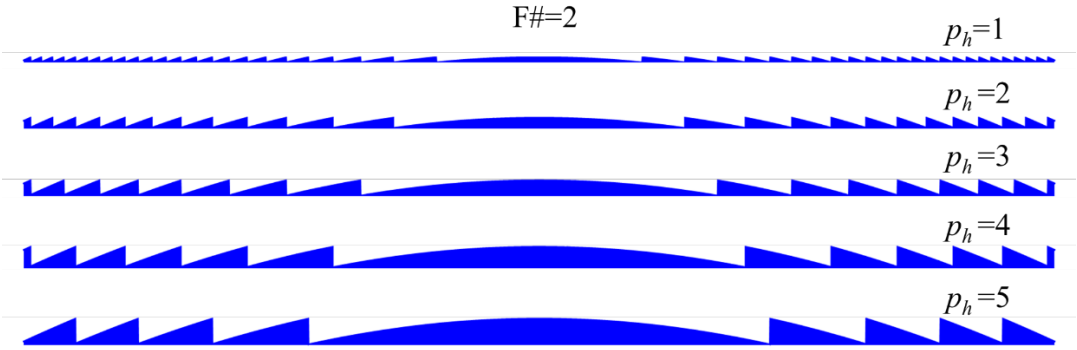



**Fig. 6.1.** Modification of the height and width of grooves when the harmonic number of a HFL increases, with a constant F/#=2.

**Table. 6.1.** Numerical values of height and harmonic wavelength for a F/#=1 HFL with different harmonic numbers.

| $p_h$ | $h_{max}$ ($\mu$m) | (m, $\lambda_p$ in nm) |
|---|---|---|
| 1 | 0.998 | Visible domain: [(1, 500)]<br>(2, 250) (3, 166.7) (4, 125) |
| 5 | 4.951 | (1, 2500) (2, 1250) (3, 833.3)<br>Visible domain: [(4, 625) (5, 500) (6, 416.6)]<br>(7, 357.1) (8, 312.5) |
| 10 | 9.809 | (1, 5000) (2, 2500) (3, 1666.7) (4, 1250)<br>(5, 1000) (6, 833.3) (7, 714.3)<br>Visible domain: [(8, 625) (9, 555.6) (10, 500) (11, 454.5) (12, 416.7)] |

**Fig. 6.2** shows a detailed cross-section of a HFL. Two zones with rotational symmetry are presented. Point O is the center of the zone numbered 0 which is defined by the semi-section OA'. Points A and B' define the zone numbered 1. To obtain a constructive interference, the radius of the first zone should be:

$$r_1 = \sqrt{2fp_h\lambda_d + (p_h\lambda_d)^2}$$

By generalization, the radius of a zone numbered $i$ is given by:

$$r_i = \sqrt{2ifp_h\lambda_d + (ip_h\lambda_d)^2} \quad (6\text{-}6)$$

where $f_d$ is the aimed focal length of the HFL, $p_h$ is the harmonic number and $\lambda_d$ is the design wavelength. Moreover, if $O$ is the origin of coordinates, the equation of the profile in each zone is:

$$x_i(y) = \frac{f - n\sqrt{f^2 + r_i^2} + \sqrt{f^2(1+n^2) + (n^2-1)y^2 + r_i^2 - 2fn\sqrt{f^2 + r_i^2}}}{1-n^2} \quad (6\text{-}7)$$

which is a part of hyperbola. The Harmonic wavelengths are still determined by Eq. (6-3). The maximum groove depth decreases as the groove number increases, which means A'A>B'B in **Fig. 6.2**.

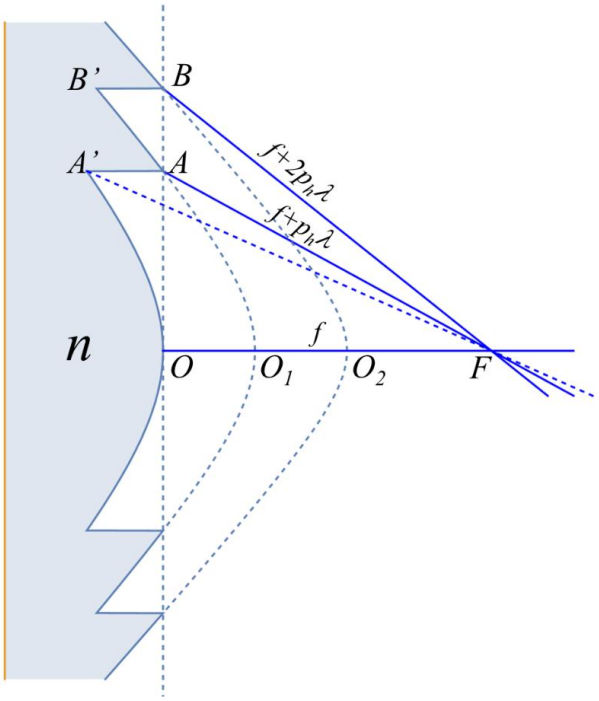



**Fig. 6.2.** Principle of a harmonic lens with a harmonic number equals to $p_h$.

A conventional perfect refractive lens composed of one plane surface and one hyperbola can focus perfectly an incident normal plane wave to the image focal point. The asymptotic line of this hyperbola determines the minimum F number as shown in **Fig. 6.2** which can be calculated by:

$$F/\#_{\min} = \frac{1}{2\sqrt{n^2-1}} \quad (6\text{-}8)$$

For a Fresnel lens, each zone numbered $i$ is a hyperbola with a different vortex $O_i$ (see **Fig. 6.2** with the zones 0, 1 and 2). Physically, the vortex should not overpass the image focal point $F$, which determines the minimum F/# of a HFL. The vortex coincides with F when:

$$r_i \geq f\sqrt{n^2-1} \quad (6\text{-}9)$$

which gives the same minimum F/# as Eq. (6-8). As an example, if n=1.5, the minimum F/# is 0.45.

As NEDs require low F-number without paraxial conditions, we will now simulate with FDTD and ASM the response of the HFL, with different F-numbers, to an incident plane wave.

# 6.3. Optical properties of the HFL

To obtain rigorous results, for all the simulations, the electromagnetic near-field distribution is obtained numerically by Lumerical FDTD. We use a maximum mesh size of $\lambda/(15n)$, where $n$ is the refractive index of the corresponding medium. We apply perfect matched layers (PMLs) as the boundary conditions in the four sides. For the computation of the far-field distribution we use ASM. **Fig. 6.3**. shows an example of output for a HFL with $p_h$=10, F/#=1, the diameter of the lens is 400μm and $\lambda_d$ =0.5 m. In the left part the geometry of the HFL is presented with the intensity distribution of the electric field inside the lens. In the right part, the intensity distribution of the electric field is given with the position of the focal image area. Several HFLs with the same diameter $D$=400 μm but different $F/\#$ have been investigated. The designed wavelength is $\lambda_d$=0.5 μm and the calculated spectral range is from 0.4 to 0.7 μm with a step size of

0.001 μm, which covers the visible spectrum. At the beginning and for sake of simplicity, we have supposed that the refractive index of the HFL is $n_d$=1.5 at all wavelengths. We define the focusing efficiency as the energy inside the central lobe over the total illumination energy, the focal length as the length between the external central point of the HFL and the maximum intensity position inside the focal spot.

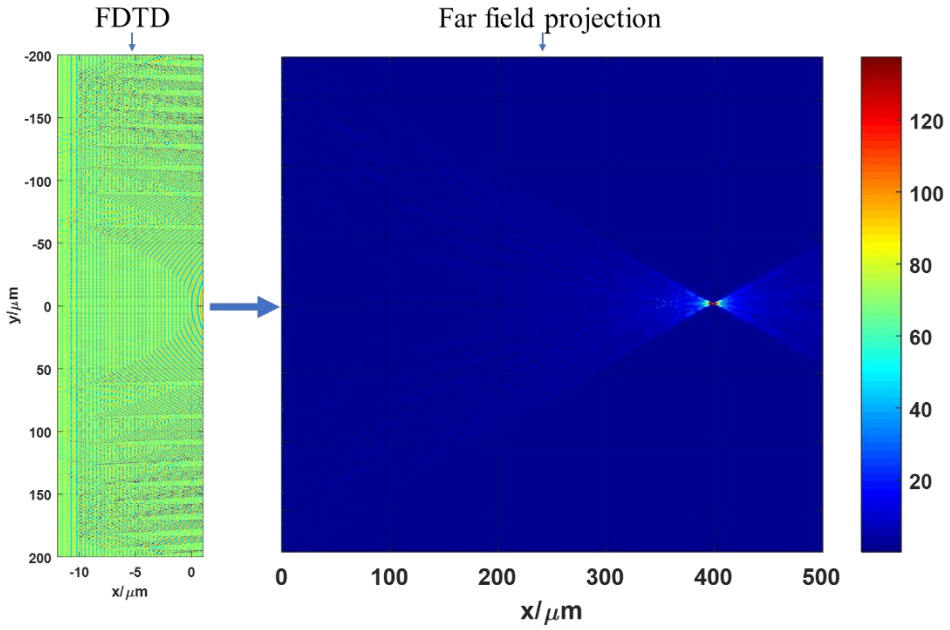



**Fig. 6.3.** FDTD approach combines with far field projection for a HFL with $p_h$=10, F#=1, D=400μm and $\lambda$=0.5 μm.

The focus efficiency, focal length, maximum intensity and the full width at half maximum (FWHM) vs wavelength are shown in **Fig. 6.4** for unitary s-polarized normal incident plane wave. All the values are given for the maximum intensity position. Three different HFLs with *F/#*=0.5, *F/#*=1 and *F/#*=3 have been simulated. A different color corresponds to a different harmonic number $p_h$.

Firstly, we notice that the harmonic wavelengths are independent of the *F/#* (see **Fig. 6.4**), and fulfill Eq. (6-3) in all cases. The focal length also verifies Eq. (6-4). Then, as the design wavelength is 0.5 μm, we obtain peak efficiency at this wavelength in all cases. The case when $p_h$=1 is an ordinary kinoform lens and we obtain only one peak at $\lambda_d$. The focal length monotonically decreases as wavelength increases, which corresponds to an Abbe number of $v_{refractive}$=-3.5 [11]. If $p_h$>1, the number of efficiency peaks increases as $p_h$ increases. The FWHM decreases as F/# diminishes which

corresponds to the diffraction theory results. As we have calculated before, F/#=0.5 is almost the limit for a refractive index n=1.5. However, the FWHM value when F/#=0.5 is larger than $\lambda/2$ in all cases, which means that a HFL cannot provide super resolution imaging. As the focal length changes with the wavelength, the real F/# is changing with wavelength. This leads to an interesting phenomenon: the FWHM almost keeps constant for an ordinary Fresnel lens for different wavelength as shown in **Fig. 6.4**(c). The optical properties when the HFL is illuminated by a unitary *p*-polarized normal incident plane wave are almost the same as the case with *s* polarization. This indicates that incident polarization for HFL is not so important. For real dispersive optical glass such as NBK7 we need to take into account the variation of the refractive index with the wavelength. In this case, we have also obtained very similar results to **Fig. 6.4**. It means that the simulation with a constant index is not a too coarse approximation.

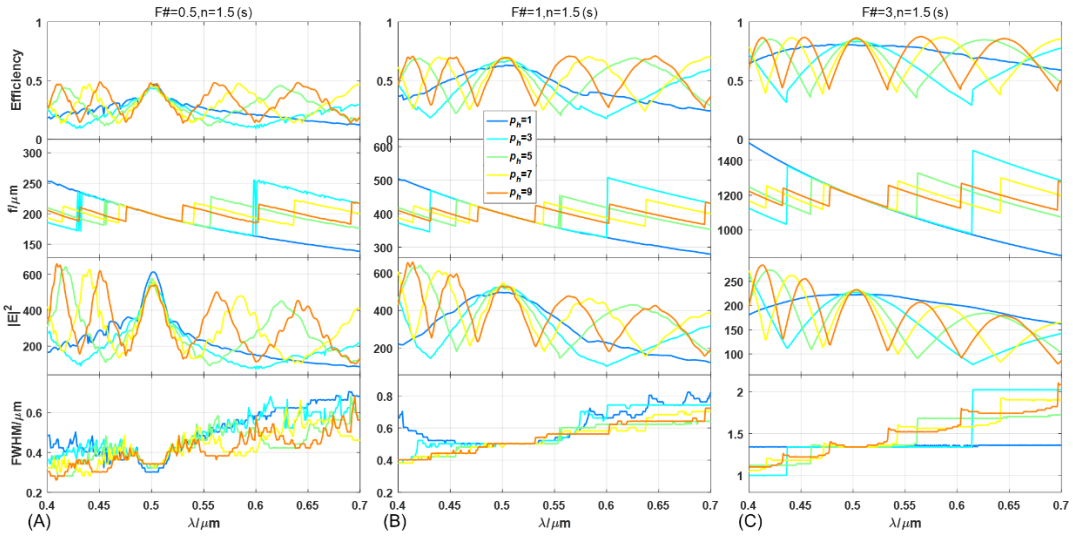



**Fig. 6.4.** Optical properties of several HFLs with different F/# for unitary s-polarized normal incident plane wave. The focus efficiency, focal length, maximum intensity and FWHM vs wavelength are plotted respectively. Each color represents a different harmonic number $p_h$ (a) F#=0.5, (b) F#=1 and (c) F#=3.

The focus efficiency increases as F/# increases. This is a phenomenon that we cannot obtain from scalar equation Eq. (6-5). To explain this effect, as F/# decreases, the incident angle $\theta_{inc}$ in each zone increases, which leads to a higher reflectance on the external surface of the HFL and then leads to a reduction of the diffraction efficiency as shown in **Fig. 6.5**.

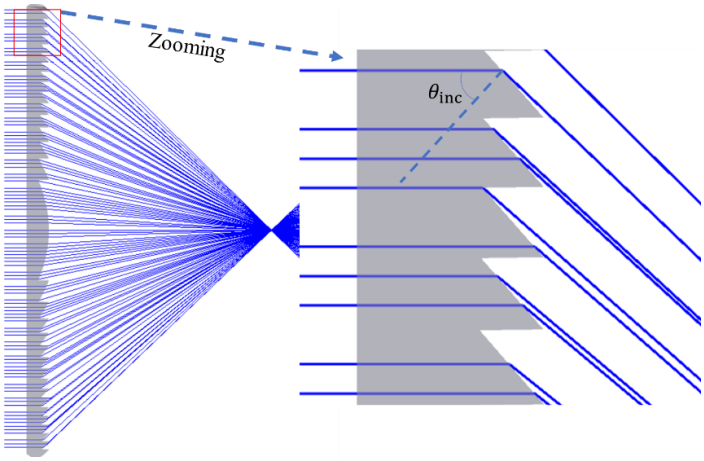



**Fig. 6.5.** Raytracing of a HFL with F#=0.5.

At harmonic wavelengths, the diffraction efficiency reaches maxima with almost the same value (see **Fig. 6.4**). The focal length jumps in the minima positions of efficiency. If $p_h$>1, there is only one focus at harmonic wavelength. But at other wavelengths, we have noticed that two obvious foci occur. For an efficiency minima wavelength, the two foci have the same intensity. Out of the efficiency minima, the foci have different intensities. Between the harmonic and the left efficiency minima wavelength, the further focus has higher intensity as shown in **Fig. 6.6**(a) where the focal length is longer. When we consider the interval between the harmonic and the right efficiency minima wavelength, the closer focus has higher intensity as shown in **Fig. 6.6**(b), where the focal length is shorter. This is the reason why the focal length jumps at efficiency minima wavelengths.

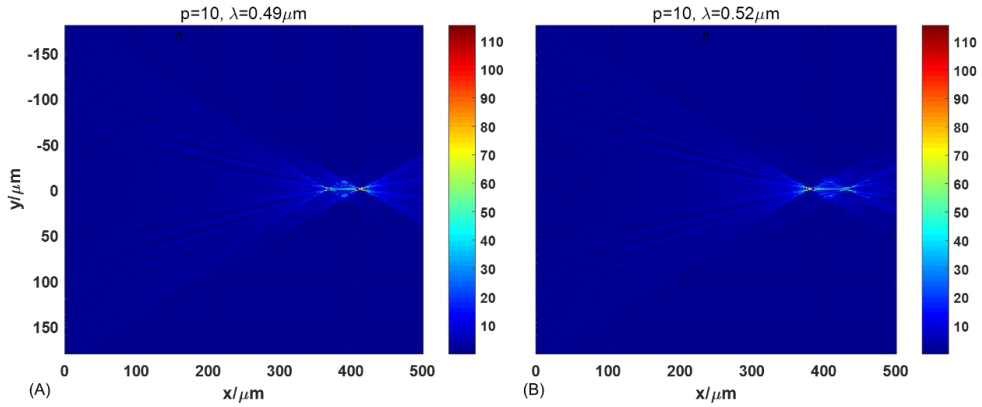



**Fig. 6.6**. Electric field intensity distribution of a harmonic lens with $p_h$=10. (a) $\lambda$=0.49 μm and (b) $\lambda$=0.52 μm.

To compare with the paraxial case, we have calculated the diffraction efficiency using

Eq. (6-5) and the focal length using Eq. (6-4). Then, we have plot the result obtained by FDTD plus ASM as shown in **Fig. 6.7**. As the calculation method for efficiency in the electromagnetic simulation is the energy inside the central lobe over the total illumination energy, the efficiency is always less than 1. However, the maximum efficiency for paraxial approximation can reach 1 at harmonic wavelength. By comparing the wavelengths corresponding to the peak diffraction efficiency, we can observe that the simulated harmonic wavelengths do not precisely respect the theoretical values. The simulated values shift towards larger wavelength as shown in **Fig. 6.7**(a), which is also the case of the jumping positions of the focal length as shown in **Fig. 6.7**(b). This may be caused by HFL boundaries which change the profile of the incident plane wave. These subtle differences can be detected by rigorous simulation but cannot be predicted by paraxial method, which is one advantage of this study.

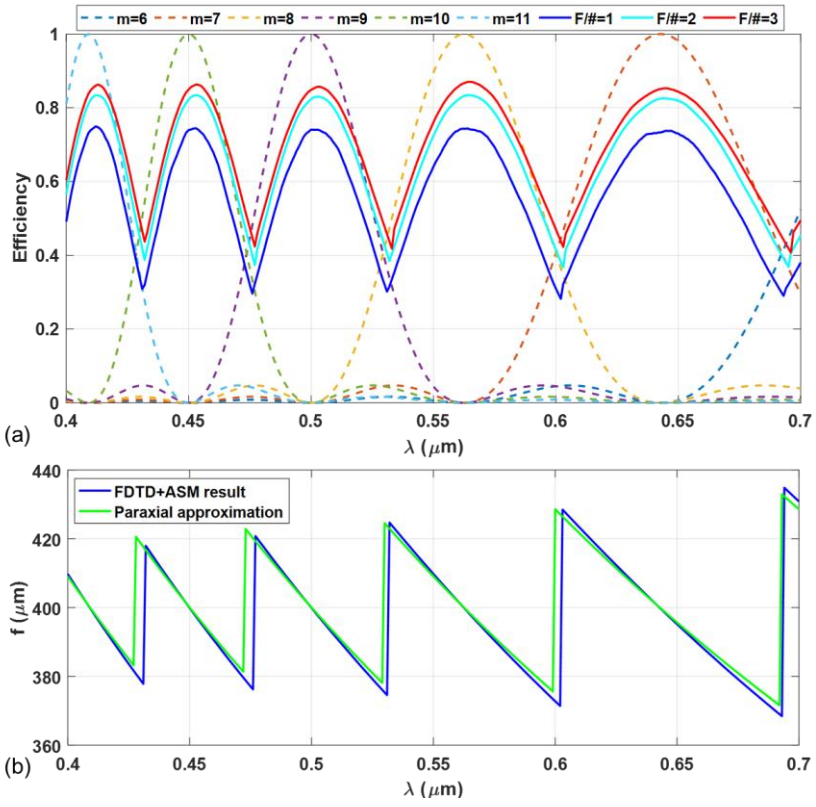



**Fig. 6.7.** Comparison between the result of paraxial approximation and FDTD plus ASM. The harmonic number $p_h$=9 in all cases. (a) Diffraction efficiency in paraxial approximation and FDTD simulation for F/#=1, F/#=2 and F/#=3. (b) Comparison of the focal length as a function of wavelength.

With the FDTD method and the far field projection, we have also calculated the case

at oblique incidence. In this case, for the computation of efficiency, we have used only the maximum intensity because the determination of the intensity envelope is difficult due to its non-regularity. We have calculated the maximum intensity of F/#=3 and F/#=1 for three different oblique angles $\alpha$=1°, $\alpha$=3°, and $\alpha$=5°respectively as shown in **Fig. 6.8**. The positions of the corresponding intensity maxima are shown in Appendix D. For all the simulations, we have used a unitary s-polarized incident plane wave.

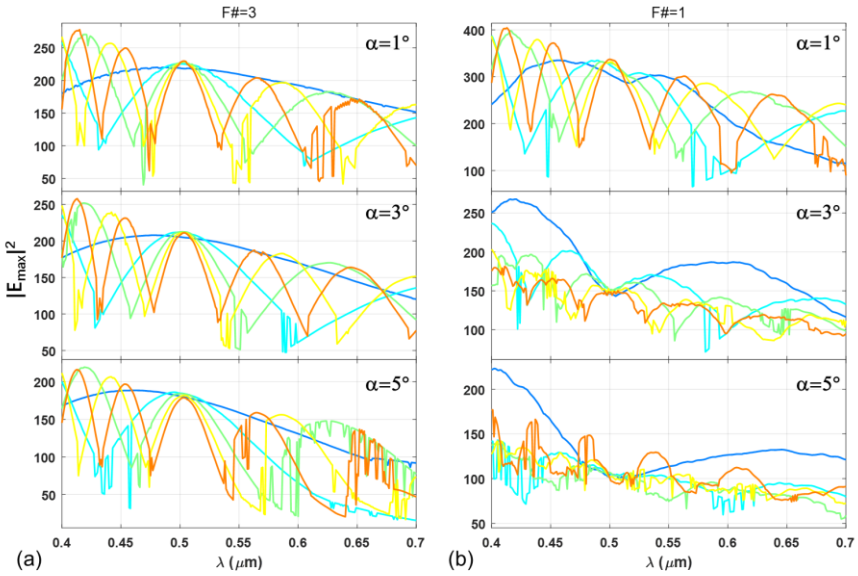



**Fig. 6.8.** The maximum electric field intensity, the y position and x position of the intensity maxima of several HFLs vs wavelength with different incident angle $\alpha$. F/#=3 in all cases. (a) $\alpha$=1° (b) $\alpha$=3° and (c) $\alpha$=5°.

The maximum intensity changes with the wavelength, the harmonic wavelengths are almost identical to the normal incidence case if the incident angle is less than a certain value (typically in our example 1° when F/#=1 and 5° when F/#=3). This angle value increases as F/# increases. Comparing **Fig. 6.8** to **Fig. 6.4** the maximum intensity decreases much less with the same oblique illumination if F/# is larger. We can say that the harmonic wavelengths are more stable when F/# is larger. If F/# is smaller, for example F/#=1, the result is more sensitive to the incident angle. It is difficult to distinguish which wavelength is the harmonic one at $\alpha$=5°. In geometrical optics, we can obtain the same conclusion using raytracing. The aberrations are more serious if F/# is small with the same incident angle as shown in **Fig. 6.9**. We also notice that the aspherical lens which can focus plane wave to a perfect point in **Fig. 6.9**(A) has similar

effect than Fresnel lens in **Fig. 6.9**(B) for the same oblique incidence, except that Fresnel lens can generate some stray light on the extreme boundary zones.

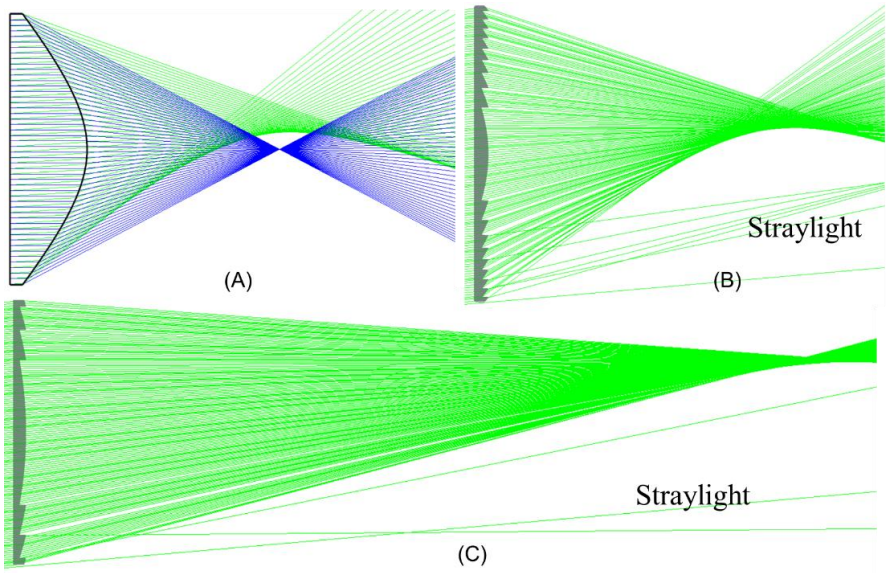



**Fig. 6.9.** Raytracing through lenses with design wavelength. (A) Aspherical lens with F/#=1, raytracing of normal incident rays (blue) and with an incident angle $\alpha$=5° (green). (B) Fresnel lens with F/#=1. Incident angle is $\alpha$=5°. (C) Fresnel lens with F/#=3, Incident angle is $\alpha$=5°.

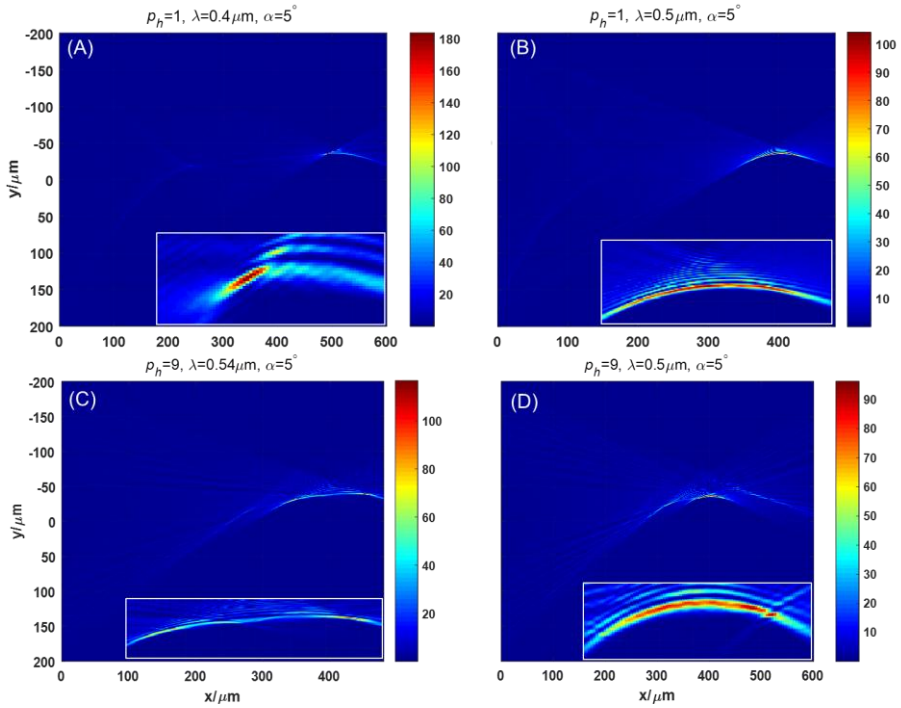



**Fig. 6.10.** Electric field intensity distribution of several cases when the incident angle is $\alpha$=5°. The inserts are the zooming around the focal position. F#=1 in all the cases. (a) $p_h$=1, $\lambda$=0.4μm, (a) $p_h$=1, $\lambda$=0.5 μm, (a) $p_h$=9, $\lambda$=0.54 μm and (a) $p_h$=9, $\lambda$=0.5 μm.

From the raytracing, we can observe clear caustic surfaces as shown in **Fig. 6.9**(a)

and (b) when $\alpha$=5°. The caustics can also be observed in the electric field intensity distributions as shown in **Fig. 6.10**. It seems that the caustics are more serious when $p_h$ is larger.

# 6.4. Conclusion

HFLs present the advantage of working well with several discrete wavelengths for normal incidence. We have investigated the efficiency, focal length, maximum focus intensity and FWHM of the focal spot for several harmonic numbers and F-numbers. We can conclude that:

-The harmonic wavelengths are not sensitive to the F/#;

-The diffraction efficiency and FWHM of the focus increase as F/# increases;

- The incident polarization has a negligible effect on the efficiency, focal length, maximum focus intensity and FWHM of the focal spot;

- The normal dispersion of optical glasses has also a negligible effect;

-The diffraction efficiency decreases as the oblique incident angle increases because of the aberrations;

- With larger F/#, HFLs can bear larger incident angles.

- The harmonic wavelengths are also unchanged with different incident angles if the aberrations are not too high.

Due to these properties, the HFLs are very promising as in-couplers in NED systems to make them more compact and additional studies dedicated to these designs should be conducted. However, to ease the design work, some specific surface models adapted to HFLs for different wavelengths should be developed for optical design software (eg. Zemax or Code V).

Before the simulation of HFLs by FDTD, we have trained on different optical systems, notably a system composed of a dielectric cuboid embedded inside a dielectric cylinder in order to generate an ultranarrow photonic jet. This work has been the subject of a publication in Optics Express [22]. The details are given in Appendix E (Ultra-narrow Photonic Nano-jets)

## Reference


1. N. Zhang, J. Liu, J. Han, X. Li, F. Yang, X. Wang, B. Hu, and Y. Wang, "Improved holographic waveguide display system," Applied Optics **54**, 3645 (2015).
2. S. R. Soomro and H. Urey, "Design, fabrication and characterization of transparent retro-reflective screen," Opt Express **24**, 24232-24241 (2016).
3. H. Mukawa, K. Akutsu, I. Matsumura, S. Nakano, T. Yoshida, M. Kuwahara, and K. Aiki, "A full-color eyewear display using planar waveguides with reflection volume holograms," Journal of the Society for Information Display **17**, 185 (2009).
4. K.-W. Z. Kai-Wei Zhao and a. J.-W. P. and Jui-Wen Pan, "Efficiency balance of a light guide plate with microstructures for a see-through head-mounted display," Chinese Optics Letters **14**, 022201-022205 (2016).
5. X. Hu and H. Hua, "Design and tolerance of a free-form optical system for an optical see-through multi-focal-plane display," Appl Opt **54**, 9990-9999 (2015).
6. J. Han, J. Liu, X. Yao, and Y. Wang, "Portable waveguide display system with a large field of view by integrating freeform elements and volume holograms," Opt Express **23**, 3534-3549 (2015).
7. J. Han, J. Liu, X. Yao, and Y. Wang, "Portable waveguide display system with a large field of view by integrating freeform elements and volume holograms," Optics express **23**, 3534-3549 (2015).
8. J. Guo, Y. Tu, L. Yang, L. Wang, and B. Wang, "Design of a multiplexing grating for color holographic waveguide," Optical Engineering **54**, 125105 (2015).
9. M. A. Golub, A. A. Friesem, and L. Eisen, "Bragg properties of efficient surface relief gratings in the resonance domain," Optics Communications **235**, 261-267 (2004).
10. T. R. Sales and G. M. Morris, "Diffractive–refractive behavior of kinoform lenses," Applied optics **36**, 253-257 (1997).
11. T. Stone and N. George, "Hybrid diffractive-refractive lenses and achromats," Applied Optics **27**, 2960-2971 (1988).
12. B. H. Kleemann, M. Seesselberg, and J. Ruoff, "Design concepts for broadband high--efficiency DOEs," Journal of the European Optical Society: Rapid Publications **3**(2008).
13. J. Jordan, P. Hirsch, L. Lesem, and D. Van Rooy, "Kinoform lenses," Applied Optics **9**, 1883-1887 (1970).
14. S. D. Mellin and G. P. Nordin, "Limits of scalar diffraction theory and an iterative angular spectrum algorithm for finite aperture diffractive optical element design," Optics Express **8**, 705-722 (2001).
15. F. Di, Y. Yingbai, J. Guofan, and W. Minxian, "Rigorous concept for the analysis of diffractive lenses with different axial resolution and high lateral resolution," Optics express **11**, 1987-1994 (2003).
16. J. Jiang and G. P. Nordin, "A rigorous unidirectional method for designing finite aperture diffractive optical elements," Optics express **7**, 237-242 (2000).
17. A. Albarazanchi, P. Gérard, P. Ambs, P. Meyrueis, G.-N. Nguyen, and K. Heggarty, "Smart multifunction diffractive lens experimental validation for future PV cell applications," Optics express **24**, A139-A145 (2016).

18. D. W. Sweeney and G. E. Sommargren, "Harmonic diffractive lenses," Applied optics **34**, 2469-2475 (1995).
19. D. Faklis and G. M. Morris, "Spectral properties of multiorder diffractive lenses," Applied Optics **34**, 2462-2468 (1995).
20. L. Lesem, P. Hirsch, and J. Jordan, "The kinoform: a new wavefront reconstruction device," IBM Journal of Research and Development **13**, 150-155 (1969).
21. D. A. Buralli, G. M. Morris, and J. R. Rogers, "Optical performance of holographic kinoforms," Applied optics **28**, 976-983 (1989).
22. J. Yang, P. Twardowski, P. Gérard, D. Yi, J. Fontaine and S. Lecler, " Ultra-narrow photonic nanojets through a glass cuboid embedded in a dielectric cylinder" Opt. Exp.(accepted for publication)

# 7. Conclusion and perspective

Near to eye displays (NEDs) for augmented reality can be widely used. It has aroused interest both commercially and academically. However, the existing architecture is not completely satisfactory for a large number of applications. The main contradiction is between comfort of wearing and optical performance. Extensive research focused on reducing the weight and dimension of NEDs is needed.

The main goal of this thesis was to design a new see-through NED that may enhance the optical performance and maintain a compact appearance. To achieve these goals, we have proposed a two-layer waveguide design. The two waveguides are stacked up with a small air gap so that light can propagate inside the two waveguides independently. For the end of the two waveguides, we have contemplated two possibilities, a cylindrical edge or a set of right-angle prisms (RAPs) along a cylindrical base. Both the two kinds of edge elements can reflect light and partially couple it into the second waveguide and finally go into observer's eyes by an out-coupler. When using a cylindrical mirror with asymmetrical optical power to obtain an output plane wave corresponding to a virtual image at infinity, we have considered a freeform prism to compensate the asymmetry. As a result, we obtain an increased FOV of H 30°× V 60° and a very large eye box of about H 15 mm × V12 mm allowing large eye movements. The polychromatic MTF values are above 0.3 at 33 lp/mm, which is sufficient for a visual system. However, the fabrication, test and alignment for freeform optics are more complex than rotationally symmetric lenses.

We then have developed another kind of transition edge. We have introduced a set of right-angle prisms (RAPs) along a cylindrical base. This permits us to suppress any optical power in the propagation of object fields inside the two waveguides and the out-coupler. Then the in-coupling subsystem can be made of only rotationally symmetric lenses. The manufacturing and design difficulties are largely reduced. Similar performance with the aforementioned optical system have been obtained.

In order to prove the concept, we have realized the second design with RAPs because the fabrication for the in-coupler is easier and the waves inside the upper waveguide are plane waves. The main improvement of our designs has been protected by an international patent. We have ordered the fabrication of the two waveguides and the transition edge. The dimensions and the surface roughness of each element have been measured. Then the three elements were glued together. The experimental setup has been built to measure the optical response of each element and the whole system. As a result, the surface roughness inside the sawtooth of the transition edge was too high compared to our requirements, so that we have observed a scattered light from the out-coupling mirror. At the time of the writing, we have not yet found a company that can manufacture this component with all our requirements. Finally, by using a 5 mm RAP, we have observed an improved and increased light intensity with expected FOVs. However, the total efficiency is only 10-20%. These low values are mainly due to the high roughness of the used optical surfaces. Hence, the main difficulty we have encountered lies in the fabrication of the transition edge. Future development will be the evaluation of the industrialization and the manufacture of the transition region with our requirements in angle and roughness.

To investigate the possibility of reducing the weight and dimension further, we have also investigated a kind of diffractive optical element (DOE) called harmonic Fresnel lens (HFL). These elements permit to obtain the same phase transfer function for several specific wavelengths, the harmonic wavelengths. By using Finite-Difference Time-Domain method (FDTD) plus angular spectrum (AS) method, we have calculated the HFL characteristics accurately. We have found that HFL can work well independently of F-number values. With oblique illumination, its properties are comparable with conventional lens and the harmonic wavelengths are unchanged. Due to these properties, the HFLs are very promising as in-couplers in NEDs systems to make them more compact and additional studies dedicated to these designs should be conducted. However, to ease the design work, some specific surface models adapted to HFLs for different wavelengths should be developed for the various optical design software (e.g. Zemax or Code V).

The thesis results present promising performance for NEDs design. In the frame of the present thesis, we have obtained some preliminary encouraging experimental results. The next step will be to find an industrial technique to fabricate the transition edge and make a complete prototype. To get a qualitative leap, the possibility of applying new optical technologies such as metamaterials or metasurfaces will be also investigated.

# Appendix A: Specifications of the Ocean Optics HL 2000 HP Source

The white light Ocean Optics HL 2000 HP Tungsten Halogen has the following specifications:

| | HL-2000 | HL-2000-LL | HL-2000-HP |
|---|---|---|---|
| Wavelength range | 360 nm – 2400 nm | | |
| Stability | 0.5 % | | |
| Drift | <0.3% per Hour | | |
| Time to stabilize | Approximately 5 Minutes | | |
| Power Consumption | 1.2A @ 12 VDC | 1.0A @ 12VDC | 1.2A @24VDC |
| Bulb life time | 1500 h | 10,000 h | 1000 h |
| Characteristic | Focused | | |
| Bulb color temperature | 2,960K | 2,800K | 3,000K |
| Room temperature | 5°C – 35°C | | |
| Humidity | 5 - 95% at 40°C | | |
| Output | 7W | 7W | 20W |
| Weight | Approximately 0.5 kg | | |
| Size | 62 x 60 x 150 mm | | |

The spectrum of the source is given by Fig. A.

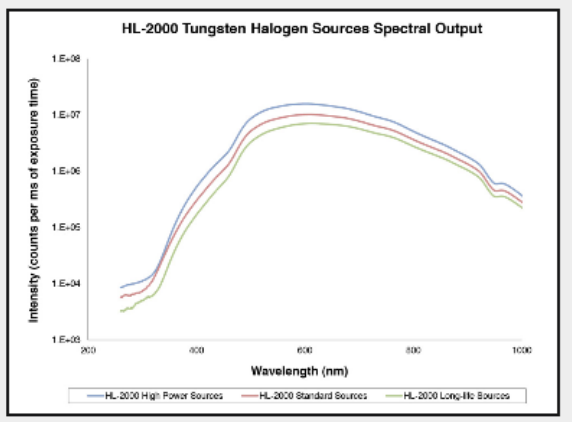



Fig. A: spectrum of the white source

# Appendix B: Specifications of the Basler Scout-sc640-70fm camera

For the intensity measurements, we have used a monochrome camera: Basler-Scout-sc640-70fm – 659 X 494 pixels – pixel Size: 7.4 x7.4 µm –Detector Size: 4.88 X 3.66 mm with the following spectral response.

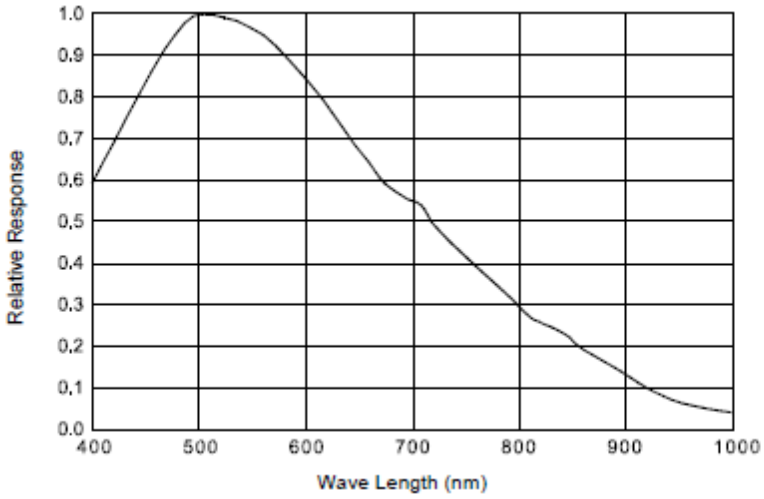



Fig. B. spectral response of the camera.

# Appendix C: Fresnel reflective coefficients and coupling by evanescent waves

As we have total internal reflection (TIR) inside the waveguide, we give the Fresnel equations here. To avoid light can propagate inside the waveguide independently, we give an equation to estimate the minimum value of the air gap.

For s-polarized light, we have:

$$r_{\perp} = \frac{n_i \cos(\theta_i) - n_t \cos(\theta_t)}{n_i \cos(\theta_i) + n_t \cos(\theta_t)}$$

$$t_{\perp} = \frac{2n_i \cos(\theta_i)}{n_i \cos(\theta_i) + n_t \cos(\theta_t)}$$

For p-polarized light, we have:

$$r_{\parallel} = \frac{n_i \cos(\theta_t) - n_t \cos(\theta_i)}{n_i \cos(\theta_t) + n_t \cos(\theta_i)}$$

$$t_{\parallel} = \frac{2n_i \cos(\theta_i)}{n_i \cos(\theta_t) + n_t \cos(\theta_i)}$$

Here we give an example of reflectance and transmittance for a glass-to-air interface:

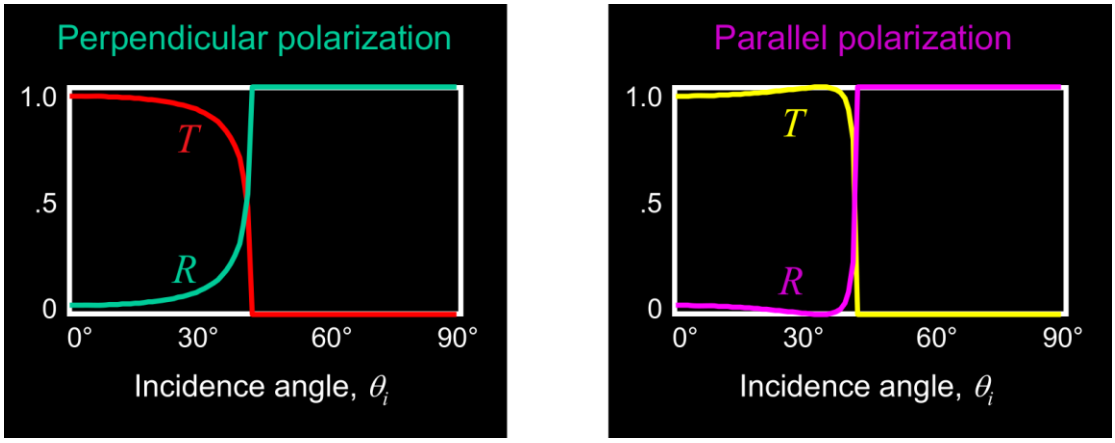



Fig.C1. Reflectance and transmittance for a glass-to-air interface.

If we consider light incident from a waveguide (higher index $n_1$) to air (lower index $n_2$), the wave equation in air when the incident angle greater than the critical angle is:

$$E = E_0 e^{-\kappa z} e^{i(k_x x - \omega t)}$$

which decay exponentially with distance $z$ from the boundary. The decay depth can be expressed as:

$$\kappa^{-1} = \frac{\lambda_1}{2\pi\sqrt{\sin^2\theta - (\frac{n_2}{n_1})^2}}$$

where $\lambda_1$ is the wavelength in the waveguide. We can obtain that the length is in the range of $\lambda$. Hence, the air gap should larger than at least one wavelength.

# Appendix D: Harmonic Lens

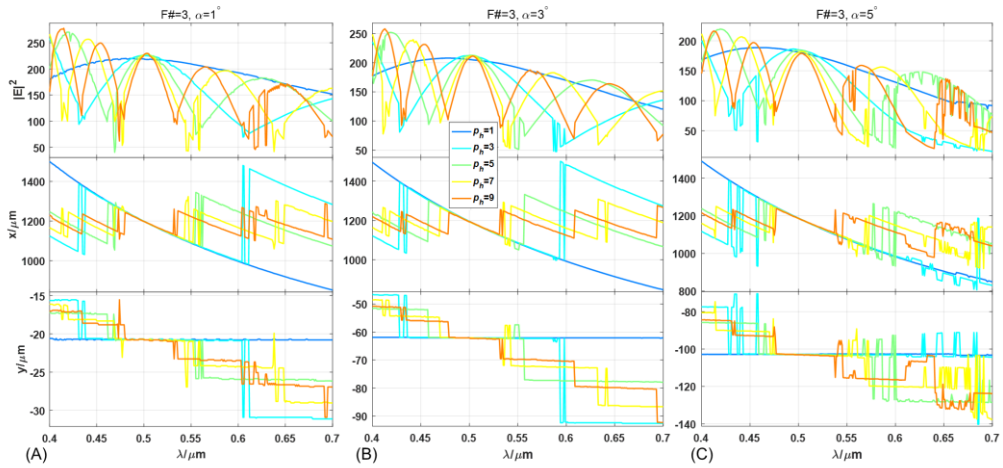



Fig. D1. The maximum electric field intensity, the y position and x position of the intensity maxima of several HFLs vs wavelength with different incident angles $\alpha$. F/#=3 in all cases. (A) $\alpha$=1°, (B) $\alpha$=3° and (C) $\alpha$=5°.

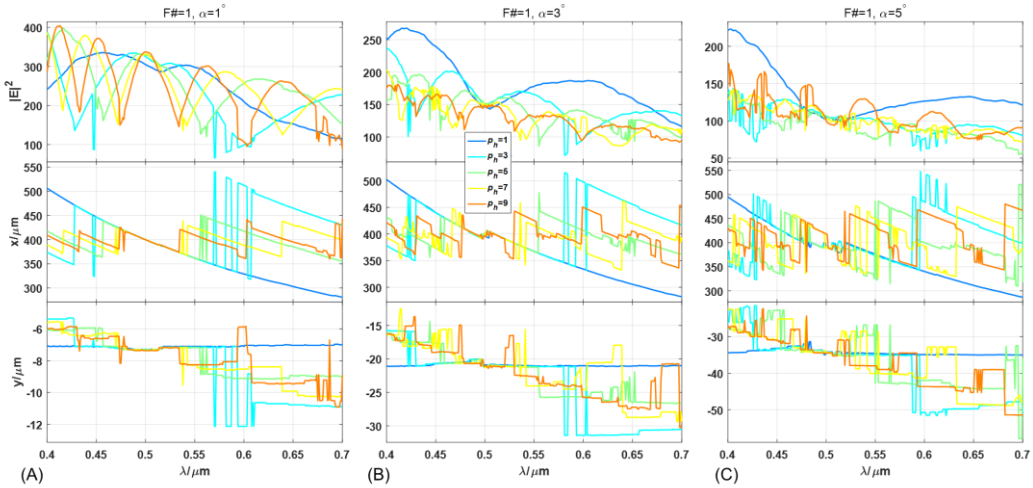



Fig. D2. The maximum electric field intensity, the *y* position and *x* position of the intensity maxima of several HFLs vs wavelength with different incident angles $\alpha$. F/#=1 in all cases. (A) $\alpha$=1°, (B) $\alpha$=3° and (C) $\alpha$=5°.

# Appendix E: Ultra-narrow Photonic Nano-jet

In the process of learning the electromagnetic simulation software (COMSOL Multiphysics), we take an example of photonic nano-jet (PNJ) which is a theme of our team. We proposed a new geometry that can reduce the full width at half maximum(FWHM) of the PNJ. Here we take this opportunity to introduce the result we have obtained.

The embedded transparent dielectric cuboid is placed near the boundary of the transparent dielectric cylinder. A FWHM of around half the Abbe diffraction limit can be obtained. Comparing to the geometry proposed in Ref [24, 25], the cuboid can be considered as locates also near a focal position. However, in our proposed design, the cuboid is immersed not in air, but in the medium of a cylinder, which produce different results. By modifying the size and refractive index of the square section cuboid, we observe: an ultra-narrow PNJ on the external surface of the cuboid, a long photonic jet and the excitation of WGMs. Then considering a rectangular section of the cuboid, we show that the width and the height of the cuboid can be used to adjust the PNJ properties. We also study several other geometries of the insert, which shows that the key parameter is the refractive index of the inserted material. Finally, with a square section cuboid we show that by changing the incident angle a curved photonic jet (hook effect) can be obtained.

### C1. Ultra-narrow Photonic nanojet

An infinite long transparent dielectric square section cuboid embedded in a transparent dielectric cylinder is illuminated by a monochromatic linear polarized plane wave with an incident wave vector $k$ parallel to the $x$ axis. Its polarization direction is parallel to the axis of the cylinder, the $z$ axis. The cylinder is tangential to the center of the outer surface of the cuboid. The electric field distribution is obtained numerically by a finite element method (FEM) with COMSOL Multiphysics. We use free triangular

mesh with a maximum mesh size of $\lambda_0/(15n)$, where $n$ is the refractive index of the corresponding medium. We apply perfect matched layers (PMLs) as the boundary conditions in the four sides. The incident field is introduced as background, and the scattered field is computed. All the results below concern the total field. Fig. 1 shows the simulated PNJ when the radius of the cylinder $r$ is 2.5 μm, the square section side of the cuboid $a$ is 1 μm and the refractive indices of the cylinder $n_1$ and the cuboid $n_2$ are 1.5 and 1.8 respectively. The illumination wavelength $\lambda_0$ is 0.5 μm. The results of the simulation give a FWHM equal to 150 nm. The electric field intensity is shown in Fig. 1(a) assuming a unitary incident intensity. Streamlines of Poynting vectors are shown in Fig. 1(b). The insert is the zoomed image around the focal position confirming that the phenomenon is related to propagating waves. The color scale represents the $Px$ values.

The corresponding electric field intensity profile (the blue curve) on the focal plane is shown in Fig. 1(c) and its distribution along the x-axis is shown in Fig. 1(d). As a reference, we also plot the case of a single full cylinder with an index of n=1.5 (the green curve). $L$ is the working distance from the external surface of the cuboid to the position of maximum intensity. $D$ is the decay length from the point of the maximum light intensity $I_{max}$ to 1/e value of $I_{max}$. Compared with the reference curve, the FWHM is reduced by 36% from 235 nm ($0.47\lambda_0$) to 150 nm ($0.3\lambda_0$), the maximum intensity is increased by about 17% and $L$ is reduced from 455 nm to 32 nm. $D$ is reduced from $2.2\lambda_0$ to $0.7\lambda_0$. We notice that both $L$ and $D$ diminish with the reduction of FWHM, as in Ref [15-20]. This phenomenon can be compared to the result obtained with a conventional lens for which higher numerical aperture corresponds to a shorter focal length for a constant lens diameter.

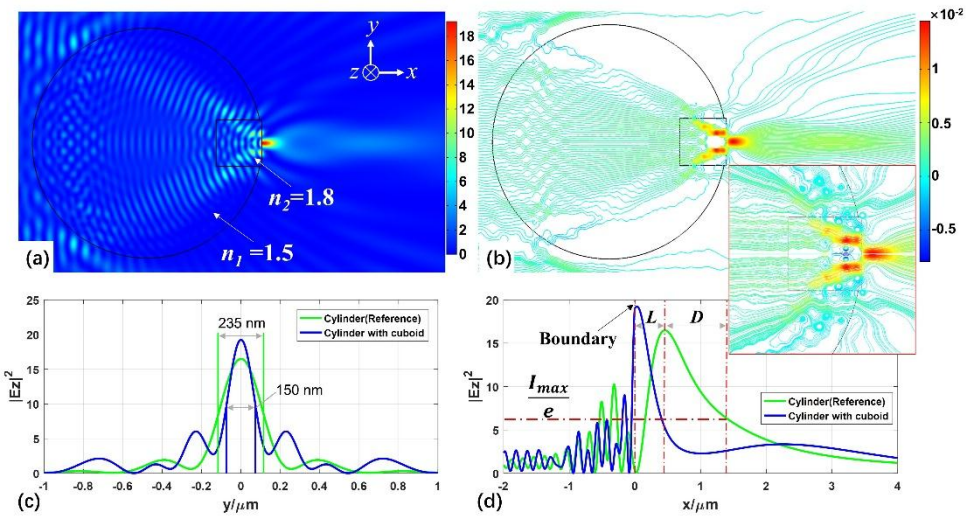



Fig. 1. Dielectric cuboid embedded in a dielectric cylinder. The wavelength of the illumination plane wave is 0.5 μm. The radius of the cylinder is r = 2.5 μm. The cuboid has a square section of side $a$ = 1 μm. The refractive indices of the cylinder and the cuboid are $n_1$ = 1.5 and $n_2$ = 1.8 respectively. (a) The electric field intensity |E|². (b) Streamlines of the Poynting vectors. The insert is the zoomed image around the focal position. The color scale represents the value of the *Px* component. (c) Electric field intensity |E|² on the focus width. In green: reference for a single full cylinder with an index of n = 1.5. In blue: new design shown in (a). (d) The electric field intensity distribution |E|² along $x$. $L$ represents the distance between the position of the maximum intensity and the external surface of the cuboid. $L$ = 455 nm for the reference case and $L$ = 32 nm for the new design. $D$ is the decay length from the point of the maximum light intensity $I_{max}$ to 1/e value of $I_{max}$.

### C2. Parametric study of the cuboid

We first consider a square section cuboid with a side increasing from $a$ = 0.1 μm to 1.5 μm with a step size $\Delta a$ = 0.02 μm. Its refractive index is changed from $n_2$=1.4 to $n_2$=1.9 with steps of $\Delta n$=0.1. The FWHM ratio ($\mathrm{FWHM}/\lambda_0$) is calculated around the maximum intensity position outside of the cuboid. The results in Fig. 2 show the FWHM, the distance of the maximum intensity position $L$ and the maximum electric field intensity |E|² outside of the cuboid as a function of the cylinder size parameter, the cuboid size and refractive. The refractive index of the cylinder is $n_1$=1.5 in all cases. The cylinder size parameter is defined as $q = 2\pi r/\lambda_0$. The curve when $n_2$=$n_1$=1.5 (green curve) serves as a reference.

From Fig. 2, a certain amount of useful information can be deduced. Firstly, the FWHM decreases with an increase in the refractive index of the cuboid. Then, the FWHM is smaller than the cylinder alone mostly when the cuboid has a higher refractive index than the cylinder. For example, the case already discussed in Fig. 1 is

indicated by a star with number 0. The smallest FWHMs occur when $n_2 \geq 1.8$, and the maximum intensity can reach the external surface of the cuboid. As we have only represented the intensity outside of the cuboid, the position of the maximum intensity may be inside. The smallest FWHM ratio is around $0.25\lambda_0$ with $n_2 = 1.9$ and with a rather large cuboid ($a > 2\lambda_0$ for $q$=39.3). One example is indicated with the red star number 1 in Fig. 2(a). Its electric field intensity distribution and Poynting vector streamlines are represented in Fig. 3 as for two other cases. We can notice in Fig. 3(a) that the light in the PNJ is still a propagating wave despite it being an ultra-narrow PNJ. Two local maxima can be observed along the optical axis, showing that constructive interference contributes to this ultra-narrow PNJ.

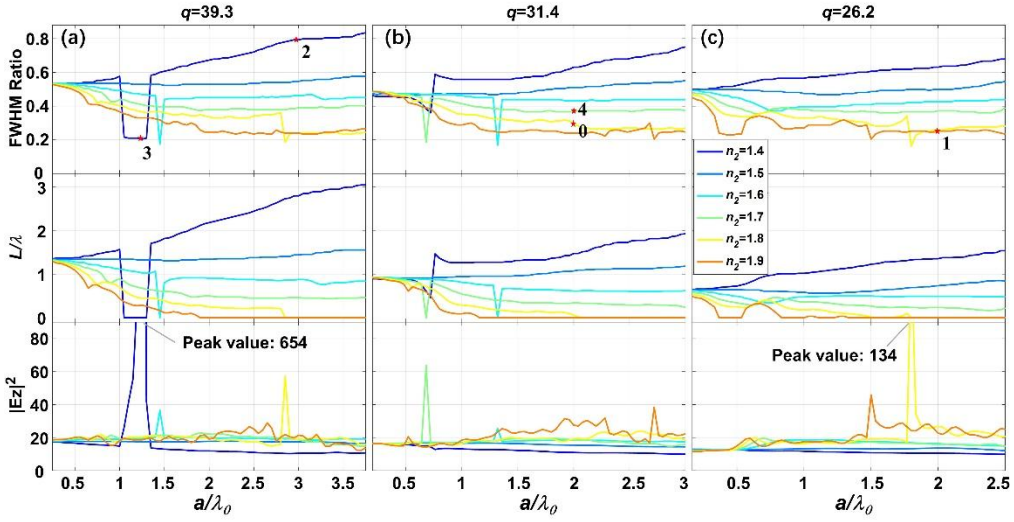



Fig. 2. FWHM ratio (FWHM/$\lambda_0$) of the maximum intensity spot, maximum intensity position $L$ and maximum electric field intensity $|E|^2$ out of the cuboid vs the cylinder size parameter $q$ and the side $a$ of the cuboid in $\lambda$ unit. Different colors represent different refractive indices $n_2$ of the cuboid, which is indicated in the center of the figure. The refractive index of the cylinder is 1.5 in all cases. The red stars are representative points that are studied separately. (a) cylinder size parameter $q$=39.3; (b) $q$=31.4 and (c) $q$=26.2.

On the other hand, when $n_2 < n_1$, Fig. 2 shows that the FWHM ratio of the PNJ is usually larger than for the cylinder alone. $L$ is also larger, and the maximum intensity is lower. In Fig. 3(b) an example is given corresponding to star point 2. This corresponds to a longer photonic jet: the decay length of the jet is 300% longer, the intensity of the jet is 36% lower and the FWHM is 145% larger than the full cylinder alone. Therefore,

we can control the FWHM ratio and $L$ flexibly by choosing required values of $n_2$ and $a$.

Finally, some very specific cases give very high peak intensities and a narrow FWHM. For example, the star point 3 is represented Fig. 3(c). This corresponds to a WGM inside the cylinder with very high intensity values compared to the previously studied PNJs. By definition, a WGMs is evanescent in the radial direction and stationary in the angular one. This is the reason why the WGM is visible on the $|E|^2$ map (Fig. 3(c) left) but not in the Poynting vector map (*Px* component of the Poynting vector in Fig. 3(c) right). We see that despite the WGM resonance, a PNJ is still happened outside. Its intensity is weak comparing to the WGM so that we cannot observe it in the figure of electric field intensity, but we can observe it from the Poynting vectors streamlines. There is an interval for size $a$ when $n_2 = 1.4$, where the FWHM ratio is around 0.2 and the distance $L$ is 0. These cases also correspond to WGMs, which indicates that WGMs can be generated more easily in this kind of configuration. For the three particular cases depicted in Fig. 3, the magnitude of the *Px* component of the Poynting vector is almost the same as when the WGMs create a high electric field intensity.

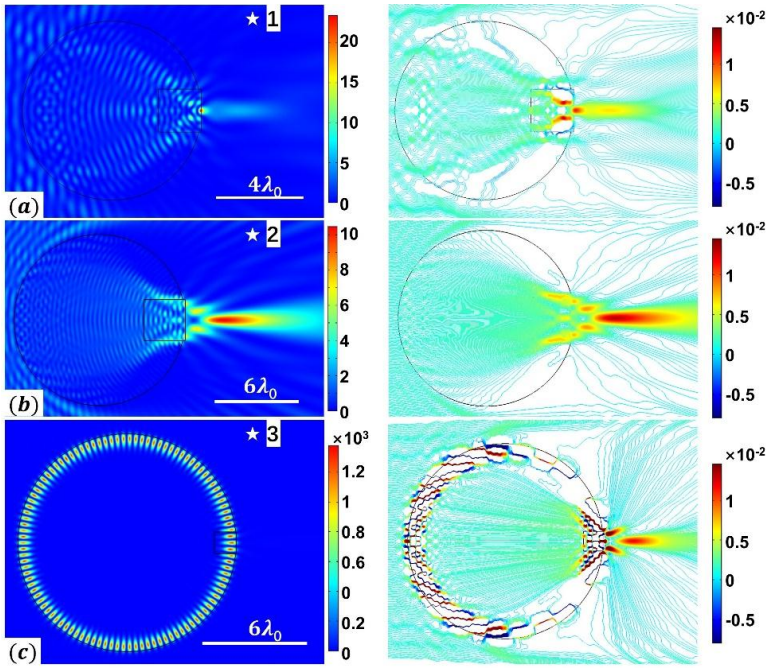



Fig. 3. Electric field intensity $|Ez|^2$ and Poynting vectors streamlines for several specific cases (marked with stars) referenced in Fig. 2. (a) $n_2 = 1.9$, a = $2\lambda_0$and $q$ =26.2. (b) $n_2 = 1.4$, $a = 3\lambda_0$ and $q = 39.3$ (c) $n_2 = 1.4$, $a = 1.25\lambda_0$ and $q = 39.3$.

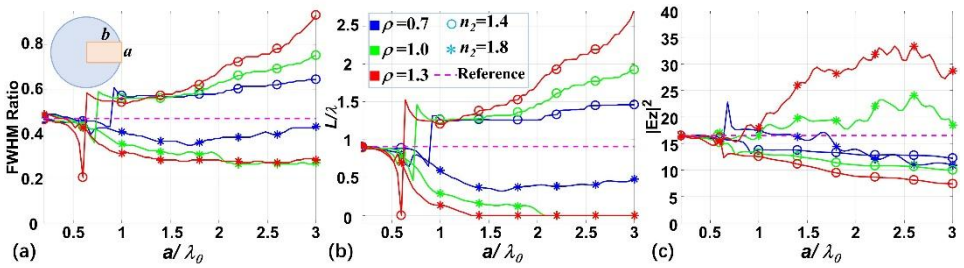



Fig. 4. Characteristics of the focal spot vs the cuboid geometry. (a) FWHM ratio of the maximum intensity spot, (b) maximum intensity position $L$ and (c) maximum electric field intensity $|Ez|^2$ out of the cuboid vs the height $a$. Different colors represent different values of $\rho$ and different marks represent different refractive indices $n_2$ of the cuboid, which is indicated in the center of the figure. The refractive index of the cylinder is 1.5 in all cases. The illumination wavelength is $\lambda_0 = 0.5\ \mu m$. The reference corresponds to a cylinder of refractive index $n_1$=1.5.

We now consider rectangular section cuboids with height $a$ and width $b$ defined along their $y$ and $x$ axes respectively as shown in the insert of Fig. 4(a). We define the parameter $\rho$ equal to $b/a$. In Fig.4, the FWHM Ratio, the distance of the maximum intensity position $L$ and the maximum electric field intensity $|E|^2$ outside of the cuboid ($L$>0) are shown as a function of $a$, $n_2$ and $\rho$. If $n_2$>$n_1$, we notice that for a constant height $a$, as $\rho$ increases, the focal spot becomes smaller and closer to the external surface of the cuboid with a higher intensity, while the focal spot becomes larger and further with a lower intensity if $n_2$< $n_1$. More generally, the width and the height of a rectangular section cuboid are additional parameters that can be used to control the PNJ properties (FWHM ratio, peak intensity and distance $L$ from the cuboid).

### C3. Alternative geometries of the dielectric insert

We have also investigated some alternative geometries that can also reduce the FWHM as shown in Fig.5. The reference case related to star number 4 in Fig. 2 is shown in Fig. 5(a) where the size of the cuboid is $2\lambda_0$. In Fig. 5(b) the insert is similar to a cuboid with a side length of $a = 2\lambda_0$, but with an external surface fitting the host cylinder shape. In Fig. 5(c) the left surface of the original cuboid ($a = 2\lambda_0$) is replaced by a cylinder with a radius equals to $\lambda_0$ and the right surface has a surface fitting the host cylinder shape. Finally, we have investigated the case where the inclusion is a cylinder with $1.4\lambda_0$ radius tangent to the larger cylinder. For all cases, the refractive index of the inserts is 1.7. The FWHM ratio and maximum intensity position $L$ are given in the upper

right corner of each figure.

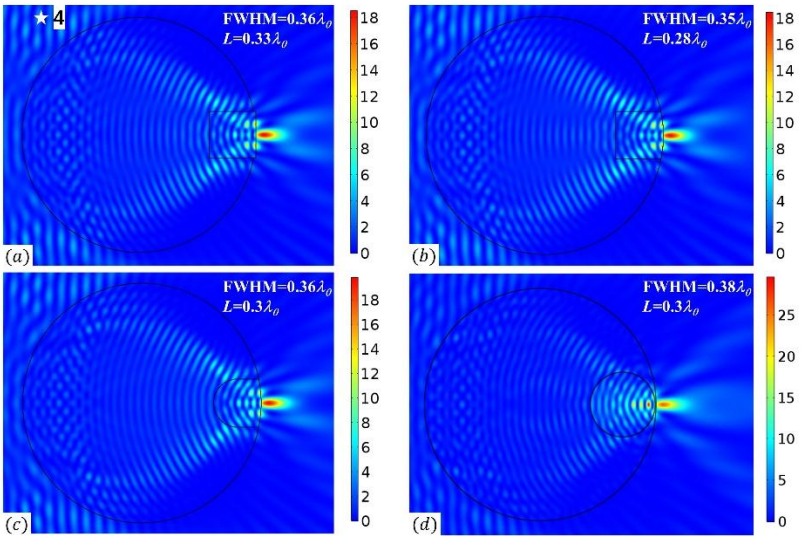



Fig. 5. Alternative geometries that can reduce the FWHM. The refractive index of all the inserts is 1.7. The FWHM ratio and maximum intensity position $L$ are given in the upper right corner of each figure. (a) Case corresponding to star number 4 in Fig. 2. The side length $a = 2\lambda_0$. (b) The insert is similar to a cuboid with a side length of $a = 2\lambda_0$, but with an external surface fitting the host cylinder shape. (c) The left surface of the original cuboid ($a = 2\lambda_0$) is replaced by a cylinder with a radius equals to $\lambda_0$ and the right surface is a surface fitting the host cylinder shape. (d) The inclusion is a cylinder with $1.4\lambda_0$ radius tangent to the larger cylinder.

We notice that the FWHM in all cases are similar and smaller than the Abbe diffraction limit. This indicates that the dominant reason for the reduction of FWHM is the larger refractive index of the insert rather than its shape. For a single cylinder, we know that the FWHM decreases with increasing of its refractive index. However, too high index will focus inside the cylinder, which limits the further reduction of the FWHM. In the proposed system, by putting the higher index insert near the focus, we benefit the reduction of the FWHM from high index and avoid a focus point inside the system. Meanwhile, as the reduction of FWHM is not sensitive to the shape of the insert, the tolerance of fabrication is better. To fabricate such a system, plasma-ion implantation or the use of an ion implantation Van der Graaf, both with mask could be used.

**C4. Incident angle and photonic hook**

When the direction of the illumination light is different from Fig. 1(along the $x$ axis), but with the same geometry and parameters, the electric field intensity distribution also

changes as shown in Fig. 6. Equivalently, for the sake of simplicity, the direction of the illumination light remains unchanged and the cuboid is rotated with an angle $\theta$. Fig. 6 represents the electric field intensity distributions and streamlines of Poynting vectors for three different angles (a) $\theta$=11°, (b) $\theta$=70°, (c) $\theta$=180°. For the case where $\theta$ is less than 11° the maximum electric field intensity and the FWHM of the jet are almost similar compared to the normal incident case as shown in Fig. 1. For $\theta$=11°, there are two or three separate focal points in Fig. 6(a). The most intense jet is due to the cuboid and is distorted. The position of the maximum intensity rotates as the cuboid rotates. The second focus is weaker and due to the cylinder. When $\theta$ is around 70°, the shape of the photonic jet is curved as shown in Fig. 6(b). This effect has been referred to as a photonic hook. This can be understood as an interference phenomenon, which takes place in the caustic and is caused by the break of the symmetry.

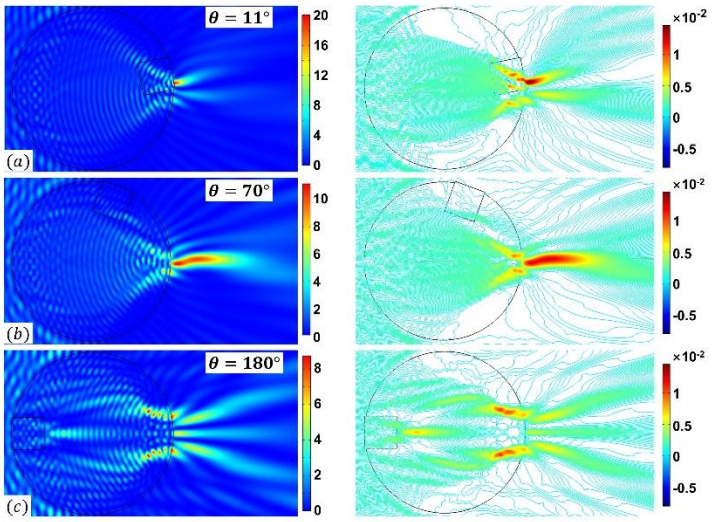



Fig. 6. Effect of the incidence angle. The illumination light flows from left to right. θ is the cuboid rotation angle compared to Fig. 1. As in Fig. 1, the wavelength of the illumination plane wave is 0.5 μm. The radius of the cylinder is r = 2.5 μm. The cuboid has a square section of side $a$ = 1 μm. The refractive indices of the cylinder and the cuboid are $n_1$ = 1.5 and $n_2$ = 1.8 respectively. (Left) Electric field intensity $|E|^2$ and (right) streamline of Poynting vectors for several rotation angles. (a) $\theta$ = 11°, (b) $\theta$ = 70°, (c) $\theta$ = 180°.

Another interpretation corresponds to an analogy with off-axis aberration, mainly coma. When the cuboid moves to the opposite side ($\theta$=180°), a photonic jet created by the cuboid inside the cylinder and three jets outside the cylinder can be observed, as shown in Fig. 6(c). If the high intensity of the focal spot and the small FWHM are the

required parameters, the maximum illumination angle should be less than 11°. If a curved photonic jet, like a photonic hook, is sought for, the shape of the jet can be modified by choosing the angle $\theta$ to be around 70°. Here we give particular angle values for the example of Fig.1. Nevertheless, they also depend on the refractive indices and the section sizes of the cuboid.

### C5. Conclusion

In summary, we have presented a design that can achieve a PNJ with a FWHM of around half that of the Abbe diffraction limit. In this geometry a dielectric cuboid is embedded inside a dielectric cylinder. In this configuration a PNJ is created and focused outside the cuboid. A parameter study of the side and refractive index of a square section cuboid has been presented, demonstrating that smaller PNJs can be obtained with higher cuboid refractive indices. Three particular cases have been selected and discussed: an ultra-narrow PNJ on the external surface of the cuboid, a long photonic jet and the excitation of WGMs. Considering a rectangular section cuboid, if the refractive index of the cuboid is greater than the cylinder index, it can be noticed that for a constant height, as the ratio width over height increases, the focal spot becomes smaller and closer to the external surface of the cuboid and with a higher intensity. We have also investigated several alternative geometries of the insert, which demonstrate that the shape of the insert is not so critical for the reduction of FWHM. To fabricate such a system, ion implantation with mask could be used. Finally, we have shown that by changing the incident angle, a curved photonic jet can be obtained, an effect known as a photonic hook. This may be used to manipulate nano-objects.