# Efficient compression for scalable transmission of digital holograms

Anas El Rhammad

INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2020IPPAT002

Thèse de doctorat

TELECOM Paris

IP PARIS

# Efficient compression for scalable transmission of digital holograms

Thèse de doctorat de l'Institut Polytechnique de Paris
préparée à Télécom Paris

École doctorale n°580 Institut Polytechnique de Paris (EDIPP)
Spécialité de doctorat : Signal, Images, Automatique et robotique

Thèse présentée et soutenue à Rennes, le 11/02/2020, par

**ANAS EL RHAMMAD**

Composition du Jury :

| | |
|---|---|
| Pascal PICART<br>Professeur, Le Mans Université | Président |
| Peter SCHELKENS<br>Professeur, Vrije Universiteit Brussel | Rapporteur |
| Christine GUILLEMOT<br>Directrice de Recherche INRIA, Rennes | Rapportrice |
| Frédéric DUFAUX<br>Directeur de Recherche CNRS, CentraleSupélec, L2S, Paris | Examinateur |
| Cristian PERRA<br>Professeur assistant, Università degli studi di Cagliari | Examinateur |
| Marco CAGNAZZO<br>Professeur, Télécom Paris, LTCI, Paris | Directeur de thèse |
| Patrick GIOIA<br>Ingénieur de recherche, Orange Labs, Rennes | Co-encadrant de thèse |
| Antonin GILLES<br>Ingénieur de recherche, IRT b-com, Rennes | Co-encadrant de thèse |

This work has been achieved within the Institute of Research and Technology b<>com, dedicated to digital technologies. It has been funded by the French government through the National Research Agency (ANR) Investment referenced ANR-A0-AIRT-07.

# Acknowledgements

First and foremost, I would like to express my sincere gratitude to my PhD supervisors, Prof. Marco Cagnazzo, Dr. Patrick Gioia and Dr. Antonin Gilles who assisted and guided me throughout this thesis. I am indebted to them for all the work I have accomplished and for this rewarding PhD degree. Their perspicacious discussions, critical feedbacks and regular involvement have been a major key for the achievement of this work.

I also acknowledge my deepest gratitude to Prof. Béatrice Pesquet-Popescu who supervised me during my first thesis year, helped me with its valuable advices and oriented me to promising research tracks.

I would like to thank warmly my dissertation committee members: Prof. Pascal Picart, Prof. Christine Guillemot, Prof. Peter Schelkens, Prof. Frédéric Dufaux and Dr. Cristian Perra. Their constructive remarks helped me to enhance the quality of this manuscript and to point out new research avenues. It is my utmost honor to have all these experts reviewing my work.

I am sincerely grateful to IRT<>Bcom for providing me the opportunity to pursue my doctorate in such elite research institute. I would like to take this occasion to thank my lab and project managers, Jean-Yves Aubié and Danièle Cloatre for their professional follow-up and support during these three years.

My earnest thanks goes to all my colleagues from bcom and Télécom-Paris for their constant assistance, enlightening discussions and for the pleasant working atmosphere. It has been a pleasure to share these part of my professional experience with you.

I dedicate a special thanks to my friends for their support in the good and the bad, which helped me to keep my goals in perspective. The list is far from being exhaustive but I would like to acknowledge Hassan, Zakaria, El Hassan, Haythem, Sabah, Soukaina, Houssam and Hamza. Thank you guys !

Last but not least, I would like to express my profound gratitude to my parents Malika and Mostafa, my brothers Mohammed and Bakr, my lovely fiancé Magalie and all my family for their unconditional encouragements and ceaseless belief in me throughout these years, which contributed to this beautiful fulfillment.

# Résumé en Français

# 1 Introduction

## 1.1 Contexte de la thèse

L'holographie numérique est une technique qui permet l'enregistrement et la reconstruction de la totalité des ondes lumineuses émises par des scènes du monde réel ou synthétique au moyen des interférences et diffraction de la lumière cohérente [168]. Plusieurs applications ont été identifiées pour l'holographie numérique, y compris la microscopie [120], mesure des particules [144], sécurité [116], stockage des données [106] et systèmes d'affichage 3D [208]. Cette thèse concerne particulièrement ce dernier cas d'usage.

Parmi les différentes technologies immersives émergentes, l'holographie numérique est considérée comme étant la solution parfaite pour une expérience 3D réaliste, authentique et confortable. En effet, contrairement aux techniques (auto-) stéréoscopiques et aux représentations multi-vues, l'holographie fournit tous les indices de perception de la profondeur permettant une visualisation à partir d'un ensemble continu de points de vue et de profondeurs. De plus, l'holographie n'est pas impactée par le conflit entre l'accommodation et convergence qui se produit dans la plupart des écrans 3D conventionnels et engendre une fatigue visuelle [107]. De plus, le port de casques spéciaux n'est plus nécessaire car les hologrammes permettent une visualisation 3D à l'œil nu. Pour toutes ces raisons, l'holographie rend l'expérience de l'utilisateur plus agréable et indiscernable de celle fournie par le système visuel humain.

Malgré ses fonctionnalités de visualisation attrayantes, la réalisation des systèmes d'affichage holographiques nécessite la résolution de plusieurs verrous techniques liés à la génération, représentation, affichage et évaluation de la qualité des hologrammes numériques [21]. Cette thèse aborde les problèmes de traitement du signal relatifs à la compression et à la transmission de données holographiques numériques.

Avant de définir un format de compression des hologrammes numériques, il sera judicieux de se demander pourquoi les motifs holographiques enregistrés devraient être codés au lieu d'encoder simplement la représentation 3D initiale de la scène et ensuite calculer

l'hologramme, à partir des données 3D décodées, au niveau du client. En effet, en compressant le modèle 3D de la scène, qui présente naturellement des redondances spatiales importantes, le système de transmission bénéficierait de l'optimalité des normes de codage lorsqu'elles sont appliquées à des contenus d'image et vidéo ordinaires. D'un autre côté, la propagation du champ d'onde lumineux de ces informations 3D, qui sont initialement bien localisées et corrélées, se traduit par des motifs complètement brouillées et complexes qui sont faiblement compressibles par les techniques de codage conventionnelles. Ainsi, des outils de transformation plus adéquats doivent être conçus afin de permettre une compression dans le domaine de l'hologramme.

Malgré l'efficacité de codage obtenue par une compression dans le domaine source, cela ne peut pas être utilisé par des architectures interopérables pour deux raisons principales. Premièrement, la charge de calcul induite par la génération d'hologrammes côté client ne permettrait pas un affichage en temps réel, notamment pour les grandes résolutions. Deuxièmement, les informations 3D représentant la scène ne sont pas toujours fournies, en particulier dans le cas d'hologrammes numériques acquis à partir des dispositifs optiques. Compte tenu de ces deux limites, il s'avère qu'une compression appropriée est requise pour les hologrammes numériques avant leur transfert. Cette thèse s'inscrit dans ce contexte et vise à concevoir un codec efficace spécifiquement dédié aux signaux holographiques numériques.

## 1.2 Motivations et objectifs

Le développement des formats de compression et de transmission efficaces pour la visualisation holographique numérique doit répondre à plusieurs exigences. Premièrement, les techniques de compression des hologrammes devraient alléger à la fois les exigences de stockage des serveurs internes et les contraintes de bande passante des réseaux de communication. Deuxièmement, le schéma de compression devrait être en mesure de coder/décoder gracieusement les parties pertinentes de l'hologramme, permettant ainsi une fonctionnalité scalable et d'autres manipulations. Enfin, la complexité de calcul du codeur et des décodeurs holographiques doit être réduite pour permettre la conception de systèmes de transmission à faible latence pour des hologrammes de haute qualité. La satisfaction de ces exigences se heurte à deux obstacles:

- **La quantité de données massive:** Les hologrammes numériques contiennent toute l'information 3D de la scène enregistrée au prix d'un très grand volume de données. À titre d'exemple, si on considère une source de laser monochromatique ($\lambda = 640\,\mathrm{nm}$) et un champ de visualisation de $\varphi_{max} = 106°$. Cela implique que la distance entre les centres de deux pixels adjacents doit être égale à $\Delta = 0.4\,\mu\mathrm{m}$. Par conséquence,

générer un hologramme de taille $10cm \times 10cm$ nécessitera une large résolution de $250K \times 250K$ [115, 210], ce qui fait plus de $60$ Gigapixels. En considérant une profondeur de bit de $2$ octets par pixel afin d'encoder les valeurs complexes de l'hologramme ($1$ octet pour chaque partie réel et imaginaire), la capacité de la mémoire demandée s'élèvera à $1$ Terabits. Un hologramme en couleur aux mêmes dimensions nécessiterait trois fois plus d'informations. De plus, cette quantité est multipliée par le nombre d'images par seconde dans le cas des hologrammes dynamiques, ce qui conduirait à un volume de données gigantesque.

- **Les propriétés de signal irrégulières:** Les hologrammes numériques présentent des propriétés particulières qui sont particulièrement différentes des images naturelles. Premièrement, les motifs holographiques ne sont pas localisés dans le domaine spatial avec de très faibles similitudes et leurs variations sont complètement non corrélées avec les changements qui se produisent dans la scène. Deuxièmement, une partie importante de l'énergie de l'hologramme est contenue dans les composantes haute fréquences qui sont difficiles à compresser. Enfin, les hologrammes numériques sont des signaux à valeurs complexes et nécessitent donc des outils de transformation et de quantification spécifiques.

# 2 État de l'art des méthodes de compression d'hologrammes

Dans cette section, nous présentons un état de l'art des différentes méthodes de compression d'hologrammes numériques. Ces approches peuvent être classifiées en deux catégories majeures. La première catégorie comporte les approches s'appuyant sur les techniques de codage classiques utilisées pour la compression des images et vidéos ordinaires, et essayant de les adapter aux caractéristiques irrégulières des signaux holographiques. La deuxième catégorie comporte les approches exploitant les caractéristiques de visualisation 3D afin de mieux interpréter la signification 3D des motifs holographiques.

## 2.1 Méthodes basées sur une adaptation des techniques conventionnelles de codage

Étant donné qu'un hologramme est représenté sous forme de motifs 2D pixélisés, la solution la plus simple et intuitive est d'appliquer directement les solutions de codage conventionnelles. Ces méthodes peuvent être classées en trois catégories: approches basées sur le codage sans perte, basées sur la quantification et basées sur la transformation.

### 2.1.1 Les approches basées sur le codage sans perte

Les premiers travaux sur la compression d'hologrammes numériques ont étudié l'utilisation des algorithmes de codage des données sans perte tels que Lempel-Ziv (LZ77) [213], Lempel-Ziv-Welch (LZW) [198], Huffman [112] et Burrows-Wheeler (BW) [31].

Malheureusement, les gains de codage obtenus par compression pure sans perte ne sont pas convaincants et ne permettent pas le codage à bas débit. Cela nécessite des méthodes de compression avec perte.

### 2.1.2 Les approches basées sur la quantification

Dans les approches basées sur la quantification, chaque valeur d'entrée est transformée en une valeur de sortie quantifiée indépendamment des autres valeurs du signal. Les méthodes de quantification sont divisées en deux classes: la quantification scalaire utilisée pour les données unidimensionnelles et la quantification vectorielle appliquée aux données d'entrée multidimensionnelles.

**Les méthodes basées sur la quantification scalaire** Dans [145], les auteurs ont appliqué plusieurs techniques de quantification scalaire aux signaux holographiques. Ils ont constaté que la quantification uniforme atteignait de meilleures performances par rapport au sous-échantillonnage.Les performances de quantification uniforme ont également été évaluées pour les reconstructions numériques et optiques des hologrammes numériques à décalage de phase [140].

Pour surmonter ce problème, plusieurs travaux ont étudié l'utilisation de quantificateurs non uniformes pour la compression des hologrammes. Dans [176], les auteurs ont exploré l'utilisation de techniques de regroupement itératif, suivies d'un codage sans perte. Leur approche proposée est basée sur l'algorithme $k$-moyennes en plaçant plus de centres dans les zones où les données sont denses et moins de centres dans les régions parcimonieuses. Les expériences ont été réalisées pour différents nombres de clusters en utilisant l'algorithme BW pour le codage sans perte, et ont montré une meilleure performance de compression par rapport à une quantification uniforme.

Malgré les résultats satisfaisants produits par les quantificateurs non uniformes, leur processus itératif est coûteux et dépend fortement des données sur lesquelles ils ont été entrainés. Pour faire face à ces limitations, les auteurs de [177, 179] ont proposé deux approches de quantification pour permettre une compression générique des hologrammes numériques. Ces méthodes bénéficient de la faible complexité de la quantification uniforme, ainsi que des performances supérieures de la méthode non uniforme. Il a été constaté que les performances

de cette approche sont comparables à celles basées sur le regroupement des $k$-moyennes et le réseau neuronal compétitif de Kohonen [62], avec l'avantage d'être non itératifs et insensibles aux changements modérés de la distribution de données holographiques.

**Les méthodes basées sur la quantification vectorielle** Étant donné que les hologrammes sont des signaux à valeurs complexes, ils peuvent être considérés comme des vecteurs de dimension 2 et conviennent alors à la quantification vectorielle. En compression d'image, la quantification vectorielle exploite la corrélation entre les blocs de pixels pour avoir plus de précision dans la définition des centres des clusters [84]. L'utilisation des approches de quantification vectorielle non uniformes a été explorée dans [178] en utilisant le regroupement des $k$-moyennes.

Dans [205], les auteurs ont proposé une quantification vectorielle en utilisant l'algorithme Linde–Buzo–Gray (LBG) pour la compression des hologrammes. En évaluant le comportement débit-distorsion des reconstructions numériques, les résultats ont montré que la compression basée sur la quantification vectorielle LBG surpasse la quantification adaptative scalaire de Lloyd Max. Les auteurs ont également démontré dans un autre travail [206] que la représentation R-I est plus pertinente pour la quantification vectorielle, tandis que la représentation A-P est mieux adaptée à la quantification scalaire. Une étude comparative détaillée des approches de quantification scalaire et vectorielle peut être trouvée dans [42, 123, 206].

Le principal inconvénient des approches basées sur la quantification est qu'elles opèrent directement sur l'hologramme original sans aucune transformation préalable. Pour de meilleures performances de compression holographique, tous les modules du processus du codage d'image doivent être appliqués.

### 2.1.3 Les approches basées sur la transformation

Ce type d'approches transforme d'abord le signal holographique d'entrée dans un autre domaine afin d'appliquer une quantification et un codage plus efficace par la suite. Étant donné que les hologrammes numériques sont codés en tant que des signaux complexes à deux dimensions, une solution intuitive consiste à appliquer les transformations déployées par les normes de codage d'image et de vidéo classiques (Joint Picture Expert Group (JPEG) et Motion Picture Expert Group (MPEG)), soit sur le format R-I ou A-P. Une deuxième solution consiste à remplacer les transformées par défaut utilisées dans ces deux normes par d'autres, afin de mieux adapter les caractéristiques irrégulières des données holographiques. Dans cette section, nous listons les transformations conventionnelles et adaptées appartenant aux formats JPEG et MPEG, qui sont utilisées dans la compression des hologrammes.

**Les transformées basées sur JPEG** Bien que JPEG soit l'un des formats de représentation référencés des images photographiques, elle n'offre des performances de compression élevées que pour les images présentant des variations lisses de texture et de couleur. En ce qui concerne les hologrammes numériques, il est clair que cette caractéristique n'est pas remplie. Par conséquent, il s'avère que la transformation par défaut déployée par le codec JPEG doit être adaptée pour répondre aux propriétés irrégulières des signaux holographiques.

Les auteurs de [211] ont proposé une version améliorée de JPEG pour la compression des hologrammes. Étant donné que la plupart de l'énergie des hologrammes est contenue dans le terme du premier ordre ayant des fréquences élevées, un masque adaptatif est appliqué à chaque bloc avant la quantification afin de rejeter tous les coefficients de Transformée en cosinus discrète (TCD) correspondant aux termes d'ordre zéro, *ie* avec des valeurs d'amplitude faibles. Les expériences montrent que la version améliorée du codec JPEG surpasse celle d'origine en termes de taux de compression et de reconstruction de phase avec et sans bruit. Or, en raison de l'utilisation de la matrice de quantification JPEG par défaut [195], la méthode introduite ne peut pas atteindre des performances optimales.

Afin de minimiser la distorsion causée par le TCD par défaut utilisé dans JPEG pour un budget de débit binaire donné, une table de quantification optimale doit être conçue. À cette fin, une table de quantification TCD adaptative efficace a été explorée dans [37]. La technique proposée est basée sur l'optimisation débit-distorsion en utilisant l'algorithme multiplicateur de Lagrange. La version optimisée a été intégrée dans le codec JPEG de base avec une efficacité de compression supérieure à la précédente et des résultats de codage compétitifs par rapport à MPEG. De plus, l'approche proposée réduit considérablement la charge de calcul des processus de codage et de décodage.

Avec l'émergence des ondelettes comme outil efficace pour la représentation temps-fréquence des signaux [58, 59], la décomposition spectrale de JPEG a été remplacée par une transformées en ondelettes, donnant naissance à une nouvelle norme appelée JPEG 2000 [46].

Grâce à la structure de décomposition multi-résolution des ondelettes et au processus du codage progressif de JPEG 2000, le flux binaire JPEG 2000 est très flexible et évolutif, avec d'excellentes performances débit-distorsion. Par conséquent, JPEG 2000 a été déployé pour la compression d'hologrammes d'interférométrie à décalage de phase [55].

Dans [23], les auteurs ont suggéré l'utilisation de JPEG 2000 avec une décomposition en paquets complets au lieu de la décomposition dyadique de Mallat. Ce choix a été motivé par le fait qu'une partie importante de l'énergie de l'hologramme est contenue dans les bandes haute-fréquence, qui doivent être décomposées davantage pour de meilleures performances de compression. En effet, il permet une décomposition plus générale où les sous-bandes

passe-bas peuvent être divisées dans des directions horizontales et/ou verticales, suivie d'une décomposition sous-niveau supplémentaire de sous-bandes passe-haut pour chaque niveau de résolution.

Bien que l'approche introduite ait permis une amélioration considérable de la compression, il est peu probable qu'elle fonctionne à l'échelle macroscopique, car les composantes haute fréquences et le bruit lié au tavelure seraient plus importants dans les signaux holographiques.

La deuxième catégorie d'approches de codage d'hologramme basées sur la transformation vise à adapter la TCD et la transformée en ondelettes déployées par JPEG/JPEG 2000 ou à utiliser d'autres transformées en ondelettes qui peuvent améliorer l'ajustement des propriétés irrégulières des hologrammes numériques.

Premièrement, l'utilisation des transformées en ondelettes populaires pour le codage numérique des hologrammes a été étudiée dans [43, 44]. À la recherche de la transformée en ondelettes la plus adaptée, les auteurs ont comparé les performances de compression obtenues par 7 schémas d'ondelettes populaires suivis d'un seuillage ou d'une quantification des coefficients d'ondelettes.

Pour améliorer l'efficacité de compression des approches basées sur les transformées en ondelettes, certains travaux ont considéré l'adaptation de schémas d'ondelettes conventionnels pour mieux analyser les franges d'interférence des hologrammes caractérisées par leurs fortes orientations.

En effet, la transformée en ondelettes directionnelles a été déployé pour la compression des hologrammes microscopiques [23, 24]. Cette transformée adapte localement les directions de filtrage aux traits d'image en fonction de leurs orientations locales [38]. En raison des orientations inhérentes présentes dans les images holographiques microscopiques, la transformée en ondelettes directionnelles améliore le compactage énergétique des franges d'interférence conduisant à de meilleures performances de compression par rapport au ondelettes standards.

Une autre catégorie de travaux a exploré l'utilisation des schémas de lifting d'ondelettes [184] introduits comme alternative aux techniques basées sur des bancs de filtres couramment utilisées pour la conception d'ondelettes. En conséquence, une nouvelle décomposition commune d'hologramme multi-échelles basée sur le concept de lifting à vecteur séparable a été développée dans [201].

**Les transformées basées sur MPEG** Depuis sa création en 1988, MPEG a lancé plusieurs normes qui sont largement déployées pour le codage audio et vidéo [197]. Deux normes MPEG sont largement utilisées pour le codage d'images et de vidéos: MPEG-4 Part10 égale-

ment connu sous le nom de H.264 ou Codage Vidéo Avancé (AVC) [199], et MPEG-H Part2 également connu sous le nom de H.265 ou Codage Vidéo Haute Efficacité (HEVC) [182].

Étant donné que AVC a été initialement développé pour le codage vidéo, la première tentative a été de l'appliquer sur des vidéos holographiques numériques [52, 204]. AVC peut également être utilisé pour coder des images holographiques en utilisant uniquement le mode intra-codage. La comparaison des méthodes de compression des hologrammes présentée dans [157] révèle qu'il a une meilleure efficacité que JPEG, mais reste surperformé par JPEG 2000 dans la plupart des cas.

Afin d'exploiter le mode d'inter-codage AVC, les auteurs de [172] ont proposé une nouvelle méthode basée sur le balayage pour générer une séquence vidéo à partir des coefficients de transformation du motif holographique segmenté. En effet, l'hologramme numérique est d'abord divisé en différents segments de tailles égales de $8 \times 8$ à $512 \times 512$ pixels, qui sont transformés dans le domaine fréquentiel à l'aide d'une 2D-TCD . Ensuite, les segments transformés sont balayés par ordre temporel pour former une vidéo.

Le deuxième standard largement déployé pour le codage vidéo est HEVC, qui est le successeur d'AVC. HEVC a été appliqué pour la première fois aux vidéos holographiques à décalage de phase en utilisant le mode d'inter-codage [204], qui atteint un gain de débit binaire plus élevé que AVC. L'efficacité de codage du mode intra a été étudiée pour les hologrammes numériques à la fois avant et après l'affichage, atteignant la meilleure performance débit-distorsion par rapport à tous les codecs susmentionnés [4, 157], y compris AVC en mode intra, JPEG, JPEG 2000 et ses extensions présentés dans [23]. Malgré ses améliorations de codage notables, l'architecture de HEVC a besoin de transformations spéciales pour mieux correspondre aux caractéristiques spécifiques des modèles holographiques.

Dans [158], les auteurs ont introduit une extension du mode intra de HEVC en utilisant des transformées adaptatives. Inspirés par [12, 174], ils ont utilisé des transformations dépendantes de la direction au lieu de la TCD conventionnel. De cette façon, le codec peut tirer parti des directionalités qui peuvent apparaître dans les hologrammes afin d'améliorer les performances de compression.

Malgré les améliorations considérables de la compression obtenues par les approches qui adaptent les transformations des normes de codage d'image et vidéo conventionnelles, certaines d'entre elles n'ont été testées que pour les hologrammes acquis à partir de scènes 2D synthétiques ou à l'échelle microscopique. De plus, la lourde charge de calcul induite par certains d'entre eux n'est pas pratique avec les hologrammes à haute résolution et dynamiques. Plus important encore, bien que les techniques de compression basées sur les transformations recherchent la conception de transformations parcimonieuses optimales et adaptées à la sémantique irrégulière des signaux holographiques, elles ne tirent pas parti de

leurs fonctionnalités de visualisation 3D. En d'autres termes, être capable d'interpréter la signification 3D des motifs holographiques pourrait permettre de meilleures performances de compression et d'autres fonctionnalités intéressantes telles que la scalabilité en termes de point de vue et de qualité ainsi que l'édition d'hologramme.

## 2.2 Méthodes exploitant les caractéristiques de visualisation 3D

### 2.2.1 Codage orienté contenu

La compréhension et l'interprétation de la signification 3D des franges holographiques sont nécessaires pour extraire et encoder les informations pertinentes de la scène enregistrée. Comme objectif ultime, l'outil de transformation devrait fournir une relation directe entre les changements qui se produisent dans la scène et les fréquences locales de l'hologramme. De cette façon, le codeur bénéficierait des redondances spatiales et temporelles initialement présentes dans le domaine source afin de réduire le coût de codage. Cela nécessite une nouvelle catégorie de méthodes de codage holographique, qui vise à développer de nouveaux codecs holographiques prenant en compte les connaissances sur la scène.

Une famille d'approches qui appartient au codage holographique orienté contenu est la compression par rapport à la profondeur de la scène. Dans [151], une dualité importante entre le paramètre de dilatation des ondelettes et la profondeur de propagation des ondes a été soulignée. Sur la base de cette découverte, une nouvelle base d'ondelettes multi-résolution appelée *Fresnelets* a été introduite dans [131], spécialement adaptée à l'analyse des hologrammes. L'utilisation de ce nouveau type d'ondelettes pour la compression d'hologramme numérique à déphasage a été étudiée dans [56].

Le codage d'hologrammes dans le plan objet (*i.e.* reconstruction) a été proposé comme solution alternative à la compression holographique prenant compte de la profondeur [14, 15, 57]. Ces approches consistent à propager l'hologramme de son plan d'acquisition vers le plan objet, puis à compresser le champ d'onde obtenu. Lorsque les hologrammes ont une grande profondeur de mise au point ou sont enregistrés à partir d'objets plats et si la distance de rétro-propagation est soigneusement choisie, l'hologramme recentré apparaît net, ce qui rend l'utilisation des codecs conventionnels d'image puissante comme dans le cas des images naturelles.

Malgré l'amélioration significative des performances de codage obtenue par l'approche de compression basée sur les Fresnelets et le codage du plan objet, leur efficacité est limitée aux scènes presque plates comme en holographie microscopique car elles ne fonctionnent que pour des distances de reconstruction fixes. Par conséquent, la modification de la profondeur de mise au point pour les scènes profondes entraînera des reconstructions gravement défor-

mées [54]. De plus, la rétro-propagation pour le codage et la propagation pour le décodage et l'affichage peuvent entraîner des coûts de calcul coûteux, ce qui empêche l'utilisation de solutions basées sur de multiples propagations en fonction des informations de profondeur.

Pour gérer la large gamme de profondeur des hologrammes macroscopiques, le concept de transformées canoniques linéaires conventionnelles (TLC), y compris les transformées de Fourier et de Fresnel, a été généralisé pour modéliser la diffraction entre des surfaces arbitraires non planes [26]. Tout d'abord, une étude approfondie des propriétés symplectiques dans le domaine temps-fréquence (TF) a permis aux auteurs de construire un opérateur inversible valide, qui est fonction de la profondeur de la scène dans certaines conditions liées à la forme de la surface, à la taille des pixels et à la longueur d'onde. Ensuite, à partir d'une approximation polygonale par morceaux du profil de profondeur, la transformée unitaire proposée peut être utilisée pour approximer l'hologramme à différentes distances de propagation.

### 2.2.2 Codage scalable

L'exploitation des propriétés de visualisation 3D des données holographiques pourrait également permettre un codage scalable d'hologramme. Ces méthodes sont essentiellement pertinentes pour les hologrammes avec un grand champ de visualisation.

Dans la méthode proposée de [128], l'hologramme est découpé spatialement en petites ouvertures, chacune correspondant à un angle de visualisation donné. Ensuite, les sous-hologrammes reconstruits sont codés en utilisant le mode intra d'AVC. Malgré la réduction significative de la bande passante de transmission permise par le schéma de codage scalable en termes de point de vue, une rétro-propagation du plan objet vers le plan d'hologramme est nécessaire après le décodage, ce qui rend le processus coûteux en termes de temps.

Pour faire face à ces limitations et gérer des positions de visualisation arbitraires, les auteurs de [191, 192] ont étudié la pertinence des ondelettes de Gabor/Morlet pour une représentation dépendante de la vue des hologrammes numériques. Dans un travail ultérieur [193], ils ont conçu un cadre pour la représentation et reconstruction holographique dépendante du point de vue, et ont expliqué les procédures d'analyse des ondelettes et la sélection des coefficients. Bien que les aspects théoriques de cette approche aient été étudiés, aucun résultat de compression significatif ni évaluation visuelle réaliste n'ont été rapportés.

# 3 Contributions

Les contributions de cette thèse sont structurées en deux grandes parties. La première partie de ce travail vise à concevoir des méthodes efficaces de compression des hologrammes

numériques pour réduire leur taille. À cette fin, nous proposons un nouveau schéma de codage qui est capable de coder des hologrammes soit dans leur intégralité ou dans un contexte dépendant de la position de l'observateur.

## 3.1 Compression des hologrammes numériques

### 3.1.1 Compression des hologrammes en utilisant le "Matching Pursuit" avec les ondelettes de Gabor

La réalisation d'un système de compression holographique efficace dépend de la façon dont les motifs holographiques sont décomposés et de la façon dont leurs fréquences spatiales sont analysées pour extraire les redondances au sein de la scène enregistrée. Il s'avère que cette tâche est très difficile, en raison des caractéristiques spécifiques des contenus holographiques. En effet, les hologrammes numériques présentent des caractéristiques non locales puisque les informations 3D de l'objet capturé sont brouillées sur tous les pixels de l'hologramme. En conséquence, les outils de compression conventionnels basés sur la décomposition en blocs de pixels ne sera pas adéquate avec les mécanismes de transformation pour prédire les variations des signaux holographiques.

Pour pallier les limitations causées par la décomposition classique basée sur les pixels, nous proposons de décomposer l'hologramme en un ensemble de faisceaux lumineux, qui est plus adéquat avec son caractère diffractif et peut permettre une meilleure efficacité de compression. Pour y parvenir, les ondelettes de Gabor sont utilisées comme outil d'analyse TF pour représenter les faisceaux lumineux diffractés par des motifs holographiques locaux. Afin de permettre une représentation compacte adaptée à la compression, l'algorithme "Matching Pursuit" (MP) est appliqué sur l'expansion redondante de Gabor.

La première étape de notre approche est de générer un ensemble redondant de coefficients en calculant les produits scalaires entre l'hologramme et une famille d'ondelettes de Gabor complexes 2D. Dans un second temps, le "Matching Pursuit" (MP) est appliqué afin de générer un sous-ensemble parcimonieux mieux adapté à la compression. En effet, le MP permet d'extraire itérativement les atomes les plus corrélés avec le signal à décomposer, permettant de le représenter sous-forme d'une combinaison linéaire de ces atomes. Deux versions du MP ont été appliquées pour décomposer les hologrammes en couleurs: $(i)$ une version scalaire qui permet d'extraire les atomes correspondants à chaque canal (R, V ou B) indépendamment, $(ii)$ une version vectorielle qui extrait les atomes de façon jointe correspondant aux trois canaux simultanément.

Une fois qu'une expansion parcimonieuse est obtenue par le MP scalaire ou vectorielle, la troisième étape consiste à générer le flux binaire encodant le signal holographique. Pour

cela, nous avons établi une étude sur les choix de quantification et encodage des coefficients complexes ainsi que les indices espace-fréquence des ondelettes de Gabor (positions 2D, rotations et dilatations). Ainsi, un schéma de codage complet est défini pour les hologrammes numériques.

Les résultats expérimentaux ont révélé que, en termes de performances RD, notre technique surpasse les approches de l'état de l'art dans la plupart des cas. De plus, une évaluation visuelle de la qualité de l'image reconstruite confirme ces résultats. Nous montrons que la version scalaire du MP est plus efficace dans le cas d'hologrammes ayant une faible corrélation entre leurs canaux source, tandis que la version vectorielle du MP est privilégiée pour les hologrammes à forte corrélation. De plus, les résultats de la simulation montrent que l'approche proposée maintiens des performances débit-distorsion et une qualité visuelle cohérentes indépendamment de la focalisation et du point de vue de l'observateur, contrairement aux codeurs classiques, dont les performances se détériorent rapidement lorsque l'observateur se déplace de la position centrale.

### 3.1.2 Schéma de codage holographique scalable en termes de point de vue

Le flux binaire généré par l'expansion de Gabor utilisant le MP ne permet aucune forme de scalabilité, ce qui signifie que tout le flux doit être décodé afin de visualiser le contenu holographique. Malheureusement, cela est incompatible avec les écrans holographiques interactifs.

Pendant la visualisation de l'hologramme, l'observateur perçoit uniquement une direction spécifique donnée par sa position et son orientation. Ainsi, la scalabilité du point de vue est très appropriée dans le cas des hologrammes à haute résolution avec un large champ de visualisation. En effet, elle permet le codage et la transmission d'un sous-ensemble d'hologrammes basé sur la perspective du spectateur, au lieu des données holographiques entières. De plus, les informations correspondant à un point de vue donné peuvent être codées en plusieurs niveaux de qualité, chacune permettant une réduction de la distorsion de la vue reconstruite. C'est ce qu'on appelle la scalabilité en termes de qualité, qui est un outil puissant pour une augmentation progressive à grain fin de la qualité de la reconstruction pendant la visualisation.

Selon l'équation des réseaux [92], les directions d'émission lumineuse sont liées aux fréquences spatiales. Cependant, en raison du caractère non local des hologrammes, chaque région locale contient des composantes spectrales superposées relatives aux contributions de lumière diffusées par plusieurs points de la scène 3D. Pour identifier rigoureusement les informations directionnelles, ces fréquences brouillées doivent être séparées en plusieurs composantes élémentaires, chacune correspondant à un faisceau lumineux émis depuis un

emplacement spatial précis vers une direction exacte. Ainsi, il s'avère que l'extraction des sous-hologrammes correspondant aux différents points de vue de l'observateur nécessite une localisation précise dans les domaines spatial et fréquentiel, simultanément. En effet, une meilleure localisation dans l'espace conduira à une meilleure précision dans la recherche de la position spatiale appropriée dans le plan de l'hologramme, et une meilleure localisation en fréquence réduira l'incertitude angulaire sur la direction du faisceau lumineux. Cette localisation espace-fréquence est contrôlée par la limite d'incertitude de Heisenberg, qui est minimisée avec les fonctions gaussiennes [77]. Étant donné que les ondelettes de Gabor sont des fonctions gaussiennes modulées par des exponentielles complexes, elles fournissent idéalement le compromis de localisation optimal dans les deux domaines.

Le schéma de codage scalable proposé dans cette deuxième contribution est basé sur cette dualité existante entre les ondelettes de Gabor et les faisceaux lumineux émis localement par l'hologramme. Premièrement, la scalabilité en termes de point de vue est obtenue en sélectionnant toutes les ondelettes qui émettent des faisceaux lumineux inscris dans la fenêtre de visualisation. La relation de correspondance entre le plan de l'hologramme et celui de l'observateur est donnée en fonction du point de l'émission ainsi que les paramètres espace-fréquence de l'ondelette.

Deuxièmement, la scalabilité en qualité est obtenue en ordonnançant les atomes de Gabor sélectionnés en fonction de leur importance pour la reconstruction du sous-hologramme. Trois critères d'ordonnancement sont proposés: $(i)$ la distance croissante entre le centre du spectre de diffraction et le centre d'observation, $(ii)$ l'intensité lumineuse décroissante donnée par l'amplitude de l'atome, $(iii)$ la taille décroissante du cône d'émission déterminée par la dispersion angulaire.

Les résultats expérimentaux ont révélé que notre schéma de codage holographique scalable surpasse clairement l'extension scalable de HEVC en termes de performances débit-distorsion et de qualité visuelle. En effet, les reconstructions numériques des sous-hologrammes générées par notre schéma scalable du point de vue fournissent la même qualité de reconstruction que celles obtenues à partir des hologrammes compressés en totalité, atteignant des gains de compression remarquables. De plus, la scalabilité en qualité proposé permet une progressivité à grain fin avec une augmentation rapide de la qualité de reconstruction, par rapport à la scalabilité en qualité de HEVC.

La deuxième partie de ce travail est consacrée à la conception de solutions efficaces pour la transmission d'hologrammes numériques. À cette fin, nous avons développé deux architectures de streaming holographique progressif basées sur le schéma de codage scalable en termes de qualité et de point de vue.

## 3.2 Transmission des hologrammes numériques

### 3.2.1 Streaming holographique basé sur le codage scalable en point de vue et qualité

L'objectif principal de la compression de données est d'alléger les capacités des supports de stockage et de réduire les besoins en bande passante pour une transmission en temps réel sur les réseaux de communication. À titre d'exemple, le téléchargement d'un hologramme monochromatique de résolution $250K \times 250K$ en utilisant une connexion Wi-Fi de $100$Mbps en bande passante prend environ 1 heure et 23 minutes. Ainsi, il est évident que les canaux de communication filaires/sans fils actuels ne seront pas en mesure de transmettre des hologrammes à haute résolution dans un temps raisonnable.

Lorsqu'il s'agit de transférer des ensembles de données massifs sur des canaux à bande passante limitée ou bruyants, le streaming est souvent considéré comme un outil puissant pour obtenir une transmission efficace. En effet, il permet un accès rapide et continu au contenu pendant la transmission des données à l'utilisateur final (avant qu'il ne soit entièrement reçu). En général, la diffusion d'image/vidéo conventionnelle est obtenue soit par des schémas de codage scalable efficaces, soit par un codage adaptatif du contenu à plusieurs niveaux de qualité afin de répondre aux contraintes de bande passante. Cependant, lorsque l'on considère des données holographiques numériques, la première solution n'est pas compatible avec les caractéristiques irrégulières des motifs holographiques, et la seconde entraînera une quantité intraitable de données à stocker.

Pour surmonter les limitations imposées par les solutions de streaming traditionnelles, nous avons développé un système de streaming holographique dans un environnement serveur-client, en utilisant le schéma de codage scalable conçu dans la deuxième contribution.

L'architecture globale serveur-client du schéma de transmission holographique proposé est composée de phases de calcul hors ligne et en ligne. La phase hors ligne correspond à la génération et à la décomposition parcimonieuse de l'hologramme en faisceaux lumineux. Premièrement, l'hologramme numérique est généré par des calculs numériques à partir du modèle 3D de la scène, ou acquis sur un dispositif à transfert de charges (CCD) connecté à un ordinateur. Ensuite, un ensemble d'atomes de Gabor redondant est produit en décomposant l'hologramme original à travers une famille d'ondelettes de Gabor. Enfin, une expansion parcimonieuse est obtenue en utilisant MP, et est stocké dans la mémoire RAM du serveur.

Dans la partie de calcul en ligne, le client envoie périodiquement une notification via le canal de retour pour informer le serveur de la position actuelle de l'observateur. Sur la base de ces informations, les atomes de Gabor permettant une représentation scalable du sous-hologramme correspondant sont extraits de l'expansion. L'ensemble obtenu est ensuite divisé en différentes couches, où le nombre d'atomes de chacune est estimé par

rapport à la bande passante disponible. Dans l'étape suivante, les atomes de chaque couche sont quantifiés et encodés en utilisant l'encodeur MP. Le flux binaire généré est ensuite progressivement transmis au client. Après décodage de chaque couche reçue par le client, ses atomes sont pondérés par leurs coefficients complexes et sommés dans le plan hologramme. Grâce au niveau de progressivité à grain fin, la qualité du sous-hologramme résultant est augmentée après la contribution de chaque atome.

Les résultats expérimentaux ont révélé que la solution de streaming proposée surpasse les solutions conventionnelles basées sur des schémas de codage vidéo scalable tels que SHVC. En effet, l'efficacité de compression couplée à une granularité fine du schéma de codage scalable permet une première visualisation rapide du contenu holographique, et permet d'obtenir de meilleures performances en termes d'augmentation progressive de la qualité de reconstruction. Ainsi, le système de streaming proposé est une première étape vers un streaming progressif efficace d'hologrammes numériques à haute résolution.

### 3.2.2 Streaming holographique à faible complexité en utilisant le codage scalable en point de vue et qualité

Malgré les performances de codage élevées et la granularité fine de l'approche précédente, elle implique une sélection, un tri et un codage en ligne des atomes de Gabor après chaque demande du client. De plus, le calcul des sous-hologrammes pour chaque position d'observation nécessite le calcul de la combinaison linéaire de synthèse du MP. Ainsi, la complexité de calcul induite par le codage scalable et la génération des sous-hologrammes peut augmenter la latence durant la transmission, et ensuite avoir un impact sur la qualité de l'expérience, en particulier dans le cas de multi-observateurs avec des changements fréquents de point de vue.

Afin de réduire le temps entre le moment où le serveur reçoit une notification et le moment où le sous-hologramme correspondant est affiché, nous avons conçu une nouvelle architecture de streaming où la sélection, le tri et le codage des atomes sont effectués au préalable côté serveur avant de démarrer le transfert. En effet, l'intégralité de l'expansion de Gabor est encodée hors ligne par le serveur, puis décodée en ligne suivant la trajectoire de l'observateur. Tout d'abord, on niveau du serveur, on divise le plan d'observation en différents blocs et les atomes de Gabor sont affectés à chaque bloc en utilisant la correspondance entre les ondelettes de Gabor et les faisceaux lumineux émis. Ensuite, chaque groupe d'atomes est codé dans différents paquets en fonction de leur importance pour la reconstruction. Côté client, les indices des paquets sont déterminés après chaque notification de la position d'observation. Ainsi, en appliquant un accès direct à la mémoire, les paquets des atomes concernés sont transmis progressivement.

Par rapport à l'architecture de streaming développée dans le chapitre précédent, où les atomes sont encodés en ligne pendant le processus de transmission, les résultats expérimentaux révèlent que l'approche proposé présente de meilleures performances en termes de compression progressive avec une réduction de la latence de transmission. Ainsi, le système de transmission conçu est une première étape vers un streaming pratique des hologrammes numériques à haute résolution.

# 4 Conclusion

Les travaux réalisés au cours de cette thèse nous ont permis de concevoir un schéma de codage et transmission scalable des hologrammes numériques. Les contributions de cette thèse sont divisées en deux parties majeures: une première partie sur la compression scalable et une deuxième partie sur le streaming progressive. Lors de la première partie, nous avons tout d'abord décomposé l'hologramme en faisceaux lumineux en utilisant les ondelettes de Gabor. Pour des fins de compression, nous avons compacté l'expansion obtenue en utilisant l'algorithme du «Matching Pursuit». Ensuite, nous avons conçu un schéma de codage spécifique aux coefficients et indices des atomes de Gabor. L'approche proposée a atteint une meilleure performance par rapport aux codeurs classiques. Pour la deuxième contribution de cette partie, nous avons exploité la dualité existante entre les ondelettes de Gabor et les faisceaux lumineux émis par l'hologramme afin de développer un schéma de codage scalable en termes de point de vue et qualité. Le débit binaire a été considérablement réduit, sans dégrader la qualité de reconstruction du sous-hologramme par rapport à celle obtenue en encodant la totalité de l'hologramme. Dans la seconde partie de cette thèse, nous avons conçu deux architectures serveur-client pour une transmission progressive d'hologrammes en fonction de la position d'observateur à l'aide du schéma de codage scalable développé. Dans la première solution, un flux binaire scalable à grain fin est généré en ligne par le serveur, après chaque notification du client concernant la position de l'utilisateur. Ensuite, pour réduire la latence causée par la charge de calcul au moment du codage, nous avons proposé une deuxième solution où l'intégralité de l'expansion de Gabor est encodée hors ligne par le serveur, puis décodée en ligne suivant la trajectoire de l'observateur. Les tests de simulation ont montré que l'architecture proposée permet une visualisation rapide en décodant les premiers atomes reçus en plus d'une augmentation progressive de la qualité tout en réduisant la latence du système. En outre, ces travaux ont été valorisés par plusieurs publications et un brevet, dont la liste est donnée en Annexe A. Cependant, nos contributions ont certaines limites qui feront l'objet de travaux futurs.

Tout d'abord, la technique d'optimisation utilisée par l'encodeur proposée pour réduire la redondance du dictionnaire de Gabor est basée sur l'algorithme du «Matching Pursuit». Théoriquement, cette technique ne fournit pas la représentation la plus optimale en termes de parcimonie. Ainsi, comme améliorations possibles de notre méthode, nous proposons d'étudier l'utilisation d'autres algorithmes d'optimisation tels que la version orthogonale du «Matching Pursuit» [61, 154] ou la poursuite de base [39, 40] qui minimise la norme $L_1$ des coefficients de la décomposition. De cette façon, nous pouvons obtenir une expansion de Gabor plus compacte des hologrammes permettant de meilleurs gains de compression.

Deuxièmement, l'approche proposée a atteint des gains de codage limités dans le cas d'une compression de l'hologramme en entier, en particulier pour les hologrammes avec de grandes tailles de pixels. Cela est dû au fait que la transformée en ondelettes de Gabor ne permet pas une compacité optimale de l'énergie du signal. Alternativement, d'autres outils d'analyse temps-fréquence peuvent être utilisés tels que la transformation de Hilbert – Huang (HHT) [110], afin de réduire d'avantage la taille globale des hologrammes. La partie fondamentale de HHT est la méthode de décomposition en mode empirique (EMD) [73, 111, 162], qui décompose les signaux non stationnaires en une base complète de composants élémentaires appelés *Fonctions du mode intrinsèque*. Ensuite, la représentation temps-fréquence est obtenue au moyen de l'analyse spectrale de Hilbert. La particularité de cette transformation est de décomposer de manière adaptative le signal analysé en fonction de ses propres informations au lieu de transformées prédéfinies. Inspiré de l'EMD et la théorie des ondelettes, un paradigme de transformation empirique en ondelettes (EWT) a été introduit dans [88]. Des tests préliminaires ont montré que la compression des hologrammes numériques basée sur l'EWT est plus performante que celle basée sur les ondelettes de Gabor dans certains cas, quand on considère l'hologramme dans son intégralité. En tant que travail futur, la compression basée sur EWT peut être étendue au contexte scalable.

Enfin, notre encodeur/décodeur proposé a été conçu et évalué uniquement pour les hologrammes numériques fixes. À l'heure actuelle, très peu de travaux abordent la compression de vidéo holographique [27, 53, 143, 171, 173]. À l'avenir, notre schéma de codage pourrait être étendu à la compression des hologrammes numériques animés. En effet, à partir de l'expansion de Gabor de la première trame d'hologramme, le domaine temps-fréquence peut être segmenté en regroupant des faisceaux lumineux de direction de diffraction et de couleur similaires, où chaque groupe correspond à une portion donnée de la scène 3D. Ensuite, en approximant une transformation rigide de chaque portion, une transformation linéaire peut être obtenue puis appliquée sur les index des atomes de la trame de référence afin de prédire la suivante. Après ce processus, un mécanisme de compensation/estimation de mouvement efficace peut être construit.

# Table of contents

# List of figures

# List of tables

# Nomenclature

**Acronyms / Abbreviations**

| | |
|---|---|
| 1D | One-dimensional |
| 2D | Two-dimensional |
| 3D | Three-dimensional |
| A-P | Amplitude-Phase |
| AR | Augmented Reality |
| AS | Angular Spectrum |
| AVC | Advanced Video Coding |
| BD | Bjontegaard Delta |
| BW | Burrows-Wheeler |
| CABAC | Context-Adaptive Binary Arithmetic Coding |
| CCD | Charge-Coupled Device |
| CDF | Cohen–Daubechies–Feauveau |
| CGH | Computer Generated Hologram |
| DA-DWT | Directional Adaptive-DWT |
| DCT | Discrete Cosine Transform |
| DFT | Discrete Fourier Transform |
| DWT | Discrete Wavelet Transform |

EBCOT Embedded Block Coding with Optimal Truncation Points

EMD Empirical Mode Decomposition

EWT Empirical Wavelet Transform

FoV Field of View

GOP Groups Of Pictures

GPU Graphics Processing Unit

HEVC High Efficiency Video Coding

HHT Hilbert–Huang Transform

HMD Head-Mounted-Displays

HP Holograme Plane

HVS Human Visual System

JPEG-PLENO Joint Picture Expert Group-Plenoptic

JPEG Joint Picture Expert Group

LBG Linde–Buzo–Gray

LCT Linear Canonical Transforms

LMQ Lloyd Max Quantizer

MB Macro-Block

MDC Multiple Description Coding

MP+AS Matching Pursuit plus Adaptive Selection

MPEG Motion Picture Expert Group

MP Matching Pursuit

MSE Mean Square Error

OP Observer Plane

PCC Pearson Correlation Coefficient

PSD Power Spectral Density

PSNR Peak Signal-to-Noise Ratio

Q-SHC Quality-SHC

Q-SHVC Quality-SHVC

QP Quantization Parameter

R-I Real-Imaginary

RDOT Rate-Distortion Optimized Transform

RD Rate-Distortion

RGB Red-Green-Blue

SHC Scalable Hologram Coding

SHVC Scalable High efficiency Video Coding

SLM Spatial Light Modulator

SMP Scalar Matching Pursuit

SNR Signal-to-Noise Ratio

SSIM Structural Similarity Index

STFT Short-Time Fourier Transform

TD Time-Distortion

TF Time-Frequency

UHD Ultra-High-Definition

UQ Uniform Quantizer

VMP Vector Matching Pursuit

VQ-SHC Viewpoint-Quality Scalable Hologram Coding

VR Virtual Reality

YCbCr Luminance-Chrominance

### Mathematical Symbols

$argmax$ Arguments of the maxima

$argmin$ Arguments of the minima

$card$ Cardinal operator

$\mathcal{F}$ Fourier transform

$\langle a, b \rangle$ Inner product of a and b

$j$ Imaginary unit

$\lambda$ Wavelength of light

$\nabla^2$ Laplace operator

$\mathcal{P}^{AS}$ Angular Spectrum propagation

# Part I

# Introduction

# Chapter 1

# Introduction

## 1.1 Thesis context

Digital holography is a technique that records and reconstructs the whole light wavefield of real-world or synthetic scenes by means of coherent light interference and diffraction [168]. Several applications have been identified for digital holography, including microscopy [120], particle measurement [144], security [116], data storage [106], and Three-dimensional (3D) display systems [208]. The scope of this thesis concerns particularly the latter use case.

Among various emerging immersive media technologies, digital holography stands out as the perfect one for a realistic, faithful and comfortable 3D experience. Indeed, contrary to (auto-)stereoscopic techniques and multiview representations, holography provides all the depth perception cues allowing a visualization from a continuous set of viewpoints and depths. Moreover, holography does not suffer from the convergence-accommodation conflict which occurs in most of conventional 3D displays and leads to visual fatigue [107]. Furthermore, the wear of special headsets is no more required since holograms allow 3D visualization to the naked eyes. For all these reasons, holography make the user's experience more pleasant and indistinguishable from the one provided by the Human Visual System (HVS).

Notwithstanding its attractive visualization features, the realization of holographic displays gave rise to several technical challenges [21], involving the generation, representation, display and quality assessment of digital holograms. This thesis addresses the signal processing issues related to the compression and transmission of digital holographic data.

Before seeking any compressed format of digital holograms, one would wonder why the recorded holographic patterns should be encoded since the initial 3D representation of the scene could be encoded instead, and followed by a hologram computation from the 3D decoded data on the client side. Indeed, by compressing the 3D scene model which

naturally exhibits high spatial redundancies, the transmission system would take benefit from the optimized coding standards which have been constructed for common image and video content. On the other hand, propagating the light wavefield of these 3D information that is initially well localized and correlated, results in completely scrambled and complex features which are weakly compressible by conventional coding techniques. Thus, more adequate transforms mechanisms must be designed for a pure compression of holographic patterns.

Despite the coding efficiency achieved by a compression in the source domain, it cannot be used by interoperable frameworks because of two main reasons. First, the computational burden induced by the hologram generation on the client side would not enable a real-time display, notably for high resolution holographic content. Second, the 3D information representing the scene is not always provided, especially in the case of digital holograms acquired from optical setups. Given these two limitations, it becomes clear that a proper compression is required for digital holograms before their transfer. This work falls within this context and aims at designing an efficient codec specifically dedicated to digital holographic signals.

## 1.2 Motivations and objectives

Thanks to its promising 3D visualization features, digital holography is expected to be the next technology deployed by 3D display systems providers. With the recent advances in optics, photonics and electronics, the design of holographic capturing and visualization devices has been extensively investigated by the research community. On the other hand, relatively few research effort has been spent on the numerical processing of digital holographic signals, especially in the field of compression and transmission [20, 102, 156, 203].

The development of efficient compression and transmission formats for digital holographic 3D visualization must fulfill several requirements. First, hologram compression techniques should alleviate both the storage exigencies of internal servers and the bandwidth constraints of communications networks. Second, the compression scheme should be able to gracefully encode/decode relevant parts of the hologram allowing then scalable functionality and other manipulations. Finally, the computational complexity of the holographic encoder and decoders must be reduced to enable the design of low-latency transmission systems for high-quality holograms. The fulfillment of these requirements is struggled by two bottlenecks:

- **Massive amount of data:** Digital holograms contain the whole 3D information about the recorded scene at the cost of a huge volume of data. For instance, let us consider a monochromatic laser ($\lambda = 640\,\mathrm{nm}$) and a target Field of View (FoV) of $\varphi_{max} = 106°$.

This results in a pixel pitch (*i.e.* distance between the centers of two adjacent pixels) of $\Delta = 0.4\,\mu\text{m}$. Thus, generating a hologram of dimensions $10\text{cm}\times10\text{cm}$ requires a large resolution of $250\text{K}\times250\text{K}$ [115, 210], that is more than $60$ Gigapixels. By considering a bit depth of $2$ bytes per pixel to encode the complex hologram values ($1$ byte for each real and imaginary parts), the required memory capacity is about $1$ Terabits. A full-color hologram with the same dimensions would require three times more information. Moreover, this quantity is multiplied by the frame rate in case of dynamic holograms, which would lead to an intractable volume of data. As a result, transmission of high-quality holograms using current wired/wireless communications channels would not be achieved in time compatible with interactive display systems.

- **Irregular signal properties:** Digital holograms exhibit particular signal properties which are substantially different from semantically meaningful natural images (*cf.* Section 2.2). First, holographic patterns are non localized in the spatial domain with very low similarities and their variations are completely uncorrelated with the changes occurring in the scene. Second, an important part of the hologram energy is carried out by the high frequency components which are difficult to compress. Finally, digital holograms are complex-valued signals and thus need specific transform and quantization tools depending on the considered representation (real-imaginary or amplitude-phase).

Because of these two hurdles, designing efficient encoding formats for holographic signals is still very challenging for state-of-the-art methods, which are mainly based on image and video codecs. Indeed, for high-resolution holograms, the compression gains that may be achieved by conventional coding standards would not be sufficient for an effective transmission when considering the entire hologram. Thus, scalability is highly desirable since it may alleviate the bandwidth requirements by encoding the relevant holographic information that needs to be displayed on the end-user device (for example, corresponding to a given viewpoint or focusing depth) in different levels of quality, allowing a quick glimpse of the hologram and progressive increase of its quality. However, because of the hologram's irregular signal properties, specific transforms are required to design efficient scalable hologram coding schemes.

Therefore, the objective of this thesis is to propose new scalable compression schemes specifically dedicated to holograms, in order to bridge the gap between the tremendous amount of holographic data and the limited bandwidth of communication networks.

## 1.3 Major contributions

The contributions of this thesis are structured into two major parts. The first part of this work aims at designing efficient digital holograms compression methods to reduce their size. To this end, we propose a novel coding scheme which is able to encode holograms either in their entirety or in a view-dependent context.

- **Full hologram compression using Matching Pursuit with Gabor wavelets:** This method enables efficient compression of color digital holograms by expanding their patterns into a sparse set of light beams. Whereas state-of-the-art image and video coding algorithms use pixel block-based decompositions, our approach consists on expanding the input hologram into an overcomplete set of light beams locally emitted by its patterns. A family of discrete Two-dimensional (2D) Gabor wavelets is used to model the light beams diffracted in different directions. Then, Matching Pursuit algorithm is applied in order to reduce the redundancy of the obtained Gabor dictionary. Finally, a complete coding framework is designed by experimentally studying the statistics and distribution of Gabor atoms' coefficients and indexes. Experimental results revealed that our method outperforms conventional image and video coding standards in terms of rate-distortion performance and visual reconstruction quality, especially for holograms with small pixel-pitches.

- **Viewpoint-quality scalable hologram coding scheme:** Alternatively to full compression, this approach allows a scalable hologram coding in a view-dependent scenario. The proposed coding scheme combines viewpoint and quality scalability by progressively encoding only the information corresponding to a given viewer. The realization of this scheme is conducted through two stages: first, the viewpoint scalability is achieved by adaptively selecting the Gabor atoms emitting light beams inside the observation window, exploiting the accurate space-frequency localization of Gabor wavelets. Second, the quality scalability is obtained by sorting the set of selected atoms according to their importance for the reconstructed view. The conducted experiments showed that our viewpoint scalable coding scheme reaches significant data reduction compared to full compression without degrading the reconstructed view. Moreover, the quality scalable scheme clearly outperforms state-of-the-art scalable codecs in terms of progressive increase of reconstruction quality.

The second part of this work is devoted to design effective solutions for view-dependent transmission of digital holograms. To this end, we developed two holographic progressive streaming architectures based on the viewpoint-quality scalable hologram coding scheme.

- **Hologram streaming using progressive coding:** This method defines a holographic streaming architecture in server-client environments able to reduce bandwidth constraints by using the viewpoint-quality scalable hologram coding scheme. Indeed, after each client notification about the observer's viewpoint and the available network' bandwidth, the server selects, sorts and encodes the Gabor atoms corresponding to the requested sub-hologram into different layers. In this way, the hologram is progressively streamed to the client following the user's trajectory. The evaluation of streaming simulations showed two main advantages of our method over conventional image/video streaming solutions. The first one is that the content may be visualized immediately by decoding the first received atoms. The second strength of our proposed solution is to enable a rapid increase of reconstruction quality due to the coding efficiency of our scalable scheme and its fine-grain level of progressivity.

- **Low-complexity progressive hologram streaming:** This approach seeks to develop a low-complexity architecture for holographic streaming in order the reduce the latency induced by the aforementioned streaming method which implies an online coding. In the proposed solution, a viewpoint-quality scalable coding scheme is designed such as the whole Gabor expansion may be encoded offline by the server. First, the observer plane is divided into spatial blocks. Then, the Gabor atoms are assigned to these blocks by exploiting the duality between Gabor wavelets and light beams. The atoms of each block are then classified into different layers according to their importance for the reconstruction and finally encoded in packets. At the decoder side, the atoms' packets are progressively decoded based on the viewer's position. Then, the corresponding sub-hologram is generated using parallel computations. Experimental results revealed that the streaming latency is notably alleviated by the proposed architecture, without decreasing significantly the coding efficiency.

## 1.4 Thesis structure

The dissertation is composed of seven chapters, and an appendix part. It is organized as follows:

**Chapter 1** introduces the technical context of the thesis, the motivations behind, and presents the main challenges and bottlenecks that are addressed by our contributions.

**Chapter 2** provides the general background essential to understand the problematic of this thesis, and lists the basic mechanisms that may be deployed by our contributions. First, we explain why holography is considered as the ultimate technology to naturally

reproduce the human visual depth cues in 3D displays. Then, a thorough comparison between photographic images and digital holograms in terms of diffraction characteristics and statistical signal properties is provided. Finally, we describe the overall transmission chain used in holographic display systems.

**Chapter 3** provides an exhaustive state-of-the-art of digital hologram compression methods. First, we classify the lossy and lossless holographic coding engines and present the approaches belonging to each category. Then, a brief overview of scalable image and video compression is provided, and the existing scalable hologram coding schemes are listed.

**Chapter 4** proposes the first contribution of this thesis, dedicated to the compression of color digital holograms using Matching Pursuit on Gabor dictionary. We first demonstrate how an efficient representation can be obtained for digital holograms by expanding its patterns into an overcomplete set of light beams using Gabor wavelets. The use of Matching Pursuit to get a sparse Gabor expansion is then laid out. We also provide the overall design choices of our proposed encoding framework. Finally, the experimental results with respect to various hologram parameters are given and analyzed.

**Chapter 5** is devoted to our second contribution, which combines viewpoint and quality scalability to further reduce the hologram size and enable progressive coding in a view-dependent context. We first present the proposed viewpoint scalable coding scheme. Then, we describe the quality scalable coding scheme. Finally, the performance of our scalable hologram scheme is evaluated through the conducted experiments.

**Chapter 6** develops our third contribution, which aims at designing a progressive streaming solution using the joint viewpoint-quality scalable hologram coding scheme. First, the proposed streaming architecture is described in a server-client environment. Then, the hologram streaming is simulated in concrete scenarios, and its performance is compared to conventional video streaming solutions.

**Chapter 7** describes our fourth contribution, which proposes a low-complexity progressive streaming solution in order to reduce the latency of the transmission chain. First, the proposed streaming architecture is described in a server-client environment. Then, the offline scalable holographic encoder as well as the online holographic decoder are designed. Finally, the latency reduction and coding efficiency of the proposed streaming solution are analyzed in the experimental results.

**Chapter 8** provides a conclusion of this dissertation, suggests some avenues to improve and extend our contributions in future researches, and finally discusses the major open issues that remain unsolved.

Finally, Appendix A lists the publications and patent published during the thesis, whereas Appendix B provides links to download the videos simulating holographic streaming.

# Part II

# Background knowledge

# Chapter 2

# General background on digital holographic data

The goal of this chapter is to introduce the reader to the basic notions of holographic imaging which distinguish it from photographic imaging in terms of 3D visualization, feature characteristics and signal processing (representation, compression, transmission). Section 2.1 emphasizes the promising 3D visualization provided by holography compared to conventional (auto-) stereoscopic techniques. Then, Section 2.2 outlines the diffraction characteristics and statistical signal properties of holographic patterns. Finally, the digital holographic transmission chain is described in Section 2.3, from content generation to display on end-user devices.

## 2.1 Reproducing human visual system depth cues in 3D displays

Visual computing technologies aim at realizing true-3D displays with the most realistic visualization experience, approaching the one provided by natural vision. Ideally, the viewer would perceive the scene as if it were physically present in front of her. To achieve this, all the physiological and psychological visual depth cues must be duplicated in a natural manner.

**Stereoscopy and auto-stereoscopy** The first attempt to enable 3D visualization was based on Stereoscopy. The principle of this technique is to feed the human eyes with two different 2D views of the scene, which are merged in the brain to create a depth sensation by means of binocular vision [118]. Despite its simple implementation and deployment for 3D television

devices, stereoscopic systems necessitate special wearable glasses with head-tracking to provide motion parallax. By using some specific optical components such as lenticular arrays and parallax barriers, auto-stereoscopic displays allow a depth perception to the naked eye [66]. Moreover, multi-scopic systems deliver more than two perspective views, allowing the viewer to perceive the 3D scene from different angles as his head moves.

The major limitation of (auto-)stereoscopic displays is their inability to provide the focus cue at the real depth of the scene. Therefore, the accommodation of the observer's eye is often in conflict with the vergence of eye-pair, which entails headaches and eye-strain for the user after prolonged utilization [107]. Furthermore, stereoscopy does not allow a true-3D visualization, since it is solely based on a projection of 2D images onto the human observer's retinas, instead of an optical generation and display of the 3D images. To overcome these limitations, alternative technologies for real 3D displays without visual fatigue have been proposed, such as volumetric displays, integral imaging and holography.

**Volumetric imaging** Volumetric imaging permits the 3D perception by illuminating localized regions of a real scene within a physical volume. In volumetric displays, the 3D illusion is created by a stack of planar images conveyed by 2D displays each at a given depth from the eyes. The elements of the volumetric representation are called *voxels*, equivalent to pixels for 2D images. Each voxel of the 3D scene emits visible light from the region in which it appears. Volumetric displays can be classified into two major categories [28]: swept and static displays.

- **Swept volumetric displays:** the 3D image is produced through the projection of decomposed slices onto a surface undergoing rotational or translatory movements. As the surface sweeps through the volume, a continuous volume of light is then perceived thanks to the human vision persistence. By this way, the viewer would be able to view the scene from different angles and focus on different parts.

- **Static-volumetric displays:** create imagery without any macroscopic moving parts in the image volume. Instead, they rely on the light emitted by activated voxels within the viewing volume. Depending on the state of each voxel, *i.e.*, luminous or transparent, it is thus possible to create a real 3D image.

Although volumetric displays do not create accommodation-vergence mismatch, their use is restricted to limited volumes and they do not provide shading nor occlusion cues.

**Integral imaging** The principle of integral imaging consists on capturing multiple perspectives of the scene using a 2D microlens array [133], resulting in a set of micro-images. When

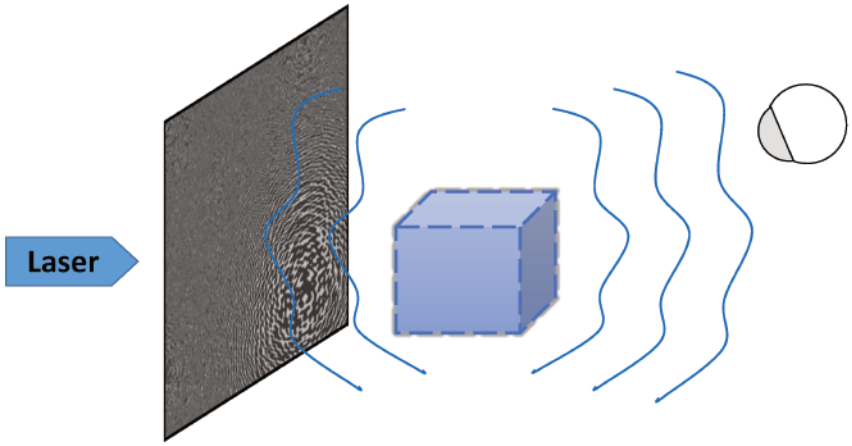



Fig. 2.1 Holographic 3D visualization principle: *by illuminating the captured hologram with the same laser used during acquisition, the scene is perceived as if it were physically present in front of the viewer.*

the latter is illuminated with an incoherent light source, each microlens enables the viewer to perceive the visible area of the scene from its current location. The advantage of using small lenses is to provide accommodation, however, it has the drawbacks of low imaging resolution and small depth range [108].

**Holography** The *concept of holography* was discovered by the Hungarian scientist Dennis Gabor in 1948 [78] during his experiment to improve the quality of electron microscope imagery. In the 1960's, the invention of the laser gave rise to practical optical holography [129]. The optical recording process relies on the interference between an object wave and a reference wave issued from a coherent light source. The obtained interference pattern is then imprinted onto a 2D photographic plate and the acquired image is called *hologram*. With the rapid development of electronic devices such as image sensors and high computational computers, the digital capture, processing and display of dynamic and colorful holograms became possible [94, 169]. Digital holograms can be either optically acquired and then encoded into an electrical signal, or generated by numerical calculations from the 3D model of the scene.

Indeed, when looking at an object, the HVS perceives the light wave that is emitted by this object. In order to provide the perfect 3D depth illusion, one has to artificially reproduce this wavefield in the absence of the object.

Figure 2.1 describes the 3D holographic visualization principle. As shown in this figure, the whole 3D information describing the light wave scattered by the object can be recorded and encoded in a 2D plane. By illuminating the captured hologram with the same laser used

for the acquisition, the object wavefront is fully reproduced by means of light diffraction, and then becomes visible to the naked eyes.

The advantage of holography is its ability to record both the amplitude and phase of the light wave emitted by the scene, while conventional photography only captures the intensity information. Therefore, it accounts for all monocular and binocular depth cues including vergence, occlusion and accommodation [89]. As a result, the reconstruction of holograms allows the exact duplication of the object light wave which can be visualized from all possible perspectives and focus depths with a sensation of physical presence.

To summarize, holography is a promising technology for real 3D displays that enables the acquisition and reproduction of the light wave scattered by a given scene. Due to the natural, authentic and comfortable 3D visualization experience which provides users, holography is expected to dethrone the (auto-) stereoscopic devices in the future [135].

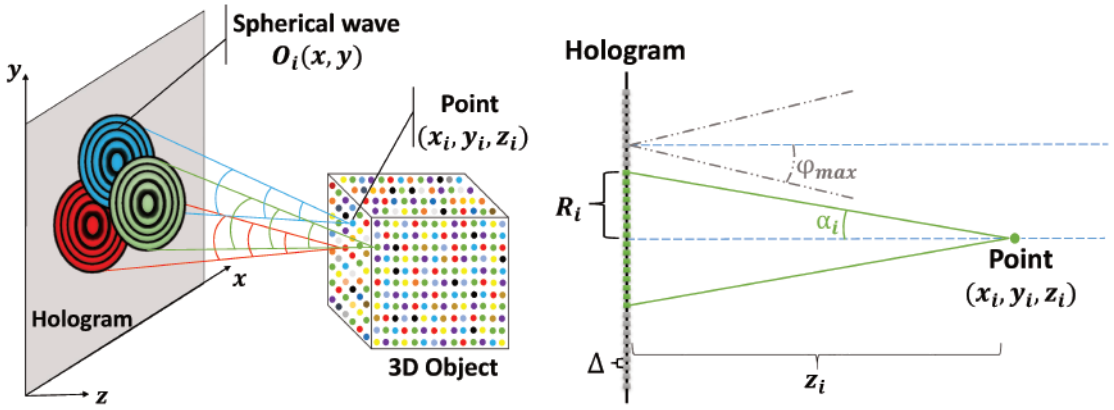
## 2.2 Characteristics of digital holograms

Because of their specific recording process based on interference of coherent light, digital holograms exhibit particular features which are substantially different from photographic images. To understand the semantics of holographic data and enable efficient processing, one needs to figure out the characteristics of their complex signal. In this section, we emphasize the diffraction characteristics and statistical signal properties of holographic patterns, which distinguish them from classical images.

### 2.2.1 Diffraction characteristics of holographic patterns

Contrary to ordinary images where the content can be intuitively recognized by visual interpretation of the spatial features, digital holograms do not present any common geometric features, such as edges, contours or textures. To determine the diffraction characteristics of holographic signals, it is important to first apprehend how such irregular patterns are formed by the propagation of light.

In photography, the light scattered by a given object is captured from a single direction, *i.e.* point of view. In contrast, digital holograms are obtained by capturing the light coming from all possible viewing directions, which are scrambled over the entire hologram plane. To illustrate this characteristic, let us consider the 3D object represented with a point cloud as shown in Figure 2.2a. The coordinate system $(x, y, z)$ is defined so that the hologram lies on the $(x, y, 0)$ plane. Every point $i$ with coordinates $(x_i, y_i, z_i)$ is considered to emit a spherical light wave when illuminated by a source of coherent light. The object wave scattered by




(a) Hologram formation from the scattered object wave.

(b) Zone of light contribution by a given point.

Fig. 2.2 Non-local character of holograms.

the scene into the hologram plane can thus be computed as the sum of all spherical waves $\{O_i(x, y)\}$ emitted by the scene points, such that

$$O(x, y) = \int_i O_i(x, y) = \int_i \frac{a_i}{r_i} e^{[j(kr_i + \varphi_i)]}, \tag{2.1}$$

where $a_i$ is the amplitude of the point $i$, $\varphi_i$ its phase, and $r_i$ is the oblique distance, given by

$$r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2 + z_i^2}. \tag{2.2}$$

As depicted from Equation 2.1, a hologram gathers intrinsically the superimposed light beams diffracted by all the object points. Thus, the 3D information spreads out over the whole hologram and the localized features are scrambled in the hologram plane.

Figure 2.2b presents the light beam emitted by a given point with an angle $\alpha_i$. To avoid aliasing, $\alpha_i$ should not exceed the maximum angle of incidence on the hologram $\varphi_{max}$, which is given by the grating equation [92], such that

$$\varphi_{max} = \arcsin(\lambda f_{max}), \tag{2.3}$$

where $\lambda$ denotes the wavelength of light, and $f_{max}$ is the maximal spatial frequency of the hologram.

According to the Nyquist sampling principle, the maximal spatial frequency that can be represented with a sampling pitch $\Delta$ is given by $f_{max} = (2\Delta)^{-1}$. Thus, according to Figure 2.2b and Equation (2.3), the maximal region that can be covered by the diffracted

beam in the hologram plane is given by its maximum radius

$$R_{i,max} = z_i \tan\left(\varphi_{max}\right) = z_i \tan\left[\arcsin\left(\frac{\lambda}{2\Delta}\right)\right]. \tag{2.4}$$

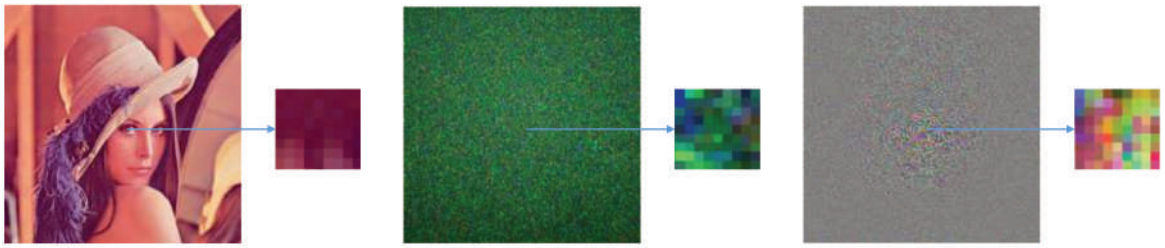
As illustrated by Equation (2.4), the non-local nature of holograms depends both on the distance between the scene and the hologram plane, and the pixel pitch. Consequently, the holographic signal present non-localized features, which are difficult to interpret in the spatial domain.

During the object wave restitution, the hologram diffracts the light towards different directions, which enables the observer to perceive the scene from different viewpoints based on its position and visualization angle. Thus, the holographic patterns could be seen as a superimposition of diffraction gratings associated to different directions of light emission. Since the local directions of diffraction are related to the spatial frequencies (*cf.* grating equation), a local spectral analysis may provide a relevant tool to physically interpret the diffractive character of holographic signals.

### 2.2.2 Statistical signal properties

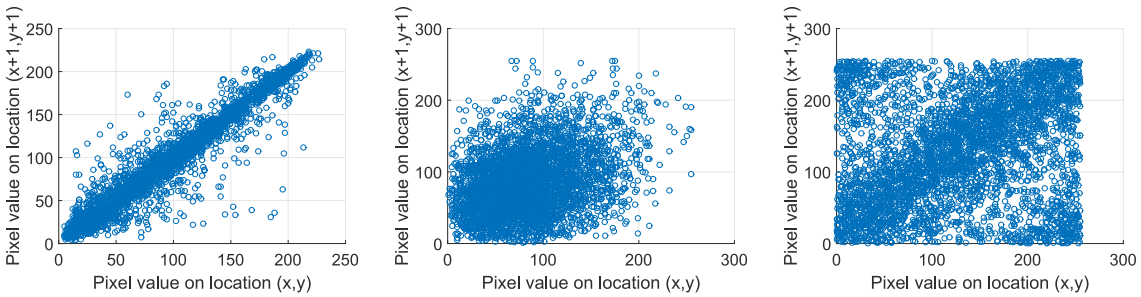
Whereas the understanding of the hologram's diffractive behavior allows a physical interpretation of its patterns, statistical properties provide a mathematical characterization of the signal variation. To emphasize this aspect, we study some statistical measurements of holographic signals in both spatial and spectral domains, and then compare them to those of regular images.

**Spatial statistics:** One of the most important statistical measurements that describes the signal nature is the correlation. It represents the degree of similarity between the samples of two observations, or between the elements of the same data. Based on the visual comparison

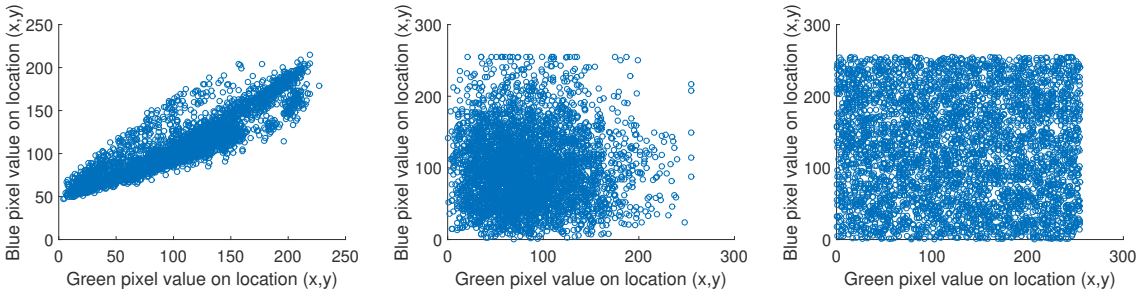


(a) *Lenna* image (b) *Car* hologram (amplitude part) (c) *Car* hologram (phase part)

Fig. 2.3 Comparison between adjacent pixels of hologram and image




(a) *Lenna* image (b) *Car* hologram (amplitude) (c) *Car* hologram (phase)




(d) *Lenna* image (e) *Car* hologram (amplitude) (f) *Car* hologram (phase)

Fig. 2.4 Correlation distributions of diagonally adjacent pixels (a-c) and green-blue channels (d-f). A patch of $64 \times 64$ pixels is considered.

showed in Figure 2.3, the correlations between the adjacent pixels of the *Car* hologram (zoomed part) for both amplitude and phase are very low compared to those present in the *Lenna* image (zoomed part). This is confirmed by the correlation distributions of the diagonally adjacent pixels are given in Figures 2.4a, 2.4b and 2.4c (the phase is renormalized between $0$ and $255$). As noticed in these figures, most of the points are clustered around the main diagonal for the *Lenna* image, whereas the pixel values of the *Car* hologram have a nearly uniform distributions especially for the phase.

The correlation between the color channels of digital holograms are also investigated. The distributions of Figures 2.4d, 2.4e and 2.4f corresponds to the correlation between the green and blue channels. Similarly to the intra-pixel correlation, the inter-channel correlation distribution of the *Lenna* image lies on a diagonal line, whereas the phase hologram exhibits a fairly uniform distribution.

In order to quantify these correlations, we calculate the Pearson Correlation Coefficient (PCC) score, defined by [155]

$$r_{xy} = \frac{\sum_i (x_i - x_m)(y_i - y_m)}{\sqrt{\sum_i (x_i - x_m)^2} \sqrt{\sum_i (x_i - x_m)^2}}, \tag{2.5}$$

| | Vertical | Horizontal | Diagonal | Red-Green | Red-Blue | Green-Blue |
|---|---|---|---|---|---|---|
| **Image** | 0.9844 | 0.9801 | 0.9583 | 0.8837 | 0.6905 | 0.9065 |
| **Amplitude hologram** | 0.5563 | 0.5669 | 0.3199 | 0.0125 | 0.0073 | 0.0133 |
| **Phase hologram** | 0.3566 | 0.4072 | 0.2287 | 0.0328 | 0.0078 | 0.0021 |

Table 2.1 Intra-pixel and inter-channel pixel correlation scores. The values are averaged over a set of $10$ patchs of $64 \times 64$ pixels.

where $x_m$ and $y_m$ are the mean values corresponding to the two observations $x$ and $y$, respectively.

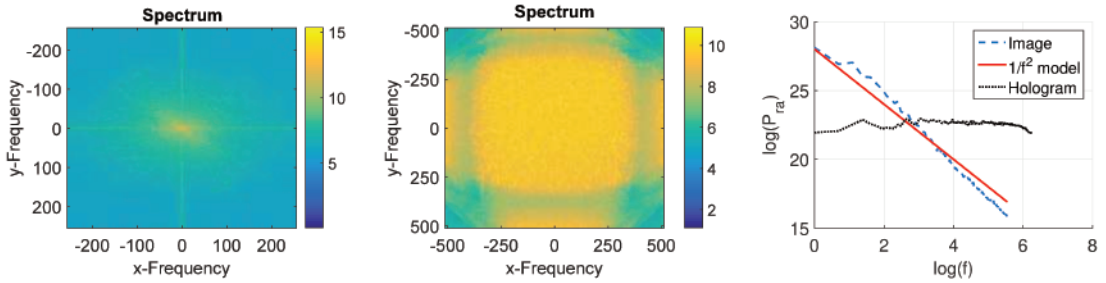
For the intra-pixel correlation, the PCC score is computed in the horizontal, vertical and diagonal directions. For the inter-channel correlation, the PCC score is computed considering the three combinations of color channels. Table 2.1 summarizes the average PCC scores obtained for a set of ten images and holograms. The values show a high correlation between the adjacent pixels of the images with a mean of $0.9743$ over the three directions, compared to $0.481$ for the amplitude holograms and $0.3308$ for the phase ones. Moreover, the mean of the PCC scores over the three color combinations reach a value of $0.8269$ for images, compared to $0.011$ and $0.0142$ for the amplitude and phase holograms, respectively. These scores confirm the very low correlation between the hologram Red-Green-Blue (RGB) channels.

**Spectral statistics:** As mentioned in the previous section, the directions of light emission depend on the local variation of frequency components. Seeking for a statistical model to describe the spectral variation, the Power Spectral Density (PSD) of holographic signals is analyzed, and compared to the one characterizing common images. As shown by the PSD of the *Lenna* image (Figure 2.5a), most of the image energy is concentrated in few low frequency coordinates. This property is very useful for compression, since the high frequencies can be omitted without causing a significant information loss. In contrast, the power distribution of the *Car* hologram (Figure 2.5b) is uniformly spread across the whole spectral domain. This means that the elimination of any frequency components would impact the directions of light emission, and then impair the quality of the reconstructed object wave.

In order to provide a one-dimensional (1D) interpretation of the PSD, the polar PSD is averaged over all orientations to obtain a radially averaged spectrum, defined by

$$P_{ra}(f) = \frac{1}{N_\theta} \sum_{i=1}^{\theta} P(f, \theta_i), \tag{2.6}$$

where $N_\theta$ is the number of orientation bins, and $P(f, \theta_i)$ represents the polar power spectra.




(a) *Lenna* image. (b) *Car* hologram. (c) Radially averaged power spectra.

Fig. 2.5 Comparison between the spectral variations of image and hologram.

As reported in [190], the average power spectrum of natural images tends to exhibit a $1/f^2$-model, where $f$ is the spatial frequency. This statistic is confirmed by the radially averaged spectrum $P_{ra}$ of the *Lenna* image presented in Figure 2.5c. Indeed, the $log(P_{ra})$ follows a linear decay in function of $log(f)$, with a slope of $-2.2$ approximately. On the other hand, the log-average power spectrum of the *Car* hologram shows narrow random fluctuations around its mean value ($log(P_{ra}) = 22.47$).

To summarize, the statistical properties of holographic signals reveal that it is complicated to formulate an appropriate model to describe their character. Indeed, holograms are non-stationary signals with a pseudo-stochastic variation in both spatial and spectral domains. Consequently, it turns out that such irregular statistics would destabilize all the foundations we have been establishing for conventional images.

## 2.3 Digital holographic transmission chain

The realization of 3D holographic display systems involves multiple hardware and software manipulations in order to achieve a good quality of experience while visualizing the 3D content. Figure 2.6 conceptually illustrates the overall block-diagram of the digital holographic transmission chain, which is composed of four blocks: creation of the digital holographic data, encoding format, compression and transmission, and finally the display for the end-user. The detailed description of each step is given in the following. Note that the compression and transmission part are the core of this thesis.

### 2.3.1 Digital hologram creation

As stated before, digital holograms can be either optically acquired or computer-generated, according to the targeted use-case. In the first case, the holographic interference patterns

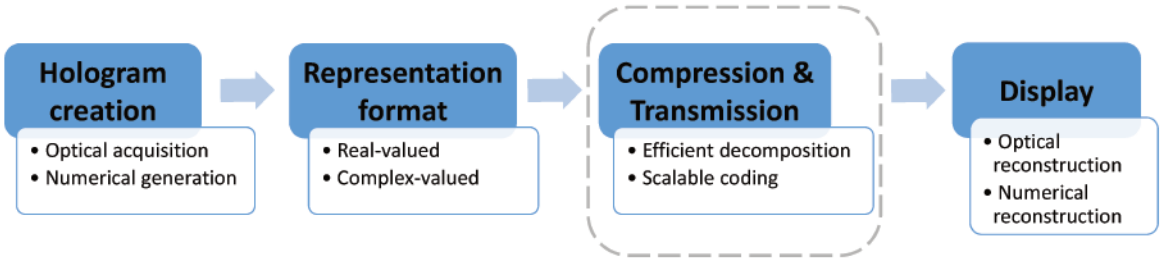



Fig. 2.6 Digital holographic transmission chain. *The dashed block corresponds to the scope of this thesis.*

are obtained optically, and the hologram is then encoded digitally using a Charge-Coupled Device (CCD) image sensor connected to a computer, as shown in Figure 2.7. The laser used by the optical setup is split into two beams: the first one is the reference wave $R$ and the second beam corresponds to the object wave $O$ emitted by the illuminated object to be recorded. Then, the phase information is captured by means of interference between $R$ and $O$. Finally, the resulting wavefield is encoded into the CCD and transferred to the computer as an array of complex numbers. It must be noted that various optical configurations have been proposed in the literature [30, 129, 194, 207]. In the second case, the hologram is generated by numerical calculations, which simulate the light propagation phenomenon. To this end, the object wave is computed using the scalar diffraction theory. The obtained hologram is called a Computer Generated Hologram (CGH). The numerical generation is typically more advantageous than the optical acquisition, since it allows not only to reconstruct but also to print and/or display the generated holographic patterns. Moreover, CGHs are not limited to real existing objects and may be generated also from virtual scenes. Several techniques have been proposed for CGH computation from 3D models [114, 130, 147, 186], and multiview-plus-depth data [86, 141, 150, 175]. Also, fast CGH synthesis methods has been developed in this work [85].

To the best of our knowledge, current digital holographic databases include b<>com [86, 87], Interfere [22] and EmergImg-HoloGrail [14, 15]. These databases contain a set of acquired digital holograms and color CGHs with different generation parameters. Recently, these holograms have been collected by the Joint Picture Expert Group-Plenoptic (JPEG-Pleno) database [167], and are considered as the benchmarking reference for research purposes.

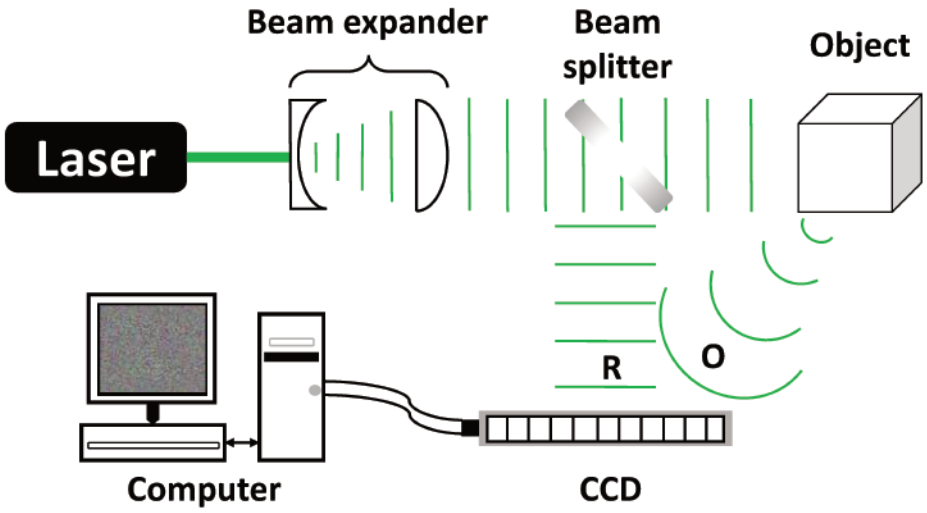



Fig. 2.7 Simplified setup of digital hologram acquisition

### 2.3.2 Representation formats

Once the object light wave is acquired or computed, it is encoded across a 2D surface corresponding to the hologram plane. Two encoding data formats of this wavefield are commonly used: real-valued and complex-valued representations.

- **Real-valued representation:** the object wave is encoded into a hologram with real positive values. The incident reference wave can be modulated using either amplitude or phase information.

  The amplitude hologram $H$ is recorded as the intensity of the interference between a reference wave $R : \mathbb{R}^2 \mapsto \mathbb{C}$ and an object wave $O : \mathbb{R}^2 \mapsto \mathbb{C}$, obtained such that

  $$H = (O + R)(O + R)^* = |O|^2 + |R|^2 + 2\Re\{OR^*\}, \tag{2.7}$$

  where $C^*$, $|C|$ and $\Re\{C\}$ are respectively the conjugate, amplitude and real parts of the complex number $C$. The numerical computation of Equation (2.7) simulates the physical phenomenon of wave interference occurring in the conventional optical acquisition process. Generally, the first two terms corresponding to the object and reference wave intensities, respectively, are not considered since they induce undesirable zero diffraction order artifacts during the reconstruction process [100]. Thus, all the necessary holographic information is contained in the third term $2\Re\{OR^*\}$.

  The modulation of the phase yields phase holograms also called *kinoforms*, which have a better diffraction efficiency than amplitude holograms. However, available optical sensors are not able to record phase information. Alternatively, iterative phase retrieval

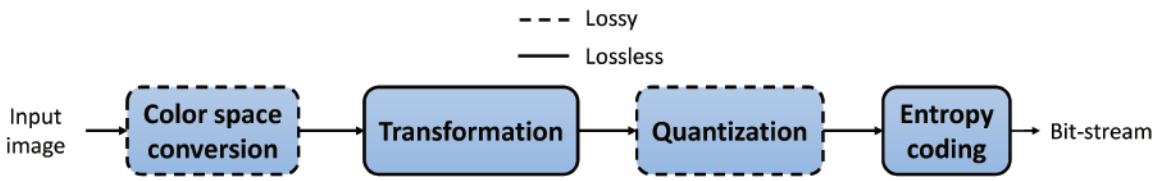



Fig. 2.8 Conventional architecture of lossy and lossless image compression.

algorithm has been proposed to record the phase of the object wave [83]. Although the encoded data is reduced by half, real-valued representation cannot reproduce faithfully the object wavefield since only amplitude or phase information is recorded.

- **Complex-valued representation:** this is the most complete representation since it allows to control jointly the amplitude and phase of the object light wave, yielding an accurate replication of the captured scene. Complex-valued holograms are represented as complex numbers using either Real-Imaginary (R-I) or Amplitude-Phase (A-P) representation. Another possible complex representation is obtained by computing the difference between multiple amplitude hologram recordings with predefined shifted phases [207]. This technique is known as *Phase-shifting holography*.

Once a suitable hologram representation format has been chosen, the real/complex values are quantized to the number of levels given by the plotting device characteristics.

### 2.3.3 Compression and transmission

The main goal of data compression is to reduce the redundancies present in a given signal, and provide the best trade-off between the compression ratio and the distortion of the reconstructed signal. Since digital holograms are 2D signals, image compression engines are applied to digital holograms in order to reduce their size and then enable their storage or transfer to high-end displays. To design an efficient holographic codec, it is important to first discern the fundamentals on which typical image coding systems are based. Figure 2.8 shows the conventional architecture of image compression in both lossy and losseless cases. It is mainly constituted of five steps, which are discussed in the following paragraphs.

**Color space conversion:** For image compression, one of the common preprocessing operations is the colorimetric conversion from the RGB color space to the Luminance-Chrominance (YCbCr) one [75]. Since the HVS is highly sensible to the luminance information (Y), the blue-difference Chroma (Cb) and red-difference Chroma (Cr) components are down-sampled with a factor two, achieving then a data reduction. It must be noted that this step induces some marginal losses of the original image.

Since the light diffraction by holograms is related to its wavelength (*cf.* Equation 2.3), applying such color conversion, would severely degrade the 3D visualization properties of digital holograms.

**Transformation:** The second step in the image compression pipeline is the transformation of the preprocessed image in order to remove the spatial correlation. The objective of this process is to concentrate the image energy using the fewest possible transform coefficients. Many types of transforms have been investigated for image coding, including Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT) [5] and Discrete Wavelet Transform (DWT). The transform coding is reversible and then considered as a lossless process. Contrary to natural images, holograms contain an important part of their energy in the high frequency components (*cf.* Section 2.2.2). Thus, employing conventional coding techniques based on high-frequency filtering would lead to a poor compression performance for digital holograms.

**Quantization:** The third module of the image compression system is a lossy process where the transform coefficients are mapped to integer values. By encoding the coefficients with fewer levels, quantization increases the compression ratio at the cost of a quality drop. Two types of quantization can be distinguished: the scalar quantization such as uniform or adaptive based on Lloyd Max algorithm [134], and the vector quantization like the iterative Linde–Buzo–Gray (LBG) algorithm based on $k$-means [132].

**Entropy coding:** The final step is a lossless process where an embedded bitstream is generated from the quantized transform coefficients. The concept of entropy coding is to assign more bits to the symbols having a low occurrence probability and less bits to those with a high probability. The entropy coding algorithms which are widely used by the image compression standards are: Huffman coding [112] and arithmetic coding [200].

One of the most useful functionality in image compression is the scalability. Indeed, it has long been known as an effective tool allowing the progressive (de-) coding of very large datasets [188]. The principle of scalable image compression is to encode the image into a unique bitstream with different layers: a basic layer containing the most important information, and enhancement layers which allow a progressive increase of the quality. Several dimensions of scalability can be considered, such as resolution, quality/Signal-to-Noise Ratio (SNR), temporal and region of interest.

Since digital holograms encode the whole 3D information of the scene, they contain a massive amount of data. Thus, it turns out that a holographic scalable coding is highly

desirable to reduce the transmission bandwidth. However, due to the irregular characteristics of holographic signals, conventional scalable image and video coding schemes would be sub-optimal when applied to holograms. Indeed, each pixel contains information about the whole 3D scene and adjacent pixels are highly complementary to form the right wavefront in the reconstruction domain. Thus, directional spatial predictors and down/up sampling algorithms used for spatial scalability are ineffective on hologram patterns. Moreover, the relationship between the distortion in the hologram and reconstruction domains is not straightforward, which makes the quality scalability more complicated. Finally, traditional image and video scalable encoders would not explicitly take into account the 3D visualization features of holographic data. Therefore, viewpoint scalability would not be possible with such schemes.

Finally, the generated bitstream is transmitted to the client for decoding, and the original image is then exactly (lossless) or approximately (lossy) retrieved. In streaming scenarios, the scalability permits the user to have a quick glimpse of the image while the refinement information is progressively downloaded. An efficient deployment of such compression architecture to holographic coding requires an appropriate transformation tool, which is able to interpret the 3D meaning of the interference patterns. State-of-the-art methods for digital hologram coding are exhaustively described in Chapter 2.

### 2.3.4 Display

Once the compressed digital hologram is decoded by the client, the light wave of the captured scene is reproduced from the retrieved hologram to enable 3D visualization. To this end, holographic displays utilize a beam shaping device called Spatial Light Modulator (SLM). Indeed, the object wavefront is replicated by spatially modulating the amplitude and/or phase of the reference beam. Theoretically, the highest reconstruction quality is achieved by SLMs with full complex modulation [74]. However, the few solutions proposed for the conception of such devices are still immature and need additional exploration. Alternatively, phase-only SLMs are commonly used to display holographic content thanks to their high diffraction efficiency. Most of the existing SLM displays in the market are based on two major technologies [99]: Liquid Crystal Devices (LCD) and Micro-Electro-Mechanical Systems (MEMS).

Given the coherent character of the reference light beam, the SLM's pixels size should be below $10\mu\mathrm{m}$ to permit its diffraction. Therefore, high resolutions are necessary to display large digital holograms with a wide FoV. To the best of our knowledge, the resolution of current commercially available SLMs do not exceed 3840 x 2160 (4K Ultra-High-Definition (UHD)), with a minimal pixel pitch of $3.74\mu\mathrm{m}$ [21]. With such restrictions, the short-term use-cases have been oriented to holographic Head-Mounted-Displays (HMD) [7]. Indeed,

since HMD display only a portion of the scene corresponding to the user's viewpoint, all the restrictions related to low-resolution and small viewing angle of digital holograms may be surpassed. Some promising prototypes have already been realized for Augmented Reality (AR) and Virtual Reality (VR) [79, 137, 209]. The expected long-term applications include high-resolution interactive projection systems, table-top holographic displays and holographic televisions.

Although holographic displays present several hardware challenges that need to be tackled, software algorithms enable the simulation of the optical digital hologram reconstruction by numerical computations. This offers a convenient way to validate the holographic generation and compression models.

The numerical calculations consist on modeling the light propagation using the scalar diffraction theory [91]. Assuming that the light wave is time-invariant, and both the diffraction aperture and propagation distance are larger than the light wavelength, the scalar wave equation is given by the well-known partial differential Helmholtz equation

$$\left(\nabla^2 + \left(\frac{2\pi}{\lambda}\right)^2\right) u = 0, \tag{2.8}$$

where $u$ is the light wave and $\nabla^2$ denotes the Laplace operator.

Let us consider a light wave propagating from a source plane to a destination plane separated by a distance $z_0$. According to the Rayleigh-Sommerfeld diffraction theorem [76], if the two planes are parallel, the light wave $u_d$ in the destination plane can be expressed as a convolution product between the light wave $u_s$ at the source plane and a kernel $\psi_{z_0}$, such that

$$u_d(x_d, y_d) = (u_s * \psi_{z_0})(x_s, y_s) = \iint_{\mathbb{R}} u_s(x_s, y_s)\psi_{z_0}(x_d - x_s, y_d - y_s)dx_s dy_s, \tag{2.9}$$

where $(x, y)$ and $(x', y')$ are the coordinates in the source and destination plane respectively.

According to the convolution theorem, the Angular Spectrum (AS) propagation provides an exact solution of the Helmholtz equation, such that

$$u_d(x_d, y_d) = \mathcal{P}^{AS}_{z_0}\{u_s\}(x_d, y_d) = \mathcal{F}^{-1}(\mathcal{F}\{u_s(x_s, y_s)\}\Psi_{z_0}(f_x, f_y)), \tag{2.10}$$

where $\mathcal{F}$ is the Fourier transform operator, $f_x$ and $f_y$ are the spatial frequencies, and $\mathcal{P}^{AS}_{z_0}$ is the AS propagation operator. $\Psi_{z_0}$ denotes the transfer function of the AS, given by

$$\Psi_{z_0}(f_x, f_y) = e^{j2\pi z_0\sqrt{\lambda^{-2} - f_x^2 - f_y^2}}. \tag{2.11}$$

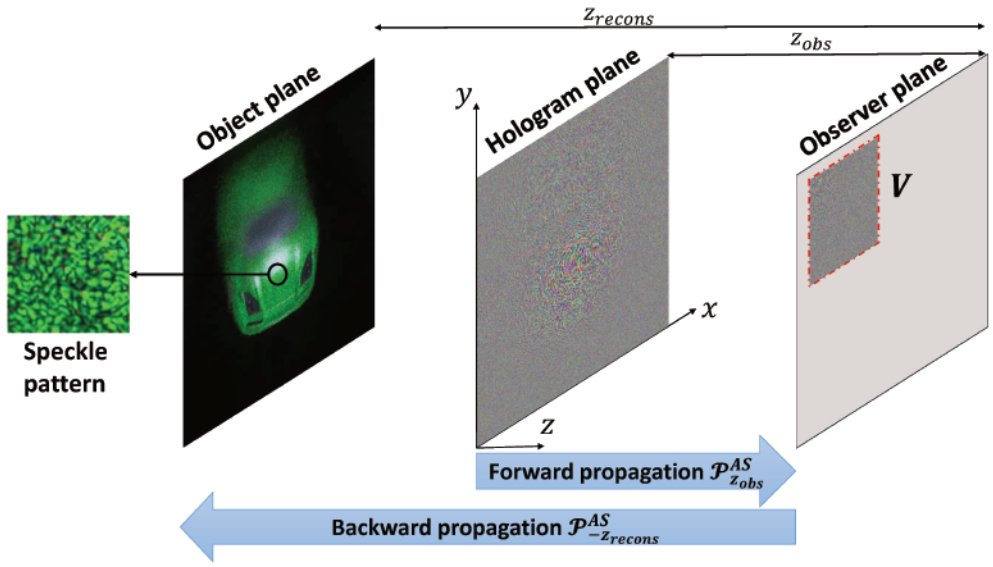



Fig. 2.9 Numerical reconstruction scheme of digital holograms: *First, the hologram is propagated to the observer plane. The obtained wavefield is cropped to a given viewing window. Finally, the numerical reconstructed view is obtained by propagating the cropped wavefield towards the object plane.*

Under some conditions on the distance $z_0$, the kernel convolution $\psi_{z_0}$ may have other formulations, such as the Fresnel and Fraunhofer approximations. More details about these solutions can be found in [92].

In the following of this thesis, the numerical reconstruction of digital holograms is performed using the AS propagation operator because of its accuracy. The numerical reconstruction scheme is illustrated by Figure 2.9. The hologram lies on the source plane, *i.e.* hologram plane ($z = 0$), the observer plane is placed at a distance $z = z_{obs}$, whereas the object (*i.e.* reconstruction) plane is located at the destination plane ($z = -z_{recons}$). First, the hologram is propagated to the observer plane using the AS propagation $\mathcal{P}^{AS}_{z_{obs}}$. Then, the diffracted wavefield is cropped to a given window corresponding to eye's pupils. The window $V$ is centered around the viewer's position. Finally, the cropped light wave is propagated to the object plane using $\mathcal{P}^{AS}_{-z_{recons}}$, and the reconstructed view is obtained.

As shown in Figure 2.9, the zoomed pattern of the numerical reconstruction presents some artifacts called *speckle*. Speckle noise is an optical phenomenon that occurs when multiple sources of coherent light interfere mutually [93]. The bright and dark dots appearing in the speckle pattern are caused by the superimposition of constructive and destructive light waves. When illuminating the hologram with a laser, the speckle effect is inherently present in the reconstruction. This effect can be partially reduced using some denoising techniques during the recording or reconstruction processes [17].

## 2.4 Conclusion

In this chapter, we first reviewed the major techniques used in real-3D display systems, and we showed that holography is considered as the most promising one. Then, we provided the diffraction characteristics and some statistical measurements of digital holographic signals, which are completely different from photographic images. Finally, we described the overall holographic transmission chain from content creation to display. The compression part of this chain will be deeply investigated in the next chapter.

# Chapter 3

# State of the art of digital hologram compression

This chapter provides an exhaustive state-of-the-art of digital hologram coding techniques. Most compression approaches have been evaluated using CGHs due to their convenient and flexible generation; however, these methods are still perfectly relevant for optically acquired digital holograms. It must be noted that several classifications of these methods have been proposed in the literature according to the compression mechanisms used by each of them [21, 51, 67, 166, 167]. The state of the art categorization proposed in this chapter is organized as follows. First, the approaches adapting conventional coding techniques for holographic data compression are described in Section 3.1. Then, hologram coding schemes exploiting 3D visualization features are listed in Section 3.2.

## 3.1 Adapting conventional coding techniques for digital holographic data compression

Since holograms are essentially pixelated 2D patterns, first attempts to compress them were based on conventional image coding techniques. These approaches can be classified into three categories: lossless coding-based, quantization-based and transform-based approaches.

### 3.1.1 Lossless coding-based approaches

Early works in digital hologram compression investigated the use of lossless data coding algorithms such as Lempel-Ziv (LZ77) [213], Lempel-Ziv-Welch (LZW) [198], Huffman [112] and Burrows-Wheeler (BW) [31]. These techniques have been applied on holograms gen-

erated by the phase-shifting interferometry, showing that BW's algorithm has the best compression ratio with an average of $1.95$ for the A-P representation and $4.66$ for the R-I one [145].

The coding gains obtained by pure lossless compression are unconvincing and would not allow encoding at low bit rates, which calls for lossy compression methods.

### 3.1.2 Quantization-based approaches

In quantization-based approaches, each input value is transformed to one quantized output value independently of other signal values. Quantization methods are split into two classes: scalar quantization used for one-dimensional data, and vector quantization applied on multi-dimensional input data.

**Scalar quantization-based methods** In [145], the authors applied several scalar quantization techniques to holographic patterns. They found that uniform quantization achieved better performance compared to downsampling: a compression ratio of $16$ was reached, using R-I representation, while maintaining a good reconstruction quality with a normalized cross-correlation of $0.98$. The performance of uniform quantization has also been evaluated for numerical and optical reconstructions of phase-shifting digital holograms [140]. Experimental results revealed that due to the uneven distribution of digital hologram values, the reconstruction quality was critically deteriorated when quantizing with less than $2$ bits, and that more than $6$ bits are necessary for visually lossless hologram coding.

To overcome this issue, several works investigated the use of non-uniform quantizers for hologram compression. In [176], the authors explored the use of iterative clustering techniques, followed by lossless encoding. Their proposed approach used the $k$-means algorithm by placing more centers in areas where the data is dense and less centers in sparse regions. The experiments were performed for different number of clusters using the BW' algorithm for lossless coding, and showed a better compression performance compared to uniform quantization.

Despite the satisfactory results produced by non-uniform quantizers, their iterative process is time-consuming and heavily depends on the data they have been trained on. To overcome these limitations, the authors of [177, 179] proposed two companding quantization approaches to enable a generic compression of digital holograms. These methods take benefit from the low complexity of uniform quantization, together with the superior performance of the non-uniform one. Indeed, a non-linear transform is first applied on fixed-interval sampling grids of the complex plane, followed by a uniform quantizer. Two companding grids are utilized, based on diamond and logarithmic spiral patterns. It has been found that

the performance of companding quantizers is comparable to the ones based on $k$-means clustering and Kohonen competitive neural network [62], with the advantage of being non-iterative and insensitive to moderate change in the holographic data distribution. In the same spirit, the technique introduced in [181] extracts clusters from the histograms of R-I or A-P hologram data to make decisions about the best quantization values.

**Vector quantization-based methods** Since holograms are complex-valued signals, they can be considered as vectors of dimension $2$, and are then suitable for vector quantization. In image compression, vector quantization exploit correlation between pixel blocks to have more accuracy in defining the clusters' centers [84]. The use of non-uniform vector quantization approaches has been explored in [178] using $k$-means clustering. The proposed quantizer tried various pixel block sizes ranging from $1 \times 1$ to $4 \times 4$. Experimental results revealed that the $1 \times 1$ vector version of $k$-means quantization outperforms the scalar one. Moreover, in contrast to common images, using larger blocks degrades the compression efficiency of the vector $k$-means quantizer.

In [205], the authors proposed a vector quantization using the LBG algorithm for compression of phase-shifting holograms. By evaluating the Rate-Distortion (RD) behavior of numerical reconstructions, results showed that compression based on LBG vector quantization outperforms adaptive scalar Lloyd Max quantization. The authors also demonstrated in other work [206] that the R-I representation is more relevant for vector quantization, whereas the A-P representation is better suited for scalar quantization. A detailed comparative study of scalar and vector quantization approaches can be found in [42, 123, 206].

The major shortcoming of quantization-based approaches is that they operate directly on the raw hologram without any prior transforms. For better holographic compression performance, all image coding pipeline modules (*cf.* Figure 2.8) must be performed.

### 3.1.3 Transform-based approaches

Transform-based approaches first transform the input holographic signal to another domain for more efficient quantization and encoding. Since digital holograms are encoded as 2D complex signals, an intuitive solution was to apply the transforms deployed by classical image and video coding standards (Joint Picture Expert Group (JPEG) and Motion Picture Expert Group (MPEG)), either on the R-I or A-P representation format. A second solution consisted of replacing the default transforms used in these two standards by others to better fit the irregular signal characteristics of holographic data. In this section, we list the conventional and adapted transforms belonging to JPEG and MPEG, which are employed in hologram compression.

**JPEG-based transforms** JPEG is one of the most popular codec for lossy image compression [195]. The JPEG specification follows the aforementioned image compression architecture depicted in Figure 2.8. First, the color space of the input image is converted from RGB to YCbCr, followed by a downsampling of the chrominance channels. Then, the image is split into blocks of $8 \times 8$ pixels. The particularity of JPEG consists on the use of 2D-DCTs, which are applied on the image blocks to obtain a representation in the frequency domain. Since the HVS is more sensitive to smooth brightness changes, the DCT coefficients belonging to high frequency components are quantized with low precision. This is done by dividing the coefficients matrix by a quantization matrix, and then rounding to the nearest integer. The quantization matrix values can be defined based on psychovisual threshold experiments. Finally, for each block, the quantized coefficients corresponding to similar frequencies are grouped together using a "zigzag" scan followed by Run-Length Encoding (RLE) algorithm [90], and Huffman coding is then performed. For a progressive bitstream, the similar-positioned coefficients of all blocks are simultaneously encoded during each scan. The important property of JPEG is to enable a wide range of compression ratios and quality, which can be controlled by a simple amplification or mitigation of the quantization matrix factor.

Although the JPEG standard is one of the referenced representation format of photographic images, it yields high compression performance only for images with smooth variations of texture and color. When it comes to digital holograms, it is clear that such characteristic is not fulfilled (*cf.* Section 2.2). Therefore, it turns out that the default transform deployed by JPEG codec must be adapted to meet the irregular properties of holographic signals.

The authors of [211] proposed an enhanced version of JPEG for off-axis holograms compression. Since most energy of off-axis holograms is contained in the first-order term having high frequencies, an adaptive mask is applied on each block prior to quantization in order to discard all the DCT coefficients corresponding to zero-order terms, *i.e.* with low amplitude values. The proposed binary mask $M$ is generated according to the following formula

$$M(f_x, f_y) = \begin{cases} 0 & \text{if } f_x, f_y = 0 \text{ or } S(f_x, f_y) \leq T \\ 1 & \text{otherwise,} \end{cases} \tag{3.1}$$

where $S(f_x, f_y)$ is the cumulated amplitude, block by block, of the coefficients corresponding to the spatial frequency coordinates $(f_x, f_y)$. $T$ is a coarse threshold defined as the mean value of $S$ over the $64$ DCT coefficients.

The experiments shows that the enhanced version of the JPEG codec outperforms the original one in terms of compression ratio and phase retrieval for both noisy and noiseless conditions. Due to the use of default JPEG quantization matrix [195], the introduced method did not attain optimal performance.

In order to minimize the distortion caused by the default DCT used in JPEG for given bit rate budget, an optimal quantization table must be designed. To this end, an efficient adaptive DCT quantization table was explored in [37]. The proposed technique is based on RD optimization using the Lagrange multiplier algorithm. The bit rate is estimated as the sum, over the $64$ spatial frequency positions, of the entropies corresponding to the quantized DCT coefficients over all the blocks. The block's distortion is expressed as the sum of the distortions induced by the $64$ quantized DCT coefficients. The optimized version has been integrated in the baseline JPEG codec achieving superior compression efficiency compared to the former, and competitive coding results compared to MPEG. Moreover, the proposed approach considerably reduces the computational burden of coding and decoding processes.

Another category of works aiming at improving the JPEG-based holographic coding relies on the post-processing of decoded holograms in order to reduce the artifacts induced by compression. A novel approach has been introduced in this direction, incorporating JPEG standard and deep convolutional network [117].

With the emergence of wavelets as an efficient tool for time-frequency signal representation [58, 59], the spectral decomposition in JPEG using DCT has been replaced by a DWT, giving rise to a new standard called JPEG 2000 [46]. A discrete wavelet is defined by real or complex coefficients forming a band-pass filter bank. DWT consists of applying a pair of low and high-pass discrete filters to the input signal, and the output of each filter is then downsampled. A multiresolution decomposition can be achieved by further expanding the low frequency sub-band. DWTs are constructed either by the lifting scheme [184] or by convolution. Two bi-orthogonal real-valued DWTs are principally deployed by JPEG 2000: **(1)** LeGall-Tabatabai $5/3$ wavelet transform [124], with $5$ coefficients for the low-pass filter and $3$ for the high-pass one. It results in integer transform coefficients, and enables lossless compression with a perfect reconstruction [34]. **(2)** Cohen–Daubechies–Feauveau (CDF) $9/7$ wavelet transform [59], which is irreversible due to its floating-point taps, making it well-suited for lossy compression [9].

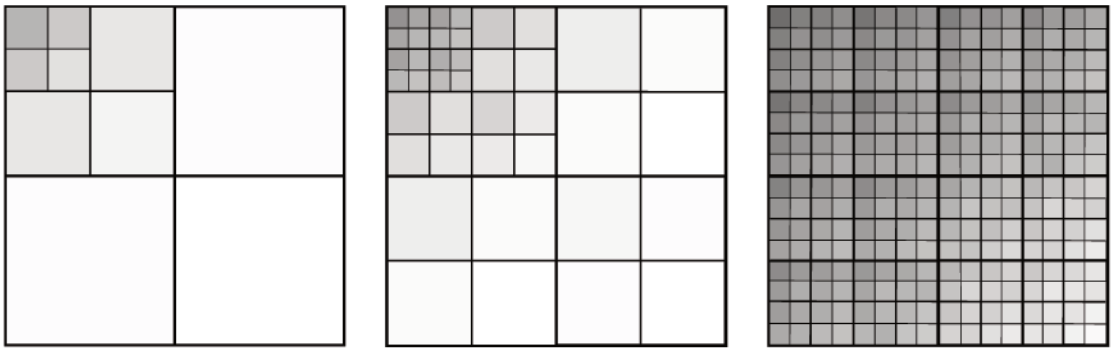
Another source of strength for the JPEG 2000 codec compared to its predecessor JPEG is its efficient entropy encoder built around the sophisticated Embedded Block Coding with Optimal Truncation Points (EBCOT) algorithm [187]. First, the sub-bands obtained by DWT are partitioned into non-overlapping rectangular blocks (*code-blocks*) of size $32 \times 32$ or $64 \times 64$. Then, the quantized coefficients of each single code-block undergo a bit plane

coding, starting with the most significant bits and progressing to less significant bits. This is performed by EBCOT through three passes, namely the significance propagation pass, the magnitude refinement pass, and the cleanup pass. For each pass, the selected bits and their corresponding contexts are then encoded using a context-based binary arithmetic coder. Finally, the bit plane coding passes of all code-blocks are grouped in packets, which are in turn distributed across one or more layers, considering all sub-bands. To form an optimal embedded bitstream, the packets length and their attribution to several layers should respect two constraints: a progressive increase of image quality with each decoded layer, and RD optimization.

Thanks to the multiresolution decomposition structure of DWT and the progressive encoding process of EBCOT, the JPEG 2000 bitstream is highly flexible and scalable, with excellent RD performance. Consequently, JPEG 2000 has been deployed for the compression of phase-shifting interferometry holograms [55], reaching a compression gain of $35\%$ over JPEG, for a normalized root Mean Square Error (MSE) of $0.7$ approximately. However, it has been shown that the efficiency of JPEG 2000-based holographic coding is restricted to high bit rates range.

In [23], the authors suggested the use of JPEG 2000 with full-packet decomposition instead of the default Mallat dyadic decomposition. This choice has been motivated by the fact that a significant part of the hologram energy is contained in the high-pass bands, which need to be further decomposed for a better compression performance. Such functionality is specified by the arbitrary decomposition structures of the JPEG 2000's Part 2 extension. Indeed, it allows a more general decomposition where the low-pass sub-bands can be split in horizontal and/or vertical directions, followed by an additional sublevel decomposition of high-pass sub-bands for each resolution level. Examples of Mallat, partial packet and full packet decomposition structures using $3$ levels are shown in Figure 3.1.

Since the full packet decomposition of the JPEG 2000 standard is limited to $3$ resolution levels within high-pass sub-bands, the authors of [23] modified the code-stream configuration and signaling syntax of the codec to handle more than two decomposition levels. To this end, they used an ordered array of split-operations performed on a pile of available sub-bands. The coding performance of the obtained extension has been compared to the default JPEG 2000's framework with a 4-level Mallat decomposition, using a database of $15$ off-axis microscopic digital holograms. In the lossless scenario, compression results report an average bit rate reduction of $0.26$ bpp when using the designed 4-level partial packet decomposition. In the lossy mode, a Peak Signal-to-Noise Ratio (PSNR) improvement of $5$dB is achieved by the 4-level full packet decomposition. Although the introduced approach enabled considerable compression enhancement, it is unlikely to perform at macroscopic



(a) Mallat dyadic. (b) Partial packet. (c) Full packet.

Fig. 3.1 Examples of 3-level decomposition structures in JPEG 2000. *The dark regions corresponds to low-frequency sub-bands*

scale, since high frequency components and speckle noise would be more prominent in the holographic patterns.

The second category of transform-based hologram coding approaches aims at adapting the default DCT/DWT deployed by JPEG/JPEG 2000 or using other wavelet transforms which may improve the fitting of hologram's untypical properties.

First, the use of popular wavelets transforms for digital hologram coding has been investigated in [43, 44]. Seeking the most suited wavelet transform, the authors compared the compression performance obtained by 7 popular wavelet schemes followed by a thresholding or quantization of wavelet coefficients. Experiments revealed that Haar's and coiflet's wavelets outperform all the other ones as well as uniform quantization. In a subsequent work [41], a library of 51 wavelets has been utilized to transform synthesized and optically recorded digital holograms. The compression efficiency using R-I and A-P parts has been compared, showing that the coding gain achieved for A-P is 2-4 times lower than R-I.

However, these works were restricted to one resolution level and did not consider lossless coding. Moreover, the experiments were performed only on digital holograms generated from 2D images. In [180], the authors compared 53 discrete 1D wavelets with several resolution levels (1, 3, 10 and 20) in order the determine the optimal wavelets parameters to compress holograms generated from 3D objects. The conducted tests showed that the best compression performance is obtained for 3 resolution levels. Furthermore, it has been proven that no wavelet function outperformed the other ones.

Although classical wavelet transforms provide a practical and easy solution for digital hologram compression, they do not achieve the same convincing coding performance as for natural images due to the specific signal properties of holographic patterns. To improve the compression efficiency of wavelet transforms-based approaches, some works considered the

adaptation of conventional wavelets schemes to better analyze the hologram interference fringes characterized by their strong orientations.

Indeed, the authors of [11] presented a novel hologram coding approach by combining classical wavelet and Bandelet transforms [126]. First, Haar wavelets are used to decompose the hologram into different sub-bands, each divided into small blocks. Then, orthogonal Bandelets [125] are used to detect the interference fringes' orientations: each block is decomposed along multiscale vectors that are elongated in the direction of a geometric flow. The latter indicates directions in which the block pixels have regular variations. By choosing the best geometric flow, the block pixels may be appoximated with minimal information loss. Finally, the transform coefficients of all blocks are entropy encoded. The proposed approach has been compared to hologram compression methods based on common wavelet transforms and DCT, achieving a higher compression ratio for the same reconstruction quality.

For the same purpose, Directional Adaptive-DWT (DA-DWT) has been deployed for compression of microscopic off-axis holograms [23, 24]. DA-DWT locally adapts the filtering directions to image features based on their local orientations [38]. The proposed compression scheme is implemented in JPEG 2000's architecture where the default DWT is replaced by DA-DWT. To avoid an encoding budget increase caused by the signaling of a large direction vectors set, DA-DWT are performed only on low-pass sub-bands considering a discrete set of $11$ vector pairs. This set is then assessed for eack codeblock in order to determine the optimal pair to be retained. Due to the inherent orientations present in microscopic holographic images, DA-DWT improves the energy compaction of interference fringes leading to better compression performance compared to standard DWT: an average bit rate reduction of $0.46$bpp was reported in the lossless mode, and a PSNR gain of $5.47$dB in the lossy mode.

Despite the remarkable improvements achieved by DA-DWT-based compression, its performance has not been evaluated in case of macroscopic holographic content. Furthermore, the coding computational complexity entailed by this method is not negligible, especially for high-resolution holograms.

An other category of work explored the use of lifting schemes [184] introduced as an alternative to the filter-bank-based techniques commonly used for wavelets design. Accordingly, a new joint multiscale hologram decomposition based on the separable vector lifting concept has been developed in [201]. Instead of independently encoding the hologram parts of complex-valued representations, the proposed scheme exploited the dependencies noticed between components of the shifted distance representation in order to enhance the compression efficiency. The wavelets are constructed through two lifting reversible processes: prediction and update, and the resulting transform coefficients are encoded using

EBCOT. Experimental results revealed that a significant gain of 2dB in terms of PSNR has been reached by the presented vector lifting scheme over the JPEG 2000-based independent compression using the CDF $9/7$ wavelet transform.

To better analyze the 2D isotropic features of holographic data, a non-separable version of the aforementioned vector lifting scheme has been designed [202]. For an efficient low-complexity decomposition, the prediction filters are optimized at each resolution level by minimizing the variance of the detail coefficients. Then, the update filter is optimized by minimizing the square error between the estimated signal and the decimated coefficients obtained by an ideal low-pass filter. The conducted tests confirmed the superiority of the non-separable vector lifting scheme compared to the separable one: up to $10\%$ and $0.8$dB of bit rate saving and PSNR of the reconstructed hologram, respectively. Moreover, it has been found that the optimal compression performance is attained for a prediction lifting operator with length of $12$: coding gain of $80\%$ and PSNR improvement of $12$dB compared to a filter with length of $2$.

In [18, 19], wave atoms have been investigated as a sparsifying transform adapted to holographic signals. Wave atoms have been proposed in the literature as oriented and multiscale transforms, alternative to 2D wavelet packets [63]. In addition to their near-optimal space-frequency localization, wave atoms form an orthonormal basis of directional wave packets and are particularly well suited for representing oscillatory patterns. The authors of [18] developed a generic hologram coding framework based on wave atoms coding using standard tensor products of the 1D transforms. To enable an efficient sparsification of holograms recorded at any distance, the obtained transform coefficients are further rearranged in the spatial or frequency domain using an adequate clustering. Finally, the rearranged coefficients are fed to the scalar quantizer and entropy coder of JPEG 2000 codec. The proposed method has been compared to JPEG 2000's extensions [23], showing significant PSNR gain of $0.9$-$7.9$dB.

**MPEG-based transforms** Since its establishment in $1988$, MPEG launched several standards which are extensively deployed for audio and video coding [197]. Each standard is composed of different parts, each covering a certain aspect of the whole specification. Two MPEG standards are widely used for image and video coding: MPEG-4 Part10 also known as H.$264$ or Advanced Video Coding (AVC) [199], and MPEG-H Part2 also known as H.$265$ or High Efficiency Video Coding (HEVC) [182]. In the following, the working principle of these two codecs is described, and the transform-based hologram coding approaches based on them are listed.

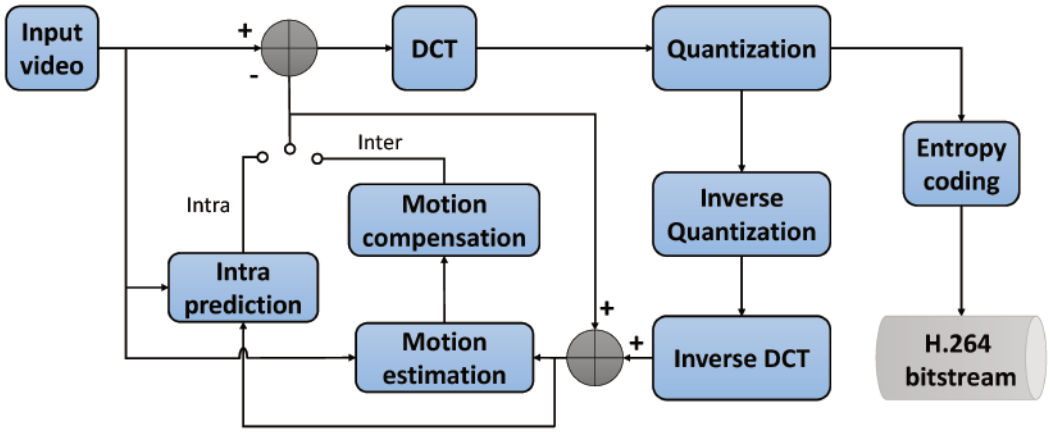



Fig. 3.2 Simplified structure of the H.264/AVC codec

AVC is mainly designed for digital video compression, and basically uses the block motion compensation technique as a prediction mechanism. Indeed, in the inter-coding mode, the temporal redundancies from previous (and future) video frames are exploited to predict the content of the current frame, called P-frame (B-frame). AVC may also be constrained to intra-coding mode, which considers only the spatial correlation whithin a single frame, called I-frame. The basic AVC's coding structure is given in Figure 3.2.

First, the video frames are assembled into several Groups Of Pictures (GOP), where the first frame of each GOP is encoded as I-frame and the next ones as (P-) B-frames. Then, each frame is partitioned into different Macro-Blocks (MB)s of size $2^i \times 2^j$ pixels, where $i, j \in \{2, 3, 4\}$. For each MB, and depending on the coding mode, the encoder searches in the previous frame or the encoded part of the current frame to find the best matching MB. Finally, the original MB is subtracted from its approximation, and the obtained residual undergo a DCT transform followed by a quantization and entropy coding to generate the H.264 bitstream.

For intra-coding mode, the prediction is performed from the already encoded neighboring blocks located in the top and left of the current MB. AVC enables up to 8 directional modes, and the encoded I-frames are used as reference for the following frames.

For inter-coding mode, the prediction is carried out with respect to previously encoded (intra-) inter-frames. First, the motion estimation process generates the motions vectors representing the estimated displacement of each MB, which are then compensated from the reference frame to form the prediction of the current frame.

Since AVC was originally developed for video coding, the first attempt was to apply it on digital holographic videos [52, 204]. AVC may also be utilized to encode still holographic images by using the intra-coding mode only. The hologram compression benchmarks

comparison presented in [157] reveals that it has better efficiency than JPEG, but still outperformed by JPEG 2000 in most cases.

In order to take benefit from the AVC inter-coding mode, the authors of [172] proposed a novel scanning-based method to generate a video sequence from the transform coefficients of the segmented holographic pattern. Indeed, the digital hologram is first divided into different segments with equal sizes from $8 \times 8$ to $512 \times 512$ pixels, which are transformed to the frequency domain using 2D-DCT. Then, the transformed segments are scanned by temporal ordering to form a video. Finally, the code-stream is obtained by an hybrid coding scheme combining three different encoders: $(i)$ AVC ensures the inter-coding of the generated video, $(ii)$ the differential pulse code modulation (DPCM) encoder is applied on DCT coefficients with very large values, $(iii)$ a conventional lossless codec, such as ZIP, is used to compress the sign bitplanes.

For a normalized correlation over $0.95$, the proposed approach achieved a compression ratio of $50$:$1$, with an average increase of $0.2$ compared to JPEG 2000. Moreover, the visual quality of the reconstructed image from the compressed hologram and the original one, at a compression ratio of $20$:$1$, is indistinguishable when the proposed coding scheme is used, whereas it is severely degraded when using JPEG 2000.

The second standard broadly deployed for video coding is HEVC, which is the successor of AVC. It enables a bit rate gain up to $50\%$ at the same level of video quality, with a resolution up to 8K UHD. The working concept of HEVC is similar to AVC with some ameliorations [149], such as the expansion of the MB's size to $64 \times 64$ pixels, improved intra-coding with $33$ directional modes plus the DC and planar prediction modes, and the advanced motion vector prediction, where several most probable candidate vectors are estimated. These changes increase the codec's computational expense especially at the encoder side.

HEVC has been first applied to phase-shifting holographic videos using the inter-coding mode [204], which attains higher bit rate gain than AVC. The coding efficiency of intra-coding mode has been investigated for still digital holograms both before and after display, showing the best trade-off between compression ratio and reconstruction quality compared to all aforementioned codecs [4, 157], including AVC intra-coding mode, JPEG, JPEG 2000 and its extensions presented in [23]. Despite its noticeable coding improvements, HEVC's architecture needs special transforms to better match the specific features of holographic patterns.

In [158], the authors introduced an extension of the HEVC intra-coding mode using adaptive transforms. Inspired by [12, 174], they used directional dependent transforms instead of conventional DCT. In this way, the codec could take benefit from the directionalities that

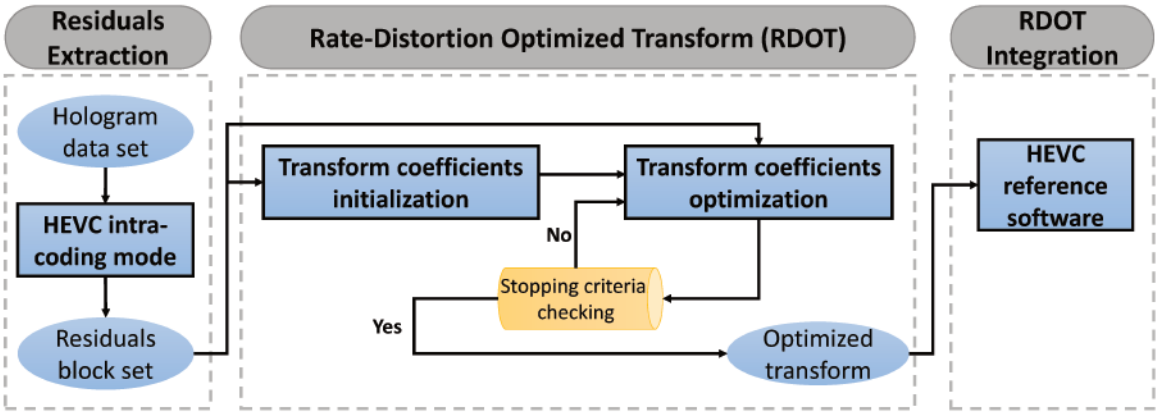



Fig. 3.3 Simplified pipeline of the HEVC's extension based on Rate-Distortion Optimized Transform

may appear in hologram features in order to enhance the compression performance. The different phases of the proposed HEVC's extension are illustrated by Figure 3.3.

The first phase of the presented architecture is **Residuals Extraction**. First, a training set of residuals is generated from the source holographic data using the HEVC intra-coding mode. Then, a representative block set of residuals is extracted for each directional prediction mode and transform unit size.

The second phase is **Rate-Distortion Optimized Transform (RDOT)**, which aims at designing an optimal set of transform basis functions from the collected intra residuals data. This phase is divided into two parts: (1) *Initialization* part, where an initial sparse set of transform coefficients is computed using the Kronecker product followed by a hard-thresholding process. The initial rate and distortion may then be estimated. (2) *Optimization* part, which involves an iterative process that refines the initial transform until it converges to optimized transform. In this part, a singular value decomposition is used to obtain the transform coefficients at each loop [174].

During the initialization and optimization parts, the rate and distortion are estimated using $L_0$ and $L_2$ norms, respectively. The selected Lagrangian cost function used by the RDOT phase is summarized by the following formula

$$T_{opt} = \underset{G}{argmin} \sum_{i} \underset{c_i}{min} \left\{ ||r_i - \hat{r}_i||_2^2 + \lambda ||c_i||_0 \right\}, \tag{3.2}$$

where $T_{opt}$ is the optimized transform, $r_i$ is the original residual block, and its reconstruction is denoted by $\hat{r}_i$, $c_i$ is the quantized transform coefficient, and $\lambda$ is the Lagrange multiplier.

After including the appropriate scaling factors to the obtained optimized transforms, they are finally integrated to the HEVC reference software in the **RDOT Integration** phase.

Notwithstanding the considerable compression improvements achieved by the approaches adapting conventional image and video coding standards transforms to better compress digital holographic data, some of them have only been tested for holograms acquired from synthetic 2D scenes or at microscopic scale. Moreover, the heavy-computational burden induced by some of them is impractical with high-resolution and dynamic holograms. Most importantly, while transform-based compression techniques seek the design of optimal sparsifying transforms well-suited to the irregular semantics of holographic signals, they do not leverage their 3D visualization features. In other words, being able to interpret the 3D meaning of holographic patterns could enable better compression performance and other interesting functionalities such as viewpoint and quality scalability as well as hologram edition.

## 3.2 Holographic data compression exploiting 3D visualization features

### 3.2.1 Content-aware coding

As concluded in the previous section, designing alternative transforms to sparsely represent the spatial holographic signal is necessary but not sufficient for an efficient compression. Indeed, understanding and interpreting the 3D meaning of holographic fringes is required to extract and encode the relevant information about the recorded scene. As an ultimate goal, the transformation tool should provide a straightforward relation between the changes occurring in the scene and the hologram's local frequencies. In this way, the encoder would benefit from the spatial and temporal redundancies initially present in the source domain to reduce the encoding cost. This calls for a new category of hologram coding methods, which aims at developing new holographic codecs taking into account knowledge about the scene.

A family of approaches that falls within holographic content-aware coding is the compression with respect to the scene depth. In [151], an important duality between the scale parameter of wavelets and the wave propagation depth was underlined, opening doors for space-depth analysis. Based on this finding, a new multiresolution wavelet-like bases called *Fresnelets* has been introduced in [131], specially tailored for digital hologram processing. Fresnelets were obtained by applying the Fresnel transform to B-spline wavelets [189], allowing the modelisation of light diffraction. The use of this new type of wavelets for phase-shifting digital hologram compression has been investigated in [56]. The authors proposed two coding schemes of the hologram Fresnelet coefficients: the first is based on

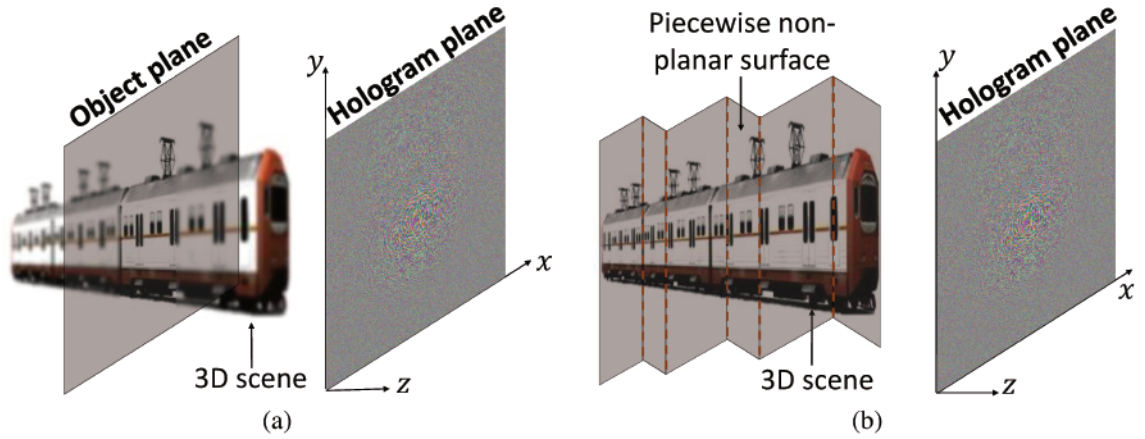



Fig. 3.4 Content-aware hologram coding approaches: *In the object plane coding method (a), the light wave emitted by the hologram is encoded in a plane crossing the scene, in which the object appears sharp and in-focus. However, deep scenes cannot be fully brought to focus using a single propagation. This problem may be addressed by using propagation over a piecewise non-planar surface fitting the depth profile of the scene (b). The train 3D model is courtesy of TurboSquid (www.turbosquid.com).*

uniform quantization followed by BW's algorithm for lossless encoding, whereas the second employed the Set Portioning In Hierarchical Trees (SPIHT) coding paradigm [165].

Encoding holograms in the object (*i.e.* reconstruction) plane has been proposed as an alternative solution to holographic depth-aware compression [14, 15, 57]. These approaches consist on back-propagating the hologram from its acquisition plane towards the object plane (*cf.* Figure 3.4a), and then compressing the obtained wavefield. When holograms have a large focusing depth or are recorded from flat objects and if the back-propagation distance is carefully chosen, the refocused hologram will appear sharp, making the use of conventional image codecs powerful as in case of natural images.

The principle of object plane coding has also been deployed for hologram lossless compression using an integer Fresnel transform [25]. The authors utilized the lifting scheme concept to implement a reversible integer version of the Fresnel transform in both convolutional and Fourier forms. They demonstrated that the inverse Fresnel transform can be computed from its complex conjugate, thereby obtaining an integer approximation suitable for lossless coding. The proposed method was evaluated on several CGHs and optically acquired digital holograms, achieving bit rate savings from $0.12$bpp up to $2.83$bpp compared to the default lossless JPEG 2000 implementation.

Despite the significant coding performance enhancement achieved by the Fresnelet-based compression approach and object plane coding, their efficiency is limited to nearly-flat scenes as in microscopic holography since they operate only for fixed reconstruction distances.

Consequently, changing the focusing depth for deep scenes will result in critically distorted reconstructions [54]. Moreover, back-propagating for encoding and forward propagation for decoding and display may lead to expensive computational cost, which prevents the use of solutions based on multiple propagations depending on the depth information.

To handle the wide depth range of macroscopic holograms, the concept of conventional Linear Canonical Transforms (LCT), including Fourier and Fresnel transforms, has been generalized to model diffraction between arbitrary non-planar surfaces [26] (*cf.* Figure 3.4b). First, a thorough study of symplectic properties in the Time-Frequency (TF) domain allowed the authors to construct a valid invertible operator, which is a function of the scene's depth under some conditions related to surface shape, hologram pixel-pitch, and wavelength. Then, from a piecewise polygonal approximation of the depth profile, the proposed unitary transform can be used to approximate the hologram at different distances of propagation. It must be noted that these transform can be further combined with other wavelets used for compression to achieve a sparse representation.

In the conducted tests, the proposed transform has been integrated in the default JPEG 2000 codec using CDF $9/7$ as sparsifying wavelets, and has been compared to Fresnelet-based approach and default JPEG 2000. The objective evaluation using RD confirms the superiority of the introduced method. Furthermore, the subjective evaluation given by numerical reconstructions was very appealing, showing that depth information and sharp details are well preserved by the proposed approach, and strongly impacted when using JPEG 2000 and Fresnelet-based especially at lower bit rates.

Although the latter work introduces a new coding paradigm with solid mathematical foundations, its validity depends heavily on the geometry of the scene and necessitates a prior approximation of its depth information. Moreover, this method will fail in case of objects placed directly behind each other in wide FoV, since a single depth map is insufficient to represent such scene.

### 3.2.2 Scalable coding

Exploiting the 3D visualization properties of holographic data could enable also a Scalable Hologram Coding (SHC). The SHC methods are essentially relevant for holograms with small pixel pitches, *i.e.* large FoV.

In the proposed method of [128], the hologram is represented as a light-field by a spatial cutting into small apertures, each corresponding to a given visualization angle. Then, the reconstructed sub-holograms are encoded using AVC's intra-coding mode. Despite the significant transmission bandwidth reduction allowed by the proposed viewpoint scalable coding scheme, a back-propagation from the object plane into the hologram one is required

after decoding, which is extremely time-consuming. Moreover, when considering a viewer distant from the hologram plane, the reconstructed sub-holograms suffer from a strong speckle noise. Thus, this approach is only valid for observers placed at the hologram plane, which is an unreasonable use-case for holographic visualization systems.

To cope with these shortcomings and handle arbitrary viewing positions, the authors of [191, 192] investigated the suitability of Gabor/Morlet wavelets for a view-dependent representation of digital holograms, emphasizing the importance of TF localization for a precise sub-holograms extraction. In a subsequent work [193], they designed a framework for view-dependent hologram representation and adaptive reconstruction, and explained the procedures of wavelets analysis and coefficients pruning. Although the theoretical aspects of this approach have been studied, no significant compression results nor realistic visual evaluation were reported.

Contrary to content-aware coding methods, the aforementioned scalabe coding schemes do not require any prior knowledge about the scene's geometry.

## 3.3 Conclusion

In this chapter, we reviewed the state of the art of hologram coding approaches, which can be grouped into two major categories. The first class consists of adapting conventional image and video coding standards for holographic data compression. The majority of methods belonging to this category principally modify the transform operator to better fit the untypical characteristics of digital holographic data, and then obtain the sparsest representation in terms of compression. The second class exploit the 3D visualization properties of digital holograms to define a straightforward relation between them and the 3D content they represent. The advantage of approaches belonging to this category is that the spatial and temporal correlations learned from the 3D scene would be beneficial not only for compression purposes, but could also provide other functionalities such as viewpoint and quality scalability, motion estimation and hologram editing.

The contributions of this thesis fall within the second class and seek the design of scalable coding schemes specifically dedicated to holographic data.

# Part III

# Digital hologram compression

# Chapter 4

# Hologram compression using Matching Pursuit with Gabor wavelets

The realization of an efficient hologram compression system depends on the way the holographic patterns are decomposed and how their spatial frequencies are analyzed to extract the redundancies within the recorded scene. It turns out that this task is highly challenging, because of the specific characteristics of holographic contents. Indeed, as explained in Section 2.2, digital holograms exhibit non-local features since 3D information of the captured object is scrambled over all the holograms pixels. As a result, conventional compression tools based on pixels decomposition did not take advantage of the transform mechanisms to predict the variations of holographic signals.

To overcome the limitations caused by the classical pixel-based decomposition, we propose to expand the hologram into a set of light beams, which is more adequate with its diffractive character and may enable a better compression efficiency. To achieve this, Gabor wavelets are used as a TF analysis tool to represent the light beams diffracted by local holographic patterns. To allow a compact representation suitable for compression, the Matching Pursuit algorithm is applied on the redundant Gabor expansion.

The remainder of this chapter is structured as follows. First, the light beams-based decomposition using Gabor wavelets is presented in Section 4.1. Second, the use of the Matching Pursuit algorithm is explained in case of holograms, and a Graphics Processing Unit (GPU) based implementation is described in Section 4.2. Then, Section 4.3 provides the different design choices of the proposed encoder. Experimental results of the introduced coding scheme are analyzed in Section 4.4.

## 4.1 Time-frequency representation of digital holograms with Gabor wavelets

The purpose of this section is to illustrate the use of Gabor wavelets for decomposing the hologram into a redundant set of light beams. Although such representation may be more suitable to apprehend the 3D meaning of the diffractive features, it implies an accurate TF analysis due to the non-local nature of holographic patterns. To this end, the basic mechanisms of TF analysis are briefly reviewed. Then, the use of Gabor wavelets as a TF analysis tool of digital holograms is explained. Finally, we describe how a redundant Gabor expansion is obtained in case of holographic signals.

### 4.1.1 Time-frequency analysis

The discipline of signal analysis involves all the transforms, representations and statistical measurements which are able to provide a meaningful interpretation of the signal content. For stationary signals which have a constant or periodic variation of the spectral information over time (or space in the case of images), the analysis can be done either from *time-domain* (*space-domain*) or *frequency-domain*. In the first case, the signal can be described by computing its temporal (spatial) statistics (mean, standard deviation, correlations...). In the second case, the FT is applied to the signal and its content is then analyzed from the Fourier spectrum. However, most of the real-life and natural signals (*e.g*: music, radar images, photographic pictures...) exhibit time-varying spectral content. Consequently, the interpretation of such signals using only spectral analysis without any temporal information or vice-versa is difficult, indeed even unrealizable.

To provide a complete representation of non-stationary signals allowing a structural description of their content, TF analysis has emerged as an efficient solution since it combines both temporal and spectral analysis simultaneously [48]. To achieve a good TF analysis, one needs to ensure the best trade-off between time and frequency localization. Early works investigated the use of Short-Time Fourier Transform (STFT) as the most basic solution for TF analysis [6]. It consists of dividing the signal into different segments of short durations using a window function, which are then Fourier-transformed. Although the STFT has the advantage of analyzing locally the spectral information of the signal, its major drawback is the lack of adaptativenenss in the TF domain. Indeed, the frequency resolution depends on the length of the analysis interval, and thus degrades substantially as the size of the interval is reduced. This is related to the *Heisenberg's uncertainty principle* [105] which assumes that a high temporal resolution and frequency resolution cannot be achieved at the same time.

This principle is illustrated by the following inequality

$$\sigma_t \sigma_f \geq \frac{1}{4\pi}, \tag{4.1}$$

where $\sigma_t$ and $\sigma_f$ represent the standard deviations in time (space) and frequency domains of the basis function, respectively. In [77], Gabor demonstrated that the inequality 4.1 becomes equality if and only if the window function used for the STFT is a complex Gaussian. This work leads to the well-known Gabor Transform providing the optimal trade-off for the TF localization, since it minimizes the uncertainty constraint.

Alternative techniques for TF analysis have been proposed, including wavelet transforms [58], Hilbert–Huang Transform [110], Wigner Distribution Function [47] and LCT [142] which are a generalization of the FT.

### 4.1.2 Gabor wavelets

The Gabor mother wavelet is the product of a complex exponential and a Gaussian function. In its 2D continuous form, it is defined by

$$g(x,y) = e^{-\frac{1}{2}\left(x^2+y^2\right)} e^{i2\pi(xf_{x_0}+yf_{y_0})}, \tag{4.2}$$

where, $(x, y)$ is the 2D spatial variable and $(f_{x_0}, f_{y_0})$ is the 2D frequency.

In their 2D form, a family of continuous Gabor wavelets is obtained by shifting, dilating and rotating the 2D mother wavelet, such that

$$g_{(s,\theta,\mathbf{t})}(x,y) = \beta\, g\left(R^{-\theta}\left(\frac{x-t_x}{s}, \frac{y-t_y}{s}\right)\right), \tag{4.3}$$

where $\beta$ is a normalization constant, $\mathbf{t} = (t_x, t_y)$ is the spatial shift vector parameter, $s$ $(s > 0)$ the frequency dilation parameter and $R^{-\theta}$ the rotation matrix defined by

$$R^{-\theta} = \begin{pmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{pmatrix}. \tag{4.4}$$

In their discrete form, 2D Gabor wavelets are defined by

$$g_{l,p}[x,y] = \beta\, e^{-\frac{f_l^2}{2}\left(a^2+b^2\right)} e^{2\pi i a f_l}, \tag{4.5}$$

where

$$\begin{cases} a = (x\cos(\theta_p) + y\sin(\theta_p))\,\Delta \\ b = (y\cos(\theta_p) - x\sin(\theta_p))\,\Delta \\ f_l = \dfrac{f_0}{s_l} \qquad\qquad 1 \leq l \leq L \\ \theta_p = \pi r_p \qquad\qquad 1 \leq p \leq P. \end{cases} \tag{4.6}$$

In Eq. (4.5), the dilated frequency $f_l$ and the orientation angle $\theta_p$ are discretized with $L$ and $P$ levels, where $l$ and $p$ are their discretization indexes, respectively. $s_l$ and $r_p$ define the frequency dilation and rotation parameters, and finally $(x, y) \in \mathbb{Z}^2$ is the 2D discretized spatial parameter, such that

$$\begin{cases} -\dfrac{M_x}{2} \leq x < \dfrac{M_x}{2} \\ -\dfrac{M_y}{2} \leq y < \dfrac{M_y}{2}, \end{cases} \tag{4.7}$$

where $(M_x, M_y)$ is the wavelet resolution.

### 4.1.3 Gabor expansion of digital holograms

As mentioned in Section 2.2.2, digital holograms fall into the category of non-stationary signals. Therefore, TF analysis is required to enable a better understanding of their patterns. Moreover, the light propagation tends to spread out the features that are initially well localized in the object domain. Thus, an adequate transform is required to minimize the energy spreading.

Among various TF analysis tools, Gabor wavelets are considered among the best candidate for hologram representation [191]. Indeed, they provide an optimal localization in both time (space) and frequency domains. Since the frequency information of holograms is related to the diffraction angle (*e.g.* grating equation), a good TF localization property would enable an accurate extraction of the local directions of diffraction. In the following, we underline how the Gabor wavelets could model the light beams diffracted by hologram patterns.

Let us consider a Gabor wavelet $g_{(s_0,\theta_0,\boldsymbol{t_0})}$ located at position $\boldsymbol{t_0} = (t_{x_0}, t_{y_0})$ and illuminated by a light source, as shown in the Figure 4.1. The light beam diffracted by this wavelet is defined by its local position of emittance $\mathbf{X}$, and its direction of diffraction given by the spherical angles $(\theta, \varphi)$. The local directional information of the diffracted light can be accurately obtained from the TF parameters of the Gabor wavelet $g_{(s_0,\theta_0,\boldsymbol{t_0})}$ given the

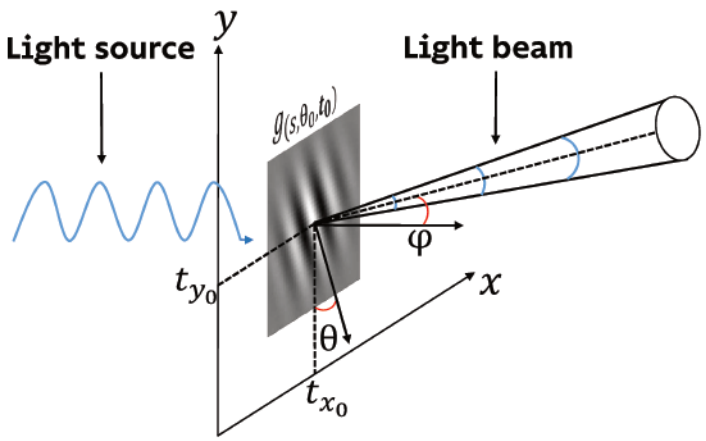



Fig. 4.1 Light diffraction by Gabor wavelets.

following correspondance

$$\begin{cases} \boldsymbol{X} = \boldsymbol{t_0} \\ \theta = \theta_0 \\ \varphi = \arcsin\left(\dfrac{\lambda\sqrt{f_{x_0}^2 + f_{y_0}^2}}{s_0}\right), \end{cases} \tag{4.8}$$

where $\lambda$ represents the wavelength of the light source.

For a given continuous holographic signal $h : \mathbb{R}^2 \to \mathbb{C}$, a Gabor TF decomposition is obtained by computing the set of convolutions between the hologram $h$ and a family of 2D continuous Gabor wavelets, such that

$$W_h(s, \theta, \boldsymbol{t}) = \iint_{\mathbb{R}} h(x, y) g_{(s,\theta,\boldsymbol{t})}(x, y) dx dy, \tag{4.9}$$

where $W_h(s, \theta, \boldsymbol{t})$ represents the transform hologram corresponding to the wavelet $g_{(s,\theta,\boldsymbol{t})}$.

Inversely, the original hologram can be recovered from $W_h$ by applying the 2D inverse continuous wavelet transform, such that

$$h(x, y) = \int_{\mathbb{R}^+} \int_0^{2\pi} \int_{\mathbb{R}^2} W_h(s, \theta, \boldsymbol{t})\ g_{(s,\theta,\boldsymbol{t})}(x, y) ds d\theta d\boldsymbol{t}. \tag{4.10}$$

In practice, digital holograms are discretized during the recording/generation process. We denote by $\mathbf{H}\{H^{(j)}\}_{j\in\{r,g,b\}}$ the color digital hologram in its discrete form, where $H^{(j)} \in \mathcal{M}_{N_x,N_y}(\mathbb{C})$ is the complex hologram corresponding to channel $j$. $\mathcal{M}_{N_x,N_y}(\mathbb{C})$ denotes

the set of $N_x$ by $N_y$ complex matrices, and thus $(N_x, N_y)$ is the spatial resolution of this hologram.

As emphasized in Eq. (4.8), the local directions of diffracted light are directly related to the intrinsic parameters (frequency, orientation and shift) of Gabor wavelets. Consequently, the hologram decomposition into a set of light beams will be highly sensitive to the choice of the Gabor wavelets family. In discrete cases, for an accurate TF hologram analysis, it is important to discretize the wavelets' parameters such that the whole hologram's spectrum (*i.e.* diffractive information) is covered. A suitable discretization of Gabor wavelets' parameters is given in the following.

In order to fit the possible directions of light diffraction, we set the frequency $f_0$ of the Gabor mother wavelet to the maximal hologram frequency $f_{max} = (2\Delta)^{-1}$, where $\Delta$ is its pixel pitch. The spatial shifting parameter $\mathbf{t}$ is sampled in the range $[1, N_x] \times [1, N_y]$, the frequency dilation parameter $s_l$ is dyadically discretized, whereas the rotation parameter $r_p$ is linearly discretized, such that

$$\begin{cases} s_l = 2^{\alpha[l]} \\ r_p = \dfrac{2p}{P}, \end{cases} \tag{4.11}$$

where $\alpha$ is an arbitrary linear function.

Aiming for a TF representation of digital holograms, a family of $S = L \times P$ discrete Gabor wavelets is built as defined in Eq. (4.5). The wavelets' parameters are discretized referring to Eqs. (4.6), (4.7) and (4.11). Then, the inner products between the hologram $\mathbf{H}$ and each wavelet of the Gabor family $F$ are computed. Finally, an overcomplete N-expansion $A_N$ ($N = N_x \times N_y \times S$ ) is obtained for each color channel, such that

$$A_N = \{a_n = (\gamma_n, c_n), \qquad 1 \leq n \leq N\}, \tag{4.12}$$

where $a_n$ is the $n^{th}$ Gabor atom, $c_n$ is the $n^{th}$ complex coefficient (*i.e.* inner product), and $\gamma_n$ is the $n^{th}$ atom's index defined by

$$\gamma_n = (l_n, p_n, \boldsymbol{t}_n). \tag{4.13}$$

In the following, we denotes by $\Gamma$ the set of atoms' indexes, and by $D$ the Gabor dictionary $\{g_\gamma\}_{g \in \Gamma}$.

Although the Gabor expansion permits an efficient analysis of hologram patterns, it may not lead to optimal results in terms of compression, since it provides an overcomplete representation. The next section explains how the Gabor dictionary's redundancy may be reduced in order to achieve a sparse representation of digital holograms.

## 4.2 Sparse Gabor expansion using Matching Pursuit

For data compression applications, it is important to find the optimal trade-off between reconstruction accuracy and sparsity of the signal representation. Indeed, the reconstruction error should be minimized using as few waveforms as possible from the signal expansion. To mathematically formulate the sparse decomposition problem, let us consider a signal $f$ from a Hilbert space $\mathcal{H}$, which can be represented as a finite linear combination of vectors (*i.e.* atoms) taken from a redundant dictionary $\{g_\gamma\}_{\gamma\in\Gamma}$. We want to find an index subset $\Gamma_0 \subset \Gamma$, which provides the sparsest representation of $f$ and minimal approximation error. To this end, we solve the following constrained minimization

$$\min_{\Gamma_0}(card(\Gamma_0)) \quad subject\ to \quad ||f - \sum_{\gamma\in\Gamma_0} c_\gamma g_\gamma||_2^2 \leq \epsilon^2, \tag{4.14}$$

where $card(.)$ represents the cardinality operator, $c_\gamma$ is the inner-product of $f$ and $g_\gamma$, and $\epsilon$ is a threshold representing the tolerated error.

If the dictionary's waveforms form an orthonormal basis, the solution of the minimization problem in Eq. (4.14) is the index subset $\Gamma_0$ corresponding to atoms with the largest inner products $\{\langle f, g_\gamma\rangle\}_{\gamma\in\Gamma}$.

On the other hand, when the dictionary's waveforms are linearly independent, finding an exact sparse decomposition of $f$ is considered as a NP-hard problem [82], which means that it has intractable complexity [80]. Therefore, sub-optimal approximation algorithms are used as an alternative for sparse representation [60].

### 4.2.1 Scalar and vector matching pursuit algorithms

Matching Pursuit (MP) is a well-known greedy algorithm applied to sparse signal approximation from overcomplete dictionaries [138]. It finds its applications in several domains such as image denoising [212], video coding [160], compressed sensing [146]. MP is an iterative algorithm that allows linear expansion by greedily selecting, from a redundant dictionary $D$, the atoms which are the most correlated with the target signal $f$. This is achieved by searching in $D$ the atom providing the maximal inner product with $f$, computing the residual $R(f)$ by subtracting from the signal an approximation that uses only that one atom. This process is repeated until the final residual's norm is below a threshold, or the wanted number of atoms are extracted. It must be noted that a wise choice of the dictionary waveforms may improve the reconstruction accuracy, as well as the convergence speed. In the following, the TF Gabor dictionary is considered.

**Algorithm 1** SMP pseudo-code.

**Require:** $\mathbf{H}$, $\boldsymbol{A}_N$
**Ensure:** $\boldsymbol{A}_K$
1: **for** $j \in \{r, g, b\}$ **do**
2: $R_0\left(H^{(j)}\right) = H^{(j)}$ ▷ Residue initialization for each channel
3: **for** $k \in \{1, \ldots, K\}$ **do**
4: $\gamma_k^{(j)} = \underset{\gamma^{(j)} \in \Gamma^{(j)}}{\arg\max} \left|\left\langle R_k\left(H^{(j)}\right), g_{\gamma^{(j)}}\right\rangle\right|$ ▷ Find $\gamma_k^{(j)}$ that maximizes inner product
5: $c_k^{(j)} = \left\langle R_k\left(H^{(j)}\right), g_{\gamma_k^{(j)}}\right\rangle$ ▷ Compute the $k^{th}$ inner product
6: $R_{k+1}\left(H^{(j)}\right) = R_k\left(H^{(j)}\right) - c_k^{(j)} g_{\gamma_k^{(j)}}$ ▷ Compute the $(k+1)^{th}$ residual
7: **end for**
8: **end for**

Let us consider a color digital hologram $\mathbf{H}$ and its overcomplete Gabor N-expansion $\boldsymbol{A}_N$, as defined in Section 4.1.3. The goal of MP is to extract an atom subset $\boldsymbol{A}_K$ from $\boldsymbol{A}_N$, where $K \ll N$, in order to achieve a sparse representation of $\mathbf{H}$. For multi-channel signals, MP algorithm has two versions: Scalar Matching Pursuit (SMP) and Vector Matching Pursuit (VMP). Their pseudo-codes are given in Algorithm 1 and Algorithm 2, respectively.

For SMP, the MP algorithm is performed on each color channel $j$ seperatly, which results in an atom set of $3K$ indexes and coefficients. On the other hand, for VMP, the atoms' indexes are chosen in a joint manner to match simultaneously the three RGB channels by maximising the $L_p$-norm of the inner products vector (*cf.* step 5 in Algorithm 2).

**Algorithm 2** VMP pseudo-code.

**Require:** $\mathbf{H}$, $\boldsymbol{A}_N$
**Ensure:** $\boldsymbol{A}_K$
1: **for** $j \in \{r, g, b\}$ **do**
2: $R_0\left(H^{(j)}\right) = H^{(j)}$ ▷ Residue initialization for each channel
3: **end for**
4: **for** $k \in \{1, \ldots, K\}$ **do**
5: $\gamma_k = \underset{\boldsymbol{\gamma} \in \boldsymbol{\Gamma}}{\arg\max} \|\langle R_k(\mathbf{H}), g_{\boldsymbol{\gamma}}\rangle\|_p$ ▷ Find $\gamma_k$ that maximizes the $L_p$-norm
6: **for** $j \in \{r, g, b\}$ **do**
7: $c_k^{(j)} = \langle R_k\left(H^{(j)}\right), g_{\gamma_k}\rangle$ ▷ Compute the $k^{th}$ inner product
8: $R_{k+1}\left(H^{(j)}\right) = R_k\left(H^{(j)}\right) - c_k^{(j)} g_{\gamma_k}$ ▷ Compute the $(k+1)^{th}$ residual
9: **end for**
10: **end for**

Consequently, VMP allows a hologram recovery from only $K$ atoms' indexes through the approximating linear combination

$$\hat{\boldsymbol{H}} = \sum_{k=1}^{K} \langle R_k(\boldsymbol{H}), g_{\gamma_k} \rangle g_{\gamma_k}. \tag{4.15}$$

VMP has been shown to be relatively efficient for color image compression, since it reduces the bit rate budget of atoms' indexes by a factor of 3, without visually degrading the quality of decoded images [113, 136]

### 4.2.2 Matching Pursuit acceleration

Despite its simple implementation, the major limitation of MP algorithm is its high computational burden. Indeed, the two main time-consuming operations, at each iteration, are: $(a)$ inner-products computation of signal residual with all dictionary atoms, $(b)$ maximum inner-product search. Moreover, the number of needed operations grows significantly when the signal's resolution and/or dictionary's size increases.

To reduce the computational complexity of SMP and VMP, we slightly modify the MP algorithm as reported in [13, 138]. In these works, the authors proposed to compute the inner-product of the $(k+1)^{th}$ residual with any $g_{\gamma'} \in D$, using a linear updating formula (*cf.* step 6 in Algorithm 1 and step 8 in Algorithm 2) as follows:

$$\langle R_{k+1}(\boldsymbol{H}), g_{\gamma'} \rangle = \langle R_k(\boldsymbol{H}), g_{\gamma'} \rangle - \langle R_k(\boldsymbol{H}), g_{\gamma_k} \rangle \langle g_{\gamma_k}, g_{\gamma'} \rangle. \tag{4.16}$$

According to Eq. (4.16), at a given iteration, the inner-products to be updated are only those for which the cross-correlations $\langle g_{\gamma_k}, g_{\gamma'} \rangle$ is non null. Since the support of discrete Gabor wavelets is compact, the set $\{\langle g_{\gamma}, g_{\gamma'} \rangle\}_{\gamma,\gamma' \in \Gamma}$ is extremely sparse. Thus, only a small fraction of the total number $M = M_x M_y S$ of coefficients is updated, thus dramatically reducing the computational cost of operation $(a)$.

For high-resolution holograms, a full search for the maximum at every iteration could make computation intractable. To overcome this hurdle, we propose a GPU-based implementation using parallel computation to reduce the execution time.

**Overview of the CUDA Thread Structure and Memory Model** We implemented the SMP and VMP in C++/CUDA on a PC equipped with an Intel Xeon e5-2620 v3 CPU operating at 2.40 GHz, a main memory of 32 GB, a GPU NVIDIA GeForce GTX 1070, under Microsoft Windows 10. In the implementation, all computations (inner products, maximum search and updating) are done by the GPU. We use the CPU threads to load

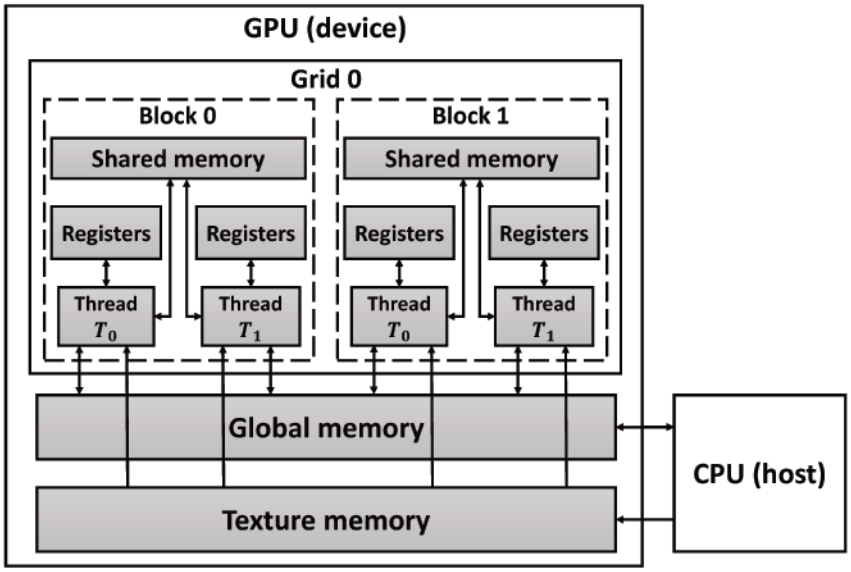



Fig. 4.2 Cuda thread structure and memory model.

hologram values, launch CUDA kernels and save the coefficients and atoms' indexes into an output file.

Fig. 4.2 illustrates the organization of threads and memory model in the device (GPU), as well as the memory transfer with the host (CPU). In CUDA, a kernel is the function to be executed with parallel computation. When invoking a kernel from the host, a grid of parallel threads which are organized in blocks is generated to simultaneously process and store data using the onchip registers. Due to its short-latency, the shared memory allows a fast access and exchange of data within a block but it suffers from a limited memory capacity. The global memory ensures transfer of large data between CPU and GPU due to its large capacity, but its long-latency implies a slow data transfer. Furthermore, the GPU threads can also access to a read-only texture memory, which is cached on-chip. Since texture cache is optimized for 2D spatial locality, threads that read adjacent texture addresses will achieve best performance.

**MP GPU-implementation** The software structure of our fast MP expansion is given in Fig. 4.3. The implementation is divided into two main parts:

- **Inner products pre-calculation:** First, the CPU loads the complex values of the color hologram $\mathbf{H}$ and the dictionary atoms $g_\gamma$ into the GPU global memory. Then, the coefficients $\{\boldsymbol{c}_n\}$ and cross-correlations $\{\boldsymbol{d}_m\}$ are computed offline and stored in the GPU global memory. For a fast calculation, we compute the inner products as a convolution product using the CUDA cuFFT library of NVIDIA [49]. Kernel *Kernel1*

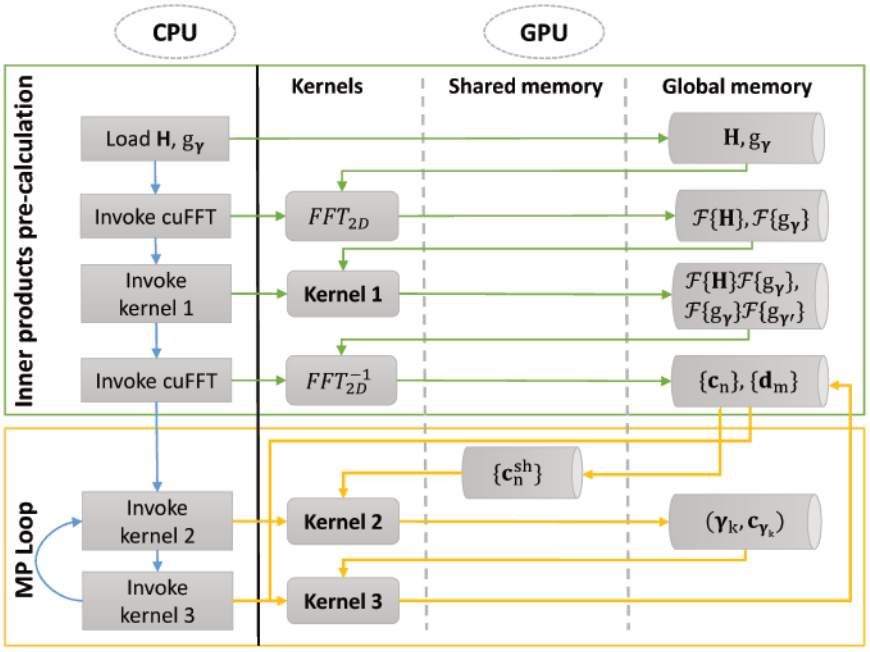



Fig. 4.3 Software structure for the Matching Pursuit algorithm.

launches a parallel multiplication in the Fourier domain with one thread per sample. This kernel is used twice, once to compute the product $\mathcal{F}\{\boldsymbol{H}\}\mathcal{F}\{g_\gamma\}$ and once for the product $\mathcal{F}\{g_\gamma\}\mathcal{F}\{g_{\gamma'}\}$.

- **MP loop:** Once all coefficients are stored in the global GPU memory, the MP expansion is obtained by iteratively selecting the atoms' indexes in the decreasing order of $\{\boldsymbol{c}_n\}$ magnitude. At every iteration $k$, the maximum search and the coefficients updating are performed by invoking kernels *Kernel2* and *Kernel3*, respectively.

  Kernel *Kernel2* returns the maximum $\boldsymbol{c}_{\gamma_k}$ of pre-calculated coefficients $\boldsymbol{c}_n$ and its corresponding atom's index $\boldsymbol{\gamma}_k$ using parallel comparison. Since every coefficient should be compared with all others, every thread of *Kernel2* must have access to set $\{\boldsymbol{c}_n\}$. To optimize data access and transfer, we start by loading all the coefficients into the shared memory which has shorter latency than the global memory. Then all the threads are synchronized to avoid memory load conflicts. Once this is done, *Kernel2* computes the maximum using a reduction algorithm [101].

  Kernel *Kernel3* launches a parallel updating of the coefficients by using the pre-calculated cross-correlations $\boldsymbol{d}_m$, according to (4.16).

| Resolution | Execution time (min) | | | | | |
|---|---|---|---|---|---|---|
| | SMP | | | VMP | | |
| | CPU | GPU | Speedup | CPU | GPU | Speedup |
| $\mathbf{1}\boldsymbol{K}$ | 93 | 12 | 7.8 | 87 | 10 | 8.7 |
| $\mathbf{2}\boldsymbol{K}$ | 379 | 45 | 8.4 | 352 | 38 | 9.3 |
| $\mathbf{4}\boldsymbol{K}$ | 1512 | 187 | 8 | 1438 | 155 | 9.3 |

Table 4.1 Comparison of SMP and VMP execution using CPU and GPU versions.

To compare our CUDA implementation of SMP/VMP with a CPU version implemented in C++, we run a K-expansion of color CGHs with different resolutions, where $K = \frac{N_x \times N_y}{2}$. The acceleration results obtained by our GPU version are summarized in Table 4.1.

We note that an acceleration of more than $8$ times is achieved by SMP and VMP when using the proposed GPU implementation. Moreover, VMP is slightly faster than SMP, since the search for the highest correlation is jointly performed on all channels. Since the SMP/VMP expansion is computed offline on the server side, the effective execution time of our codec will be only related to the actual encoding time of the selected number of MP atoms. As it will be shown in the next section, this step is very similar to conventional image coding algorithms such as JPEG2000, since it is also based on quantization and entropy coding. Therefore, as for the encoding time, we expect the same order of magnitude as these classical algorithms.

## 4.3 Design choices of the encoding framework

For transmission purposes, a compressed bitstream should be generated from the MP expansion, which requires a study on the quantization and encoding choices. This compressed bitstream is made up of the quantized MP coefficients and the corresponding indexes.

In this section, we study the statistical properties, the encoding order and format of MP components (*i.e.* the coefficients and the indexes) of the complex holographic signal. We propose a framework for representing this information with the aim of optimizing the trade-off between the total coding rate and the distortion on the *reconstructed image*. Note that this is a difficult task, since the relationship between the reconstructed image and the reconstructed hologram $\hat{\mathbf{H}}$ (*cf.* (4.15)) is not straightforward.

For an efficient compression, it is essential to estimate the probability distribution of MP components, and then to design the entropy encoder accordingly. For each Gabor atom $a_k$ of the MP approximation, we need to encode four components, as listed below

- **Dilation index $l_k$:** integer in the range $[1, S]$ corresponding to the $k^{th}$ atom of MP expansion. In our experiments $S = 8$.

- **Rotation index $p_k$:** integer in the range $[1, P]$ corresponding to the $k^{th}$ atom of MP expansion. In our experiments $P = 8$.

- **Atom position $\mathbf{t}_k = (t_{x_k}, t_{y_k})$** the spatial 2D position of the $k^{th}$ atom, where $1 \leq t_{x_k} \leq N_x$ and $1 \leq t_{y_k} \leq N_y$.

- **Atom coefficient $c_k$:** the complex inner-product value of the $k^{th}$ selected atom.

Therefore, the bitstream $\mathcal{B}$ that encodes the MP K-expansion of the hologram is given by

$$\mathcal{B} = \{a_k = (l_k, p_k, \boldsymbol{t}_k, [c_k]); \qquad 1 \leq k \leq K\}, \tag{4.17}$$

where $[c_k]$ is the quantized value of $c_k$.

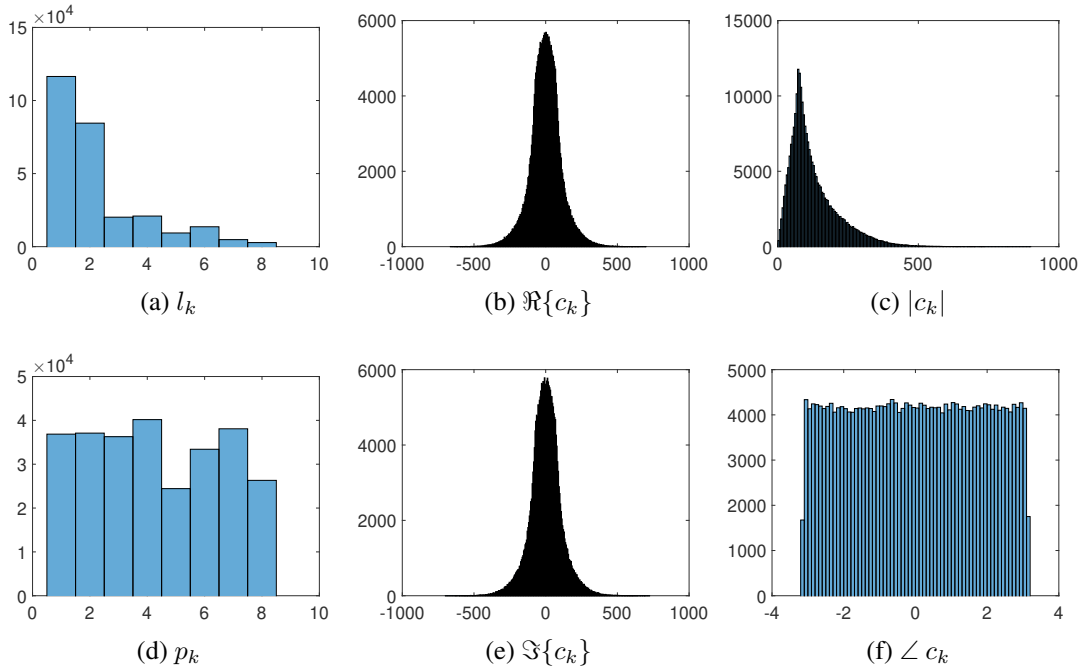
Several choices must be made for these parameters. First and foremost is the atom ordering; a possible solution would be to encode the atoms in decreasing norm order. This would induce on the encoded bitstream a certain degree of progressivity, meaning that the first bits of the stream conveys the most of the information (*i.e.* the amplitude of the most energetic atoms). However, some preliminary experiment shows that in this case the encoding of the atom position $\mathbf{t}_k$ is very costly, approaching the maximal rate $\log_2(N_x) + \log_2(N_y)$. This means that, only for encoding the position, we need $20$ bits per atom for a 1K-resolution hologram, or 24 bits per atom in the case of 4K. A more efficient solution is to encode the atoms in raster scan order. As we show later on, we can reduce the cost of the position to a shear $3$ bits per atom for 1K holograms. In the following, we only consider the case of raster-scan ordering of atoms.

### 4.3.1 Distribution of the atoms' components

In Figure 4.4, we provide the histograms of the MP components for a 1K CGH. For all our tests, we set a Gabor family with $64$ wavelets where $L = P = 8$. The arbitray function $\alpha$ defining the frequency dilation (*cf.* Eq. (4.11)) is given by

$$\alpha(l) = 0.4(l - 1). \tag{4.18}$$

We computed the same histograms for several other CGHs, and all share the same main characteristics allowing us to draw relevant conclusions.




(a) $l_k$ (b) $\Re\{c_k\}$ (c) $|c_k|$

(d) $p_k$ (e) $\Im\{c_k\}$ (f) $\angle\, c_k$

Fig. 4.4 Histograms of the MP components.

First, the dilation coefficient distribution (*cf.* Figure 4.4a) tells us that MP mainly selects high-frequency atoms. Since the distribution is rather non-uniform, entropy coding can achieve some rate reduction. On the contrary, the rotation index distribution is almost uniform, as shown in Figure 4.4d, which confirms that hologram patterns have no specific contours and no orientation is privileged for the diffracted light beams. Therefore, no gain is expected from entropy coding of $p_k$. Considering the histograms of the R-I and A-P parts of the MP coefficients $c_k$, we observe that, as expected, the phase histogram is very flat, while the others are more peaky. However, in order to decide the most appropriate representation, we would rather consider the entropy and the conditional entropy of the MP components, since these represent a good estimation of the coding rate that can be achieved respectively by an entropy encoder and a context-based entropy encoder. This is detailed in what follows.

### 4.3.2 Entropies of the atoms' components

For the same set of holograms as in the previous subsection and for each MP component, we estimate the first-order entropy $H(X)$, and the conditional entropies $H(X|Y)$ and $H(X|Y,Z)$, where $X$ is the MP component (*e.g.*, the dilation index of a selected atom), $Y$ is its immediate neighbour (in the example,the dilation index of the previously selected atom) and $Z$ the previous neighbour. This is shown in Table 4.2. The corresponding entropies are

| **Component** | $H(X)$ | $H(X\|Y)$ | $H(X\|Y,Z)$ |
|---|---|---|---|
| $l_k$ | 2.27 | 2.27 | 2.26 |
| $p_k$ | 2.98 | 2.98 | 2.97 |
| $t_k$ | 10.04 | 10.03 | 10 |
| $\delta_k$ | 3.28 | 3.17 | 2.98 |
| $[\|c_k\|]$ | 4.85 | 3.11 | 1.98 |
| $[\angle\, c_k]$ | 7.29 | 6.29 | 6.02 |
| $[\Re c_k]$ | 5.01 | 4.10 | 1.94 |
| $[\Im c_k]$ | 5.02 | 4.10 | 1.94 |

Table 4.2 Conditional entropies of MP components.

good estimations of the lower coding-rate bound of entropy encoders which respectively use no context, a one-neighbour context and two-neighbour context.

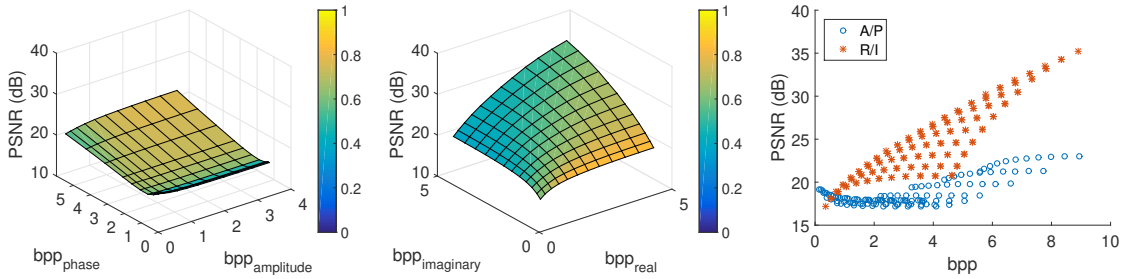
The entropies of $l_k$ tell us that it is effective to use an entropy coder for this component, since the coding rate is reduced from $\log_2(S) = 3$ to $H(l_k) = 2.27$, *i.e.* a $24\ \%$ rate reduction. However, we observe that the $l_k$'s are poorly correlated, since the conditional entropy is very close to the first order entropy. Therefore, it is not worth using context-based coding on this component.

As for the rotation parameter, we see that $H(p_k)$ is very close to $\log_2(P) = 3$, which means that a variable-length coder would most probably achieve no gain. Moreover, the $p_k$'s are also poorly correlated among them. In conclusion, the rotation indexes shall be coded with a simple 3-bit fixed-length code.

The next parameter is $\mathbf{t}_k$. We see that its entropy is around 10 bits, which is close to the half of the rate needed to encode $t_{x_k}$ and $t_{y_k}$ (since $N_x = N_y = 1024$, we would need $\log_2 N_x + \log_2 N_y = 20$ bits). No correlation is shown among the position indexes, since the conditional entropy gives no improvement. However, we wonder whether there is not a smarter way the encode the atoms' positions. We consider $\delta_k$ which is the difference between two consecutive atom positions in raster scan order, defined by

$$\delta_k = [t_{x_k} + (N_x - 1) \cdot t_{y_k}] - [t_{x_{k-1}} + (N_x - 1) \cdot t_{y_{k-1}}]. \tag{4.19}$$

We find out that $\delta_k$ is much more convenient to encode than the absolute position. Furthermore, some correlation exists between neighbouring values of $\delta_k$, justifying the resort to a context-based entropy coder for this component. Globally we can reduce the position information from $20$ bits to about $3$ bits, with a rate reduction of $85\ \%$.




(a) Amplitude-Phase

(b) Real-Imaginary

(c) Amplitude-Phase vs Real-Imaginary. For each point the rate is the sum of the two component rates.

Fig. 4.5 HEVC rate-distortion curves for the Amplitude-Phase and Real-Imaginary formats.

### 4.3.3 Coding of the atoms' coefficients

The encoding of MP coefficients consists into two steps. First they are quantized using a suitable quantization step (or number of bits); then we use lossless coding on the resulting quantized coefficients. Moreover, we have the choice between two possible representations: R-I and A-P. In order to address this first issue, we conducted a simple yet informative experiment. We encoded the R-I and A-P data as they were ordinary images using HEVC in intra-coding mode. We registered the coding rates for each component and as a quality metric, we computed the PSNR of the reconstructed images with respect to the image reconstructed from the original hologram[1]. The results are shown in Figure 4.5. More precisely, in Figures. 4.5a and 4.5b we show the separate rate contributions of the two components, and the resulting PSNRs. In Figure 4.5c, we compute the global rate as the sum of the two component rates. The result is very clear: for practically all the bit-rates, the A-P representation is largely less effective than the R-I one. Moreover, we found a very simple and nice result, also visible from Figure 4.5b: for each total rate $R_I + R_R = R_T$, the largest PSNR is obtained for $R_R = R_I = R_T/2$. This is a very simple and effective practical rule for deciding the rate allocation between the two components. We also observe that in the case of A-P representation, there is no such a simple rule, which makes even less appealing its use. Since similar results have been obtained for several other histograms, we conclude that the best choice is to use the R-I representation with equal rate allocation between the real and imaginary parts.

[1]The PSNR – Peak Signal-to-Noise Ratio – of an image $f$ with respect to a reference image $g$ is a logarithmic measure of the mean square error: $\text{PSNR} = 10\log_{10}\frac{(2^b-1)^2}{\text{MSE}}$, with $b$ being the bit depth of the images (typically, $b = 8$ bits) and $\text{MSE} = \frac{1}{N_x N_y}\sum_{x,y}[f(x,y) - g(x,y)]^2$.

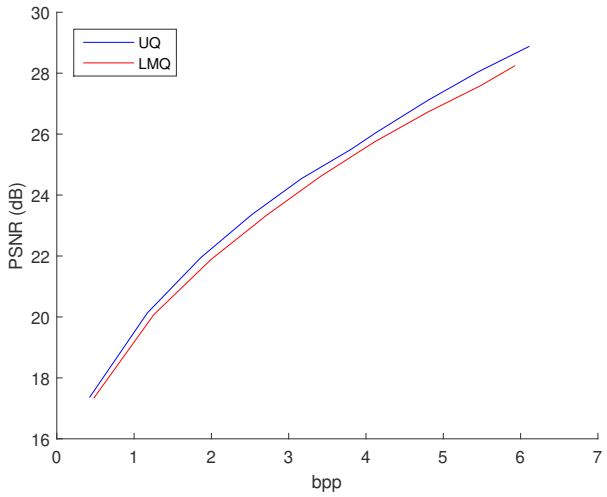



Fig. 4.6 RD comparison between LMQ and UQ.

We also have to consider the relationship among the quantizer resolution $R_Q$ (number of bits per coefficient), the number of atoms $K$ and the total available bit-rate. For the sake of simplicity, we consider only a small number $N_R$ of possible quantizer resolutions, and for each of them we can plot a RD curve by adjusting the number of selected atoms $K$. The global RD curve of our system will be the convex hull of these $N_R$ curves. We mention explicitly that, in preliminary tests, we observed that a large value of $N_R$ does not improve significantly the performance, while a small value of $N_R$ keeps the complexity in a reasonable range. Therefore in our system we typically have $N_R = 2$ or $3$.

The next step is to design the quantizer of the MP coefficients. We consider two possible choices: a Lloyd Max Quantizer (LMQ) and an Uniform Quantizer (UQ), since the first is theoretically optimal while the second is very simple. In order to select between LMQ and UQ, we perform an encoding of several holograms using our complete system, using in a first case the LMQ, and in a second case the UQ. Even though the LMQ assures the lowest distortion on the decoded hologram, this does not necessarily means that the *reconstructed image* will be better than the one obtained by UQ, since the reconstruction process is not an isometric operator. Indeed, as shown in Figure 4.6, UQ may have better performance than LMQ. More in general, on several holograms we observed close performances between LMQ and UQ, so we selected the latter in sight of its simplicity.

Finally, we consider the entropy of the quantized MP coefficients. We consider a 8-bit uniform quantization per component, which can be seen as a high-quality representation of the signal. In this case, the A-P representation needs $11.3$ bits per atom, a value that can be reduced to $\approx 8$ bits when using two neighbours as context. However, the R-I representation is

more efficient, needing as low as ≈4 bits per atom when a context-based encoder is used. Moreover we observe a trend in the conditional entropy of MP coefficients which would suggest to look to the case of larger neighbourhoods. However, using larger neighbourhoods makes the estimation of the entropy less and less reliable [33], without mentioning the exponentially-increased complexity. Therefore we cannot go behind a size of two neighbours.

Finally, the main design choices for our encoder are summarized here:

**Coding order** is by raster-scan of atoms.

**Dilation indexes** are encoded with an arithmetic encoder without context.

**Rotation indexes** are encoded with a fixed length code using $\log_2(P)$ bits.

**Atom's position** is represented by the differential position $\delta_k$ which is encoded using a context-based arithmetic encoder with a two-neighbours context

**MP coefficients** are quantized using a UQ with a number of bits which depends on the target bit rate. Then, the quantized coefficients are represented as real and imaginary parts, and each of them is encoded using a context-based arithmetic encoder with a two-neighbours context.

We explicitly note that this coding framework is easily adapted (with minor changes) to both SMP and VMP cases. Moreover, for the VMP case, we may change the norm used in Algorithm 2. This fine-tuning of the proposed method is detailed in the following section.

## 4.4 Compression performance

In this section we evaluate the effectiveness of the proposed method (both SMP and VMP versions) to compress color digital holograms. As stated before, the target of hologram compression is not to preserve the complex hologram signal, but rather the quality of the reconstructed image. Therefore, in all our experimental tests, we use as quality metric the PSNR between two images: the one reconstructed from the original (ie. non-compressed) hologram, and the one reconstructed from the compressed one. The cost metric is the total rate needed to encode the holographic signal.

First of all, we introduce the set of CGHs used in our tests, as well as the fine-tuning of the proposed encoder. Then, we present the compression encoders used for benchmarking. Finally, the rate-distortion graphs and some sample reconstructed images are presented and discussed.

| **Parameter** | **$\Delta$ size** | |
|---|---|---|
| | **Large** | **Small** |
| **Hologram's name** | *Dices2K*, *Car1k* | *Ring4K*, *DiffuseCar4K* |
| **Resolution** | 2K, 1K | 4K |
| **Pixel pitch** ($\mu$m) | 4.8 | 0.4 |
| **RGB-Wavelength** (nm) | (640, 532, 470) | (640, 532, 470) |
| **RGB-Diffraction** ($^\circ$) | (3.8, 3.2, 2.8) | (53.1, 41.7, 36) |

Table 4.3 Color CGH database parameters.

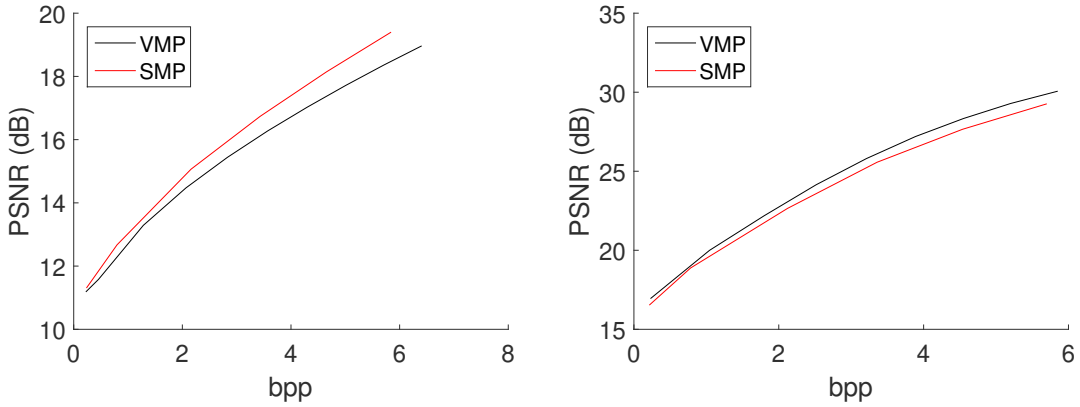
### 4.4.1 Holographic database

Table 4.3 summarizes the parameters of our test color CGHs, which can be classified into two categories according to their pixel pitches ($\Delta$). The chrominance format of CGH is set to 4:4:4 with 8-bit depth for each channel.

As referred in Table 4.3, since their maximal diffraction angle is bigger, color CGHs with small pixel pitch ($\Delta \approx \lambda$) provide a wider field of view than the ones with large pixel pitch. They also contain more high frequency components. Moreover, a high resolution is needed for CGHs with small pixel pitch in order to obtain a good quality (*i.e.* with less speckle noise) reconstruction of the object wave.

### 4.4.2 Comparison between the scalar and vector encoders

As discussed in Section 4.2, the efficiency of SMP and VMP is related to the correlation among color channels. In order to give some enlightenment about this issue, we show in Figure 4.7 the RD performance corresponding to the two variants of the proposed method using respectively SMP and VMP. For the *Dices2K* hologram, the scene is composed of one red, one blue and one green dice, each having different depth. In that case, the three color channels are weakly correlated in the object domain. Consequently, a disjoint decomposition of the hologram using SMP is expected to have better performance. This is confirmed by the RD graph shown in Figure 4.7a. However, in most cases, there is a strong correlation between the RGB channels of the source, and VMP is then privileged. The *Ring4K* hologram is a good example since it is a mix of RGB colors. As expected, the RD graph in Figure 4.7b shows better performance for the VMP version in this case.




(a) *Dices2K* CGH (b) *Ring4K* CGH

Fig. 4.7 Comparison of SMP and VMP RD curves.

### 4.4.3 Selection of the $L_p$-norm for the vector encoder

According to Algorithm 2, the selection criterion of the best atom is based on the $L_p$-norm of the scalar products vector between the atom and the given color channels. Using our database CGHs, we compared the effectiveness of $L_1$, $L_2$ and $L_\infty$ norms in terms of RD performance. We found that $L_2$ and $L_\infty$ have roughly the same RD performance, while $L_1$ performs slightly worse. Given its smaller computational complexity we chose the $L_2$ norm.

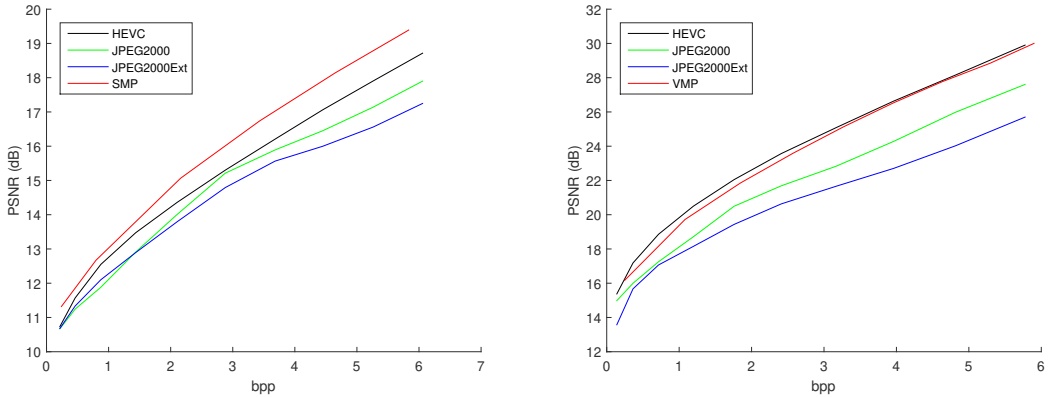
### 4.4.4 Evaluation of the Rate-Distortion performance

To evaluate our approach, we compare the SMP/VMP performance with three of the state-of-the-art codecs used for hologram compression. A brief description of the three codecs settings is given below:

**JPEG2000:** Each hologram's channel undergoes four levels of Mallat dyadic decomposition using the biorthogonal CDF9/7 filterbank. The codeblock's size of the used JPEG2000 encoder is set to $64 \times 64$ pixels.

**JPEG2000Ext [23]:** The extension is based on efficient compaction of the hologram's energy by using a 4-level full packet transform. Indeed, it allows to decompose high-pass sub-bands that convey an important part of the holographm energy. Moreover, the directional adaptive wavelet transform may be effective in case of holograms where more directional features appear.

**HEVC:** The codec is used in intra-coding mode for the real and imaginary part separately. The bit rate is controlled by varying the Quantization Parameter (QP) values.




(a) *Dices2K* (b) *Car1K*

Fig. 4.8 RD curves for holograms with large pixel pitches.

• **Holograms with large pixel pitches** As visible from Table 4.3, holograms with large pixel pitches have a small diffraction angle. Thus, we consider in our tests a central view reconstructed at a distance $z_{recons_1}$ = 0.543cm and $z_{recons_2}$ = 0.355cm, which correspond to an observer focusing on the red dice and the front of the car holograms, respectively.

**a) RD graphs:** The RD graphs of the *Dices2K* and *Car1K* holograms are given in Figures 4.8a and 4.8b, respectively. Our proposed method (with $R_Q \in \{5, 6\}$ bits), clearly outperforms JPEG2000 and JPEG2000Ext in both cases. For the *Dices2K* hologram, the proposed method also outperforms HEVC, whereas for the *Car1K* one, the performance is quiet similar.

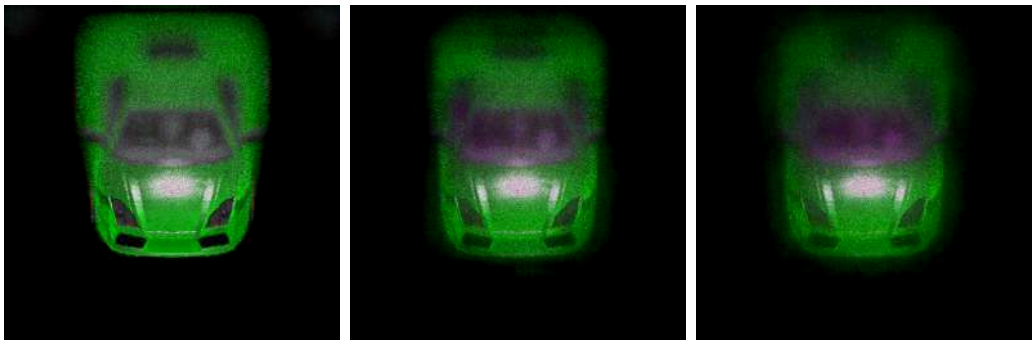
**b) Visual results:** The bit rates considered for the compressed *Dices2K* and *Car1K* holograms are $bpp_1$ = 2.15 and $bpp_2$ = 1.75, respectively. As shown in Figures 4.9 and 4.10, we clearly see that HEVC and VMP/SMP provide a better visual quality compared to JPEG2000 and JPEG2000Ext, which confirm the RD performance. Indeed, the HEVC reconstruction of the red dice appears darker than the VMP/SMP. In addition, the contrast of the SMP reconstruction is better than VMP. Also, the front of the car shows more aliasing for the HEVC reconstruction. Overall, VMP and SMP seem having a better visual quality than HEVC.

• **Holograms with small pixel pitches** Having a small pixel pitch, the *Ring4K* and *Diffuse-Car4K* holograms exhibit larger diffraction angles. Thus, they allow a reconstruction from different points of views. In this following tests, we consider two reconstructed viewpoints:

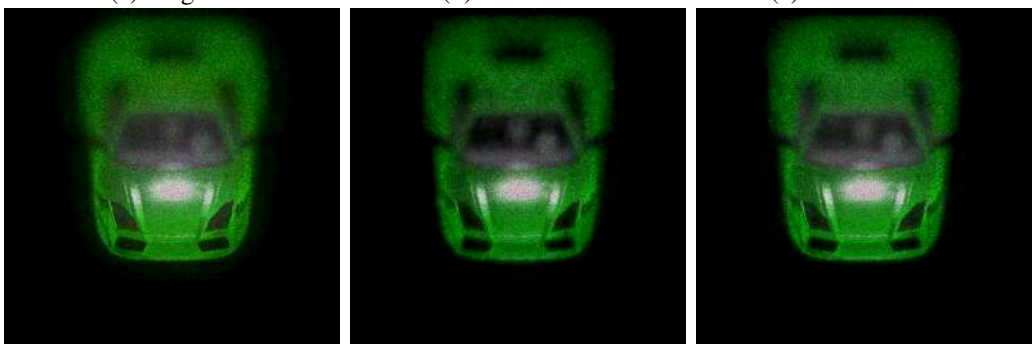
(a) Original (b) JPEG2000 (c) JPEG2000Ext

(d) HEVC (e) SMP (f) VMP

Fig. 4.9 Numerical reconstructions of *Dices2K*: (a) original, (b-f) compressed with 2.15 bpp.

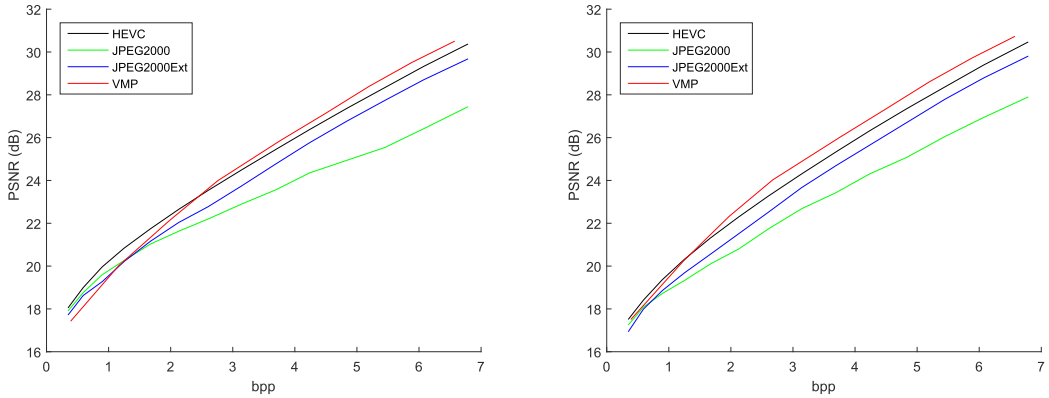


(a) Original (b) JPEG2000 (c) JPEG2000Ext

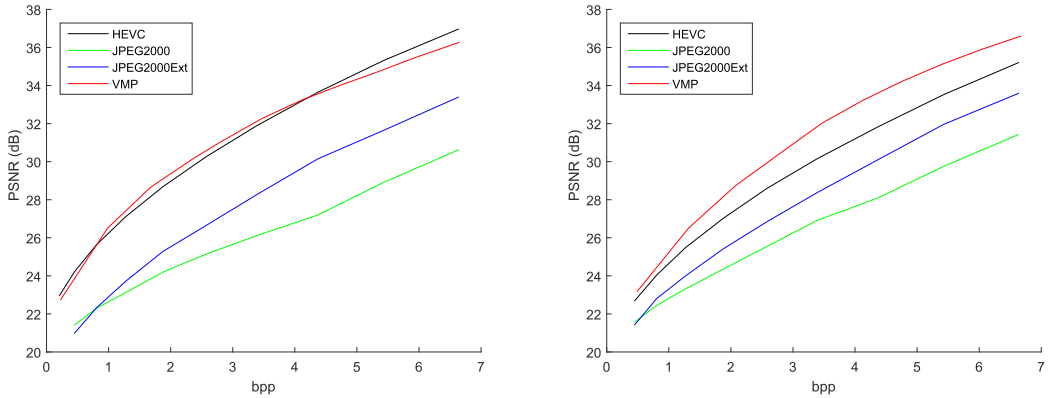


(d) HEVC (e) SMP (f) VMP

Fig. 4.10 Numerical reconstructions of *Car1K*: (a) original, (b-f) compressed with 1.75 bpp.




(a) Central view (b) Right view




(c) Central view (d) Right view

Fig. 4.11 RD curves for holograms with small pixel pitches: (a-b) *Ring4K*, (c-d) *DiffuseCar4K*.

one from a central position and a second from the right side. The distances are $z_3 = 0.16$cm and $z_4 = 0.11$cm, which correspond to an observer focusing at the front part of the ring and car, respectively.

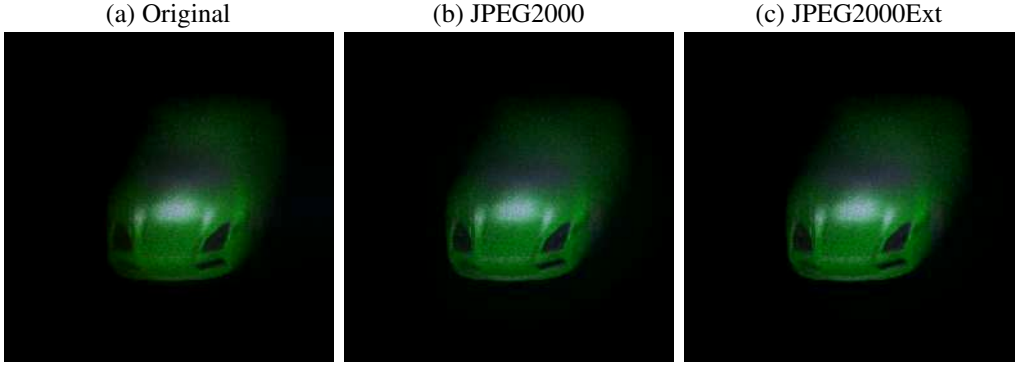
**a) RD graphs:** Figure 4.11 shows the RD graphs obtained for the *Ring4K* and *DiffuseCar4K* holograms, respectively. VMP, with $R_Q \in \{4, 5, 6\}$ bits, clearly outperforms HEVC, JPEG2000 and JPEG2000Ext in both cases for the right view. For the central view of *DiffuseCar4K*, VMP provides better RD trade-off compared to JPEG2000 and JPEG2000Ext, and slightly outperforms HEVC at medium bitrates. For the central perspective of *Ring4K*, VMP has a better compression performance than other approaches excepted at low bitrates for HEVC, and very low bitrates for JPEG2000 and JPEG2000Ext.

(a) Original (b) JPEG2000 (c) JPEG2000Ext

(d) HEVC (e) SMP (f) VMP

Fig. 4.12 Numerical reconstructions of *Ring4K*: (a) original, (b-f) compressed with 2.65 bpp.

(a) Original (b) JPEG2000 (c) JPEG2000Ext



(d) HEVC (e) SMP (f) VMP

Fig. 4.13 Numerical reconstructions of *DiffuseCar4K*: (a) original, (b-f) compressed with 1.95 bpp.

**b) Visual results:** The reconstructed images from the *Ring4K* and *DiffuseCar4K* holograms are given in Figures 4.12 and 4.13, respectively. The bit rates of the compressed holograms are $bpp_{\mathrm{Ring}} = 2.65$ and $bpp_{\mathrm{Diffuse\ Car}} = 1.95$. In both cases we show the reconstruction from the right-side point of view. The image reconstructed with VMP shows a better visual quality for both holograms, especially in the front left areas. Indeed, the retrieved wavefield of this region corresponds to the light beams having a higher angle of diffraction. Thus, higher frequencies are needed for reconstructing peripheral views, which may impair classical compression methods but not ours.

## 4.5 Conclusion

In this chapter, we proposed a novel approach for lossy compression of color digital holograms based on a sparse light beams decomposition using Scalar and Vector Matching Pursuit. Thanks to their good space-frequency localization, Gabor wavelets were used to model the light beams locally diffracted by holographic patterns. A detailed GPU implementation of the Matching Pursuit algorithm has been presented for the decomposition of high-resolution holograms. Moreover, the lossless encoding, quantization, order and format of the MP components have been studied and optimized in order to achieve the best RD performance for the proposed method.

Experimental results revealed that, in terms of RD performance, our technique outperforms state-of-the-art approaches in most cases. Moreover, a visual inspection of the reconstructed image quality confirm this findings. We show that Scalar Matching Pursuit is more efficient in case of holograms having a low correlation between their source channels, whereas Vector Matching Pursuit is to be privileged for holograms with strong correlation. Moreover, simulation results show that the proposed approach keeps consistent RD performance and visual quality independently of the observer's focus and point of view, in contrast to state-of-the-art methods, whose performance quickly deteriorate as the observer moves from the central position.

This characteristic of the proposed scheme is particularly interesting for scalable compression of digital holograms, and will be deeply studied in the next chapter.

# Chapter 5

# Viewpoint-quality scalable compression scheme of digital holograms

As stated in the previous chapter, digital hologram compression from a Gabor expansion using matching pursuit outperforms conventional image and video coding standards (JPEG 2000 and HEVC), especially for holograms with small pixel pitches. However, the bitstream generated by this method does not allow any form of scalability, which means that the whole bitstream must be decoded in order to visualize the holographic content. Unfortunately, this is incompatible with interactive holographic displays.

During hologram visualization, the observer is only watching at a specific direction given by its position and gaze. Thus, the viewpoint scalability is very suitable in the case of high-resolution holograms with wide FoV. Indeed, it enables the coding and transmission of a hologram subset based on the viewer's perspective, instead of the full holographic data. Furthermore, the information corresponding to a given viewpoint may be encoded into several quality layers, each allowing a distortion reduction of the reconstructed view. This is called quality scalability, which is a powerful tool for fine-grained progressive increase of reconstruction quality during visualization.

Accordingly, this chapter presents a compression scheme that combines viewpoint and quality scalability based on the viewing direction. Given the Gabor decomposition obtained previously, the viewpoint scalability is achieved by selecting the Gabor atoms that emit light beams into the observer's field of view. Then, the quality scalability is obtained by sorting these atoms with respect to their importance for the reconstructed perspective. In this chapter, we will show by experiments and simulations the effectiveness of this approach.

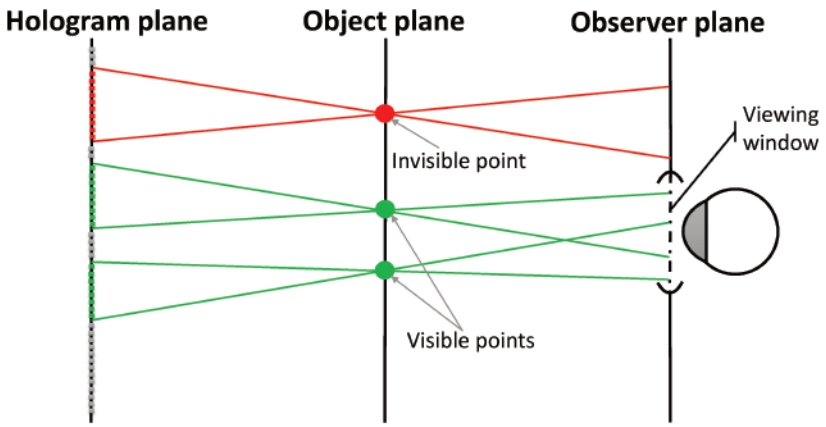



Fig. 5.1 Schematic principle of the holographic view-dependent representation. *The green points corresponds to visible points by the eye's pupil, whereas red points are invisible.*

## 5.1 Viewpoint scalable compression from Gabor expansion

Viewpoint scalability is very important for digital holographic data compression, especially for holograms with large FoV. Indeed, the hologram may gracefully degraded by encoding only the information corresponding to the light wave reached by a given localized observer instead of the full hologram. This functionality can be reached by expanding the original hologram according to a set of analyzing wavelets whose intrinsic properties match the way information is organized inside the patterns; hence this set of functions may not be optimal in terms of compression, *i.e.* it may not lead to the sparsest representation when considering the hologram in its entirety. However, to allow adaptive view-dependent reconstruction, it should provide a convenient way to generate a partial hologram corresponding to the viewer position. In the following, Section 5.1.1 emphasizes the view-dependent representation principle as well as the importance of TF localization for a concise extraction of directional information. Then, Section 5.1.2 provides the duality between Gabor wavelets and diffracted light beams. Finally, Section 5.1.3 explains how a viewpoint scalable coding scheme can be achieved by adaptively selecting relevant Gabor atoms from the MP expansion.

### 5.1.1 View-dependent representation of digital holograms

**Principle** When considering a hologram in its entirety, it is seen as an overlapping of diffraction gratings associated to various directions of light emission. However, when displaying it, only a subset of light beams reach the observer's eyes. As shown in Figure 5.1, for a given viewing window, the human eye's pupil would perceives only the light emanating

from visible points. The latter are reconstructed from specific regions of the hologram called *sub-holograms*.

Accordingly, a view-dependent representation can be achieved by separating sub-holograms corresponding to the observer's perspective.

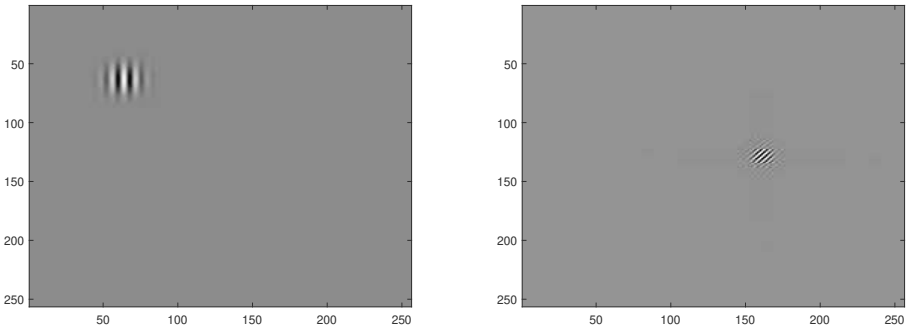
For such a representation, the reconstructed hologram only depends on the position and orientation of the observer; hence, from a representation point of view, it might be interesting to deal with data formats that are suited to real-time adaptation of a global holographic scene into a specific reconstruction corresponding to the viewing conditions at a given time. This would allow to significantly reduce the amount of information to encode, transmit, and display.

The seminal approach of view-dependent representation was proposed in [103, 104, 161]. In [161], a holographic display prototype based on the adaptation to the viewing window of an observer is presented. The minimal relevant features are gathered in sub-holograms. However, those are mostly pre-computed according to clustered directional features. In order to allow precise view-dependent selection of data to reconstruct into the displayed hologram, the representation should be based on a concise extraction of the directional information locally present in the hologram. In the next paragraph, we emphasize the importance of TF localization for an accurate extraction of sub-holograms.

**Importance of TF localization** According to the grating equation, the directions of light emission are related to spatial frequencies. However, because of the non-local character of holograms, each local region contains superimposed spectral components pertaining to light contributions scattered by several points of the 3D scene. To rigorously identify the directional information, this scrambled frequencies must be separated into multiple elementary components, each corresponding to a light beam emitted from a precise spatial location towards an exact direction.

Thus, it turns out that the extraction of sub-holograms corresponding to different observer's viewpoints requires an accurate localization in both spatial and frequency domains, simultaneously. Indeed, a better localization in space will lead to better precision in finding the appropriate spatial position in the hologram plane, and a better frequency localization will narrow down the angular uncertainty on the light beam direction.

From a wavelet point of view, a good TF localization implies that the wavelet spectrum must be concentrated around its proper frequency, while maintaining a good localization in the spatial domain. As discussed in Section 4.1.1, TF localization is controlled by the Heisenberg's uncertainty bound, which is minimized with Gaussian functions [77]. Since Gabor wavelets are Gaussian functions modulated by complex exponentials, they ideally




(a) Spatial domain (b) Frequency domain

Fig. 5.2 Time-frequency localization of a 2D Gabor wavelet (only real part is displayed).

provide the optimal localization trade-off in both domains. Figure 5.2 gives an example of Gabor wavelets TF localization, showing a compact support in both spatial and frequency domains.

### 5.1.2 Duality between Gabor wavelets and light beams

As shown in Figure 4.1, Gabor wavelets can model the light beams scattered by holographic patterns, leading to a duality between Gabor wavelets and diffracted light beams.

This duality is illustrated by Figure 5.3, showing the example of a light beam emitted by a Gabor wavelet $g_{(s,\theta,\boldsymbol{t})}$ from $\mathbf{t} = (t_x, t_y)$ in the Holograme Plane (HP). The direction of diffraction is given by ($\theta$, $\varphi$), where $\varphi$ is related to the frequency of the Gabor wavelet according to the grating equation (2.3).

In the Observer Plane (OP), the emitted light beam forms a diffraction spectrum $S$ centered around $C$ and having an angular dispersion $\Delta\varphi$. The latter depends on the frequency dilation parameter.

The coordinates of the diffraction spectrum's center $C(t'_x, t'_y)$ emitted by a Gabor atom $a$ from a position $(t_x, t_y)$ is given by the following relation

$$\begin{cases} t'_x = t_x + z_{obs} \tan(\varphi) \cos(\theta) \\ t'_y = t_y + z_{obs} \tan(\varphi) \sin(\theta), \end{cases} \tag{5.1}$$

where $z_{obs}$ is the distance between HP and OP.

To verify the aforementioned duality, let us consider a sub-hologram generated from one Gabor wavelet located at a position $\mathbf{t}_0(t_{x_0}, t_{y_0})$. By performing a forward propagation $\mathcal{P}^{AS}_{z_{obs}}$

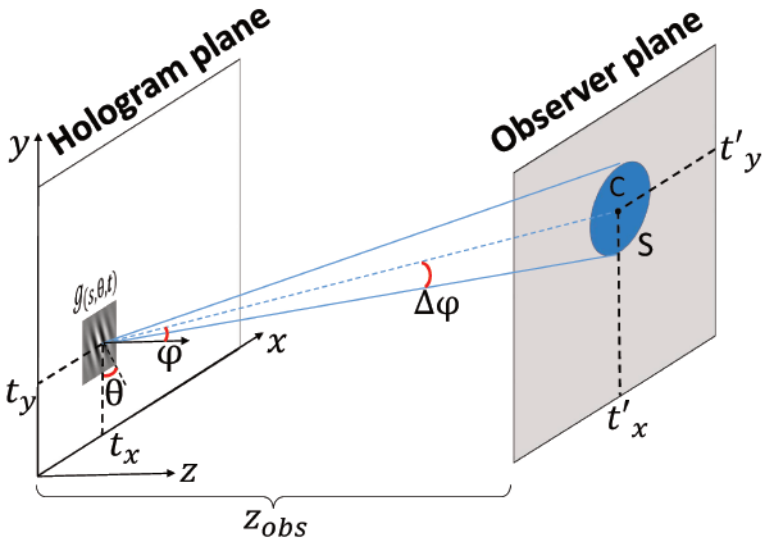



Fig. 5.3 Duality between Gabor wavelets and light beams. *A diffraction spectrum $S$ is formed by the light diffracted from the Gabor wavelet $g_{(s,\theta,t)}$.*

from the HP into OP. The coordinates of the obtained diffraction spectrum's center in the OP will exactly correspond to the values given by the relation of Eq. (5.1).

### 5.1.3 Adaptive selection for sub-hologram generation

The duality between Gabor wavelets and diffracted light beams, coupled with their excellent TF frequency localization, can be exploited for an accurate view-dependent representation of digital holograms. To this end, we design an adaptive selection process extracting a suitable subset of Gabor atoms from the K-expansion $A_K$ generated by MP. This process should meet the following requirements

**Fast atoms selection:** the relation between the atoms' indexes and the viewing parameters should be straightforward to avoid intensive computations during visualization. This is particularly the case for the relation given by Eq. (5.1).

**Faithful sub-hologram reconstruction:** the degradation induced by the atoms selection should not significantly impact the image quality within the viewing window.

**Reasonable compression:** although the reduction of the transmitted data amount is sought from the atoms pruning, an efficient encoder is required to allow an online transmission. A measure of the compressibility can be evaluated with respect to the overall compressed hologram size.

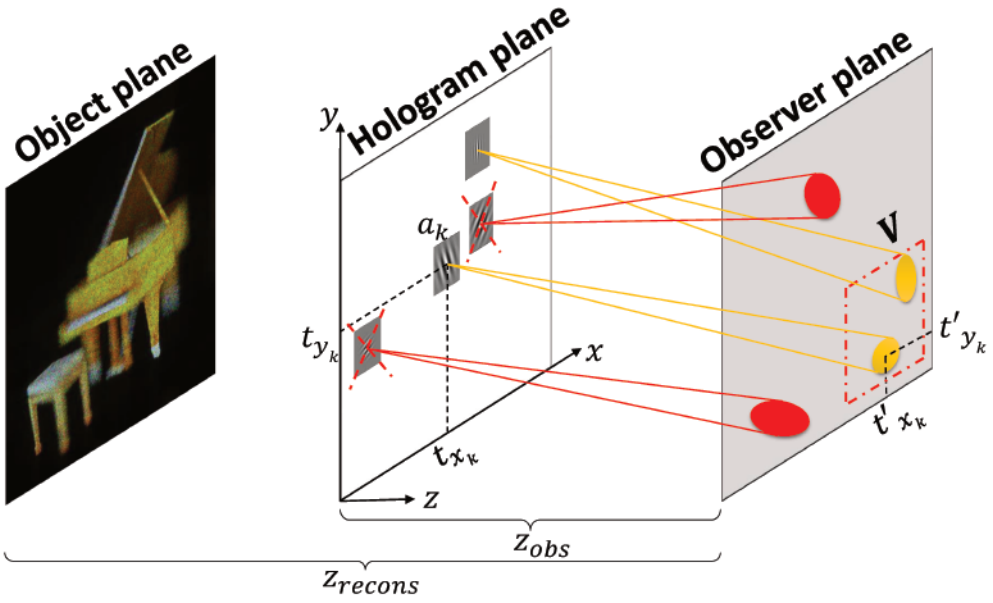



Fig. 5.4 Schematic principle of the adaptive selection process. *The sub-hologram $\boldsymbol{H}_V$ is generated from the Gabor atoms having their diffraction spectrums' centers inside the viewing window V*

**Reducing the impact of seeking the sparsest representation:** choosing wavelets that provides good trade-off between space and frequency localization guarantees an identification of the hologram's diffraction properties, but leads to a large choice of possible representations, some of which may not visually preserve the directional information.

As depicted in Figure 5.1, a sub-hologram corresponds to all diffractions gratings emitting light into the eye's pupils. Accordingly, Figure 5.4 illustrates the process of selecting the relevant atoms to generate a sub-hologram based on the observer's position. It consists on selecting all the Gabor atoms having diffraction spectrums' centers inside the viewing window $V$ (yellow ones) and discard all the remaining atoms (red ones).

Thus, the subset of $M$ atoms defining the sub-hologram $\mathbf{H}_V$ corresponding to the visualization window $V$ is given by

$$A_V = \{a_k \in A_K \quad | \quad C_k \in V, \quad 1 \leq k \leq K\}. \tag{5.2}$$

For transmission purposes, the partial bitstream encoding the sub-holograms can be generated using the coding framework developed in Section 4.3 using the same quantization and entropy coding choices. In this way, an efficient viewpoint scalable compression scheme is achieved.

## 5.2 Quality scalable compression of generated sub-holograms

Despite the significant data reduction brought by viewpoint scalability, the bitstream encoding each sub-hologram does not permit a progressive incrementation of the reconstruction quality. To achieve such functionality, we propose to sort the Gabor atoms of each sub-hologram according to their importance for the reconstructed view.

First, the principle of conventional quality scalability is provided in Section 5.2.1. Then, the proposed quality scalable compression scheme for digital holograms is described in Section 5.2.2.

### 5.2.1 Quality scalability

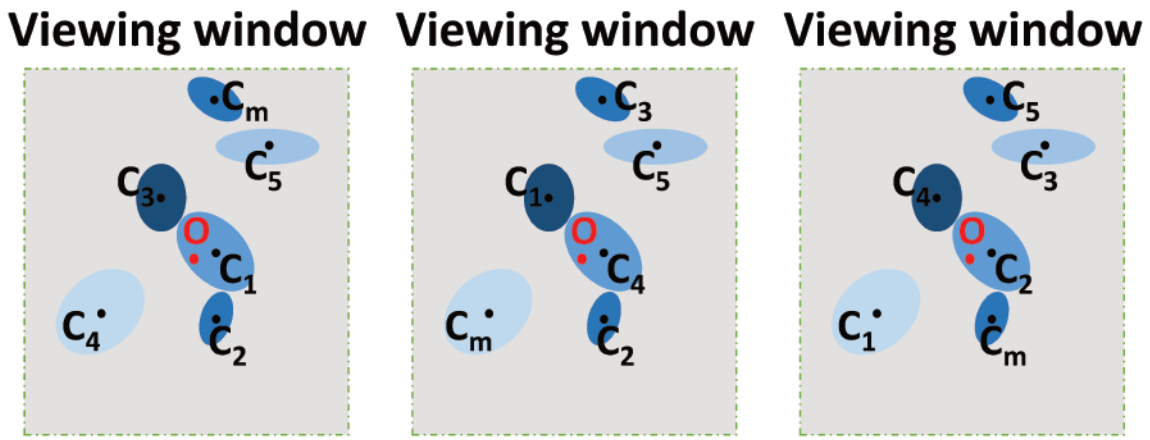
The purpose of quality scalability (*i.e.* SNR scalability) is to encode the original image / hologram into a single embedded bitstream having several layers: a base layer to enable a rapid visualization of the content, and refinement layers to progressively enhance the visual quality while the data being downloaded. This functionality is highly useful for networks with time-varying bandwidth, since it allows decoding at any target bit rate by a simple truncation of the bitstream.

As discussed in Section 3.2.2, quality scalability provided by image and video codecs heavily relies on the RD optimization techniques such as in JPEG 2000, or the inter-layer prediction tools employed by H.264/H.265. Although such mechanisms may enable a quality SHC, they lead to weak compression gains because of the specific signal properties of holographic signals. Moreover, the relationship between the distortion in the hologram and reconstruction domains is not straightforward, which makes the design of an efficient quality scalable scheme more complicated.

Two quality scalable coding schemes can be considered for digital holograms. The first one aims at generating an embedded bitstream of the full hologram without taking into account the user's viewing conditions. The second one seeks for an embedded partial bitstream per viewpoint; *i.e.* sub-hologram. The latter scheme would be more beneficial for view-dependent displays, especially when the hologram's resolution is very high. In the following, we propose a solution for the second scalable coding scheme, using the sub-hologram representation introduced in the previous section.

### 5.2.2 Proposed method

**Sorting criteria** To achieve a quality scalability for each generated sub-hologram, we propose to sort the atoms of the set $A_V$ according to their importance for the reconstructed




(a) Distance $|OC_m|$ (b) Amplitude $|c_m|$ (c) Angular dispersion $\Delta\varphi_m$

Fig. 5.5 Sorting criteria for quality scalability of sub-holograms. (*Darker diffraction spectrums correspond to atoms with high amplitudes*)

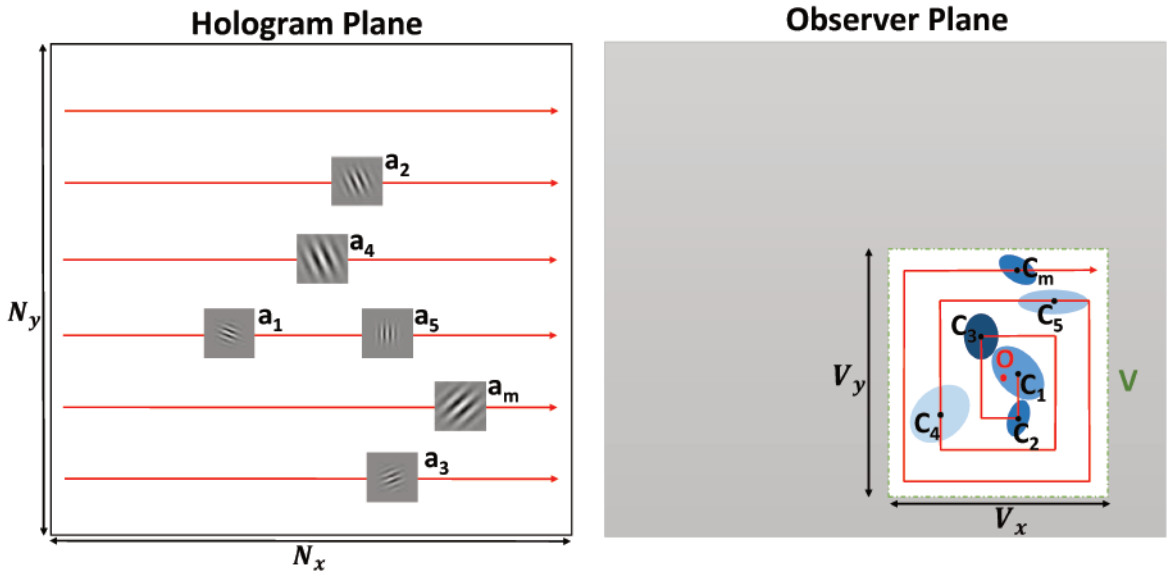
view. Since the quality of a given sub-hologram depends on the amount of light emitted by the Gabor wavelets into the viewing window, the sorting order depends on three criteria: $(1)$ increasing distance of the diffraction spectrum's center $C_m$ from the observation's center $O$ (*cf.* Figure 5.5a), $(2)$ decreasing light intensity given by the atom's amplitude $|c_m|$ (*cf.* Figure 5.5b), $(3)$ decreasing emission cone's size determined by the angular dispersion $\Delta\varphi_m$ (*cf.* Figure 5.5c).

Accordingly, we define three sorting functions $f_i : m \in [1, M] \mapsto [1, M]$ (where $M = Card(A_V)$), that give the order of the $m$-th atom according to criterion $(i)$. Then, the sorted set $A_V^i$ allowing the quality scalability for the sub-hologram $\mathbf{H}_V$ is given by

$$A_V^i = \{a_{f_i(m)}, \qquad 1 \leq m \leq M\}. \tag{5.3}$$

Moreover, a hybrid sorting can be defined by combining the three criteria. For example, $A_V^i$ may be divided into different subsets following a decreasing threshold on amplitude and/or angular dispersion, and then sorting each obtained subset according to criterion $(1)$.

The sorted M-expansion depicted by Eq. (5.3) provides a fine-grain level of quality scalability. Indeed, each compressed sub-hologram can be decoded from the first received atom, *i.e.* from the base layer. Then, the visual quality is progressively enhanced by decoding the remaining atoms until the target quality or the downloading time is achieved. The set $A_V^i$ may be also split into different layers with a given number of atoms, achieving a coarse-grain level of quality scalability. In such configuration, the atoms can be encoded by packets.




(a) Raster scan (b) Spiral scan

Fig. 5.6 Raster scan in the hologram plane (a) vs Spiral scan in the observer plane (b).

**Encoding in the observer plane** As reported in 4.3.3, the atoms of the MP expansion are encoded in the HP following a raster scan order (*cf.* Figure 5.6a). Although a differential coding of 2D positions allows a significant bit rate reduction, the encoded bitstream does not enable a proper progressive decoding.

To allow an efficient coding of atoms' positions without discarding the quality scalability, we propose to encode the positions of diffraction spectrums' centers in the OP instead of atoms' positions in the HP, by taking benefit from the correspondence given in Eq. (5.1). Accordingly, the centers' positions may be encoded in a differential spiral scan order as shown in Figure 5.6b. It must be noted that such encoding is efficient only in case of a quality scalability based on the distance criterion. For a proper progressivity based on the amplitude or angular dispersion criterion, the centers' positions are encoded with a fixed length code, *i.e.* using $\log_2(V_x) + \log_2(V_y)$ bits per atom's position.

Hereafter, the $m^{th}$ atom's index of the set $A_M^i$ is defined by

$$\gamma_m = (l_m, p_m, \boldsymbol{t}'_m), \tag{5.4}$$

where the coordinates $(t'_{x_m}, t'_{y_m})$ of the $m^{th}$ diffraction spectrum's center are related to the $m^{th}$ atom's position ones $(t_{x_m}, t_{y_m})$ by Eq. (5.1).

| **Hologram** | **DiffuseCar4K** | | | **Ring4K** | | | **Piano4K** | |
|---|---|---|---|---|---|---|---|---|
| **View** | *TopLeft* | *Central* | *Right* | *TopLeft* | *Central* | *Right* | *BottomLeft* | *BottomRight* |
| $\mathbf{O}_x$ (mm) | 0.41 | 0.82 | 1.22 | 0.41 | 0.82 | 1.22 | 0.41 | 1.22 |
| $\mathbf{O}_y$ (mm) | 0.41 | 0.82 | 0.82 | 0.41 | 0.82 | 0.82 | 1.22 | 1.22 |
| $\mathbf{z}_{obs}$ (mm) | 0.5 | 0.83 | 0.62 | 0.73 | 1.22 | 0.91 | 1.05 | 1.05 |
| $\mathbf{z}_{recons}$ (mm) | 1.69 | 2.02 | 1.81 | 2.32 | 2.81 | 2.5 | 2.41 | 2.08 |

Table 5.1 Observer parameters.

## 5.3 Experimental results

In this section, the effectiveness of the proposed scalable holographic compression approach, in terms of data reduction and progressive coding, is evaluated in a view-dependent context.

The holographic database as well as the observation parameters used for the experiments are described in Section 5.3.1. Then, the performance of the introduced viewpoint and quality scalable coding schemes is assessed in Sections 5.3.2 and 5.3.3, respectively.

### 5.3.1 Hologram and observer parameters

Three CGHs are used in the conducted experiments, including the *DiffuseCar4K* and *Ring4K* holograms generated from multiview-plus-depth projections of the 3D scene, and the *Piano4K* hologram generated from the synthetic scene representing a piano. These holograms have a 4K resolution, a pixel pitch of $0.4\mu\mathrm{m}$, and a wide FoV of $53.1°$, allowing numerical reconstructions from different point of views.

Table 5.1 summarizes the observer parameters for the three considered viewpoints. These values are given in accordance with the system coordinates defined in Figure 5.4.

The partial reconstruction $\mathbf{R_V}$ corresponding to a given view $V$ is obtained by propagating the cropped wavefield in the OP into the reconstruction plane. Using the AS propagation, $\mathbf{R_V}$ is given by

$$\mathbf{R_V} = \mathcal{P}^{AS}_{-z_{recons}}(\mathcal{P}^{AS}_{z_{obs}}(\mathbf{H}_V)). \tag{5.5}$$

To reduce the speckle effect in numerical reconstructions, we use a large cropping window of size $1600 \times 1600$ pixels. Moreover, it must be noted that the adaptive selection window is slightly larger than the cropping window used for the reconstruction, in order to take into account the diffraction spectrums situated at the border of the visualization area.

### 5.3.2 Performance evaluation of the viewpoint scalable coding scheme

**RD performance** To evaluate the effectiveness of our viewpoint scalable coding scheme, we compare the RD graphs corresponding to the numerical reconstructions obtained from:

- The sub-hologram $\mathbf{H}_V$ encoded from the Gabor M-expansion using Matching Pursuit plus Adaptive Selection (MP+AS).
- The entire hologram **H** encoded from the Gabor K-expansion using the MP encoder (MP-only).
- The entire hologram **H** encoded using the HEVC codec in intra-mode.

It is worth mentioning that a sub-hologram can be generated by propagating the hologram to the OP, cropping the obtained wavefield to the observer window and then propagating it back the HP. Then, HEVC can be applied on the generated sub-hologram instead of the full hologram. However, the forward and backward light propagation are very time-consuming, and thus unpractical for real-time view-dependent displays.

Alternatively, sub-holograms can be approximated by a simple cut in the HP following the viewing window. Nevertheless, because of the non-local character of hologram, light may be diffracted from any region of the hologram, especially for observer far from the HP. This justify our choice to use HEVC on the entire hologram as a benchmark solution.

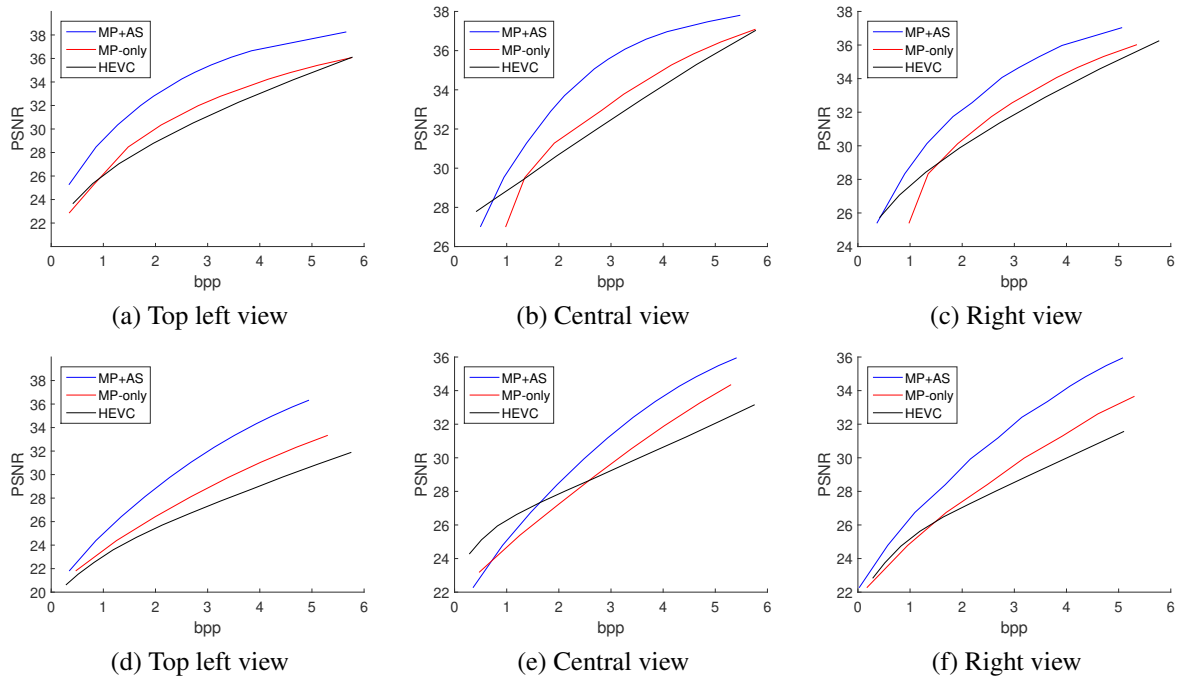
The RD graphs obtained by MP+AS, MP-only and HEVC for the three considered reconstructed views of the *DiffuseCar4K* and *Ring4K* holograms are presented in Figure 5.7.

As shown by RD curves, our approach clearly outperforms HEVC in terms of compression performance for *Right* and *Top left* views at all bit rate ranges, and for *Central* view especially at medium and high bit rates. Moreover, a significant gain is achieved by MP+AS compared to MP-only for a given reconstruction quality, since less atoms are encoded.

Table 5.2 summarizes the atoms reduction ratio ($1 - \frac{M}{K}$), as well as the compression gain ($1 - \frac{R_{MP+AS}}{R_{MP-only}}$) obtained for a reference PSNR value $P_{ref} = 34.8$dB. As shown by this table, the bit rate gains achieved by the pruning process depend on the considered point of view. For the same reconstruction quality, a compression gain of $45.8\%$ is attained for the *Top left* view of *DiffuseCar4K*.

To confirm the objective compression results obtained by the RD graphs, the three reconstructed views corresponding to are presented in Figure 5.8. The *Top left*, *Central* and *Right* reconstructed views are shown in the first, second and third columns, respectively. The considered bit rate of the compressed *Ring4K* holograms is $bpp_2 = 2.65$.

For all reconstructed views, the sub-holograms compressed using MP+AS expansion presents a better visual quality compared to HEVC and MP-only, which confirms the RD




(a) Top left view (b) Central view (c) Right view

(d) Top left view (e) Central view (f) Right view

Fig. 5.7 RD graphs corresponding to numerical reconstructions of *DiffuseCar4K* (first row) and *Ring4K* (second row), compressed using: HEVC, MP-only and MP+AS (our approach).

| **Hologram** | **DiffuseCar4K** | | | **Ring4K** | | |
|---|---|---|---|---|---|---|
| **View** | *Top left* | *Central* | *Right* | *Top left* | *Central* | *Right* |
| **Atoms reduction** ($\in [0,1]$) | 0.46 | 0.41 | 0.35 | 0.39 | 0.26 | 0.37 |
| **Bit rate gain** (%) | 45.8 | 36.4 | 26.8 | 33.4 | 17.7 | 29.3 |

Table 5.2 Comparison of MP+AS and MP-only: Atoms reduction and bit rate gains.

results. For instance, when considering the *Top left* reconstructions of *Ring4K*, a gain of 4.5dB is reached by the atoms pruning compared to HEVC, where only 3dB is achieved without the adaptive selection process.

Based on the obtained results, we conclude that the sparse MP representation does not affect the adaptive selection process introduced in Section 5.1.3. Indeed, encoding sub-holograms generated by the combination of MP and the adaptive atoms pruning enables a significant coding gains compared to full hologram compression, without degrading the quality of the reconstructed view. Thus, an efficient viewpoint scalable coding scheme is achieved for digital holograms.

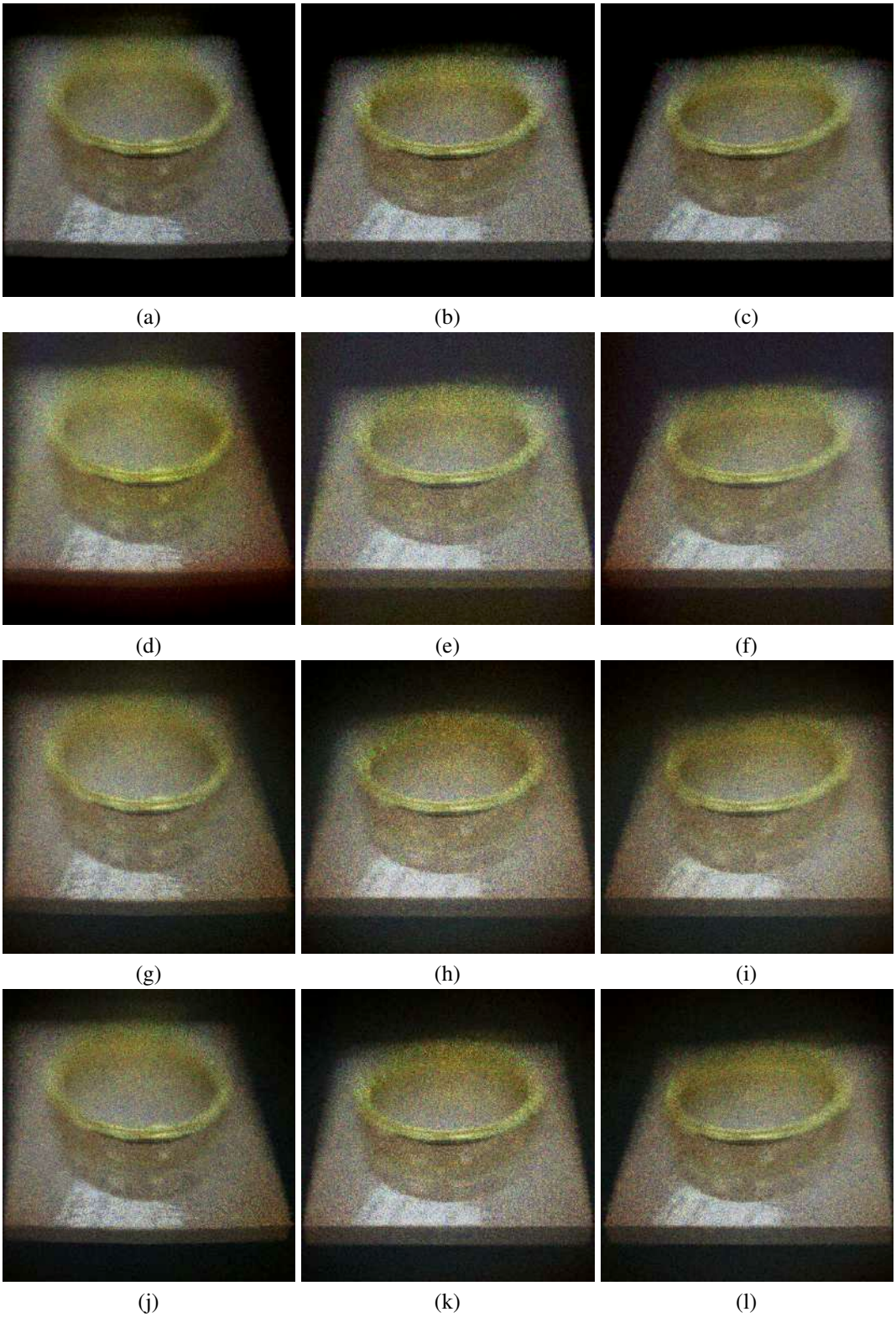


Fig. 5.8 Numerical reconstructions of original *Ring4K* (a-c), and compressed using: (d-f) HEVC, (g-i) MP-only, (j-l) MP+AS (proposed approach). The bit-rate is $2.65$bpp

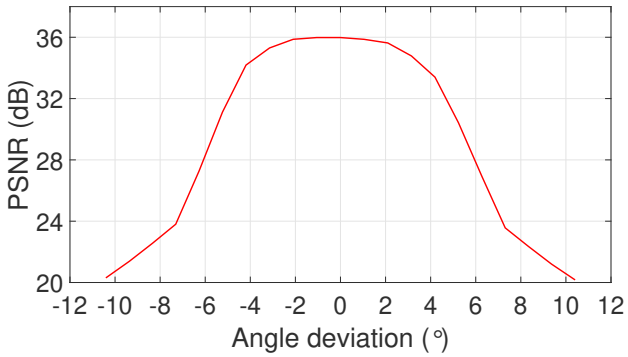


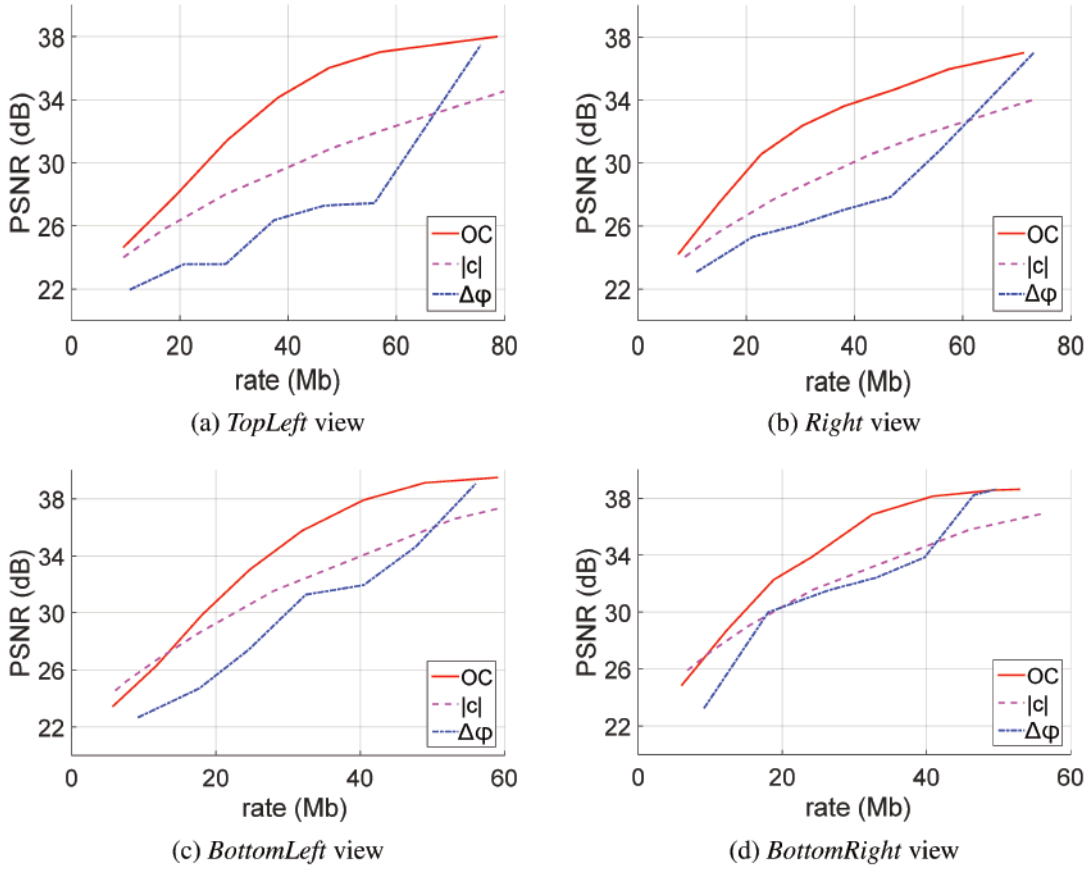

Fig. 5.9 Quality degradation of the neighboring reconstructed views from a sub-hologram corresponding to central view.

**Quality degradation of the neighboring views** Figure 5.9 shows the distortion related to numerical reconstructions of neighboring views obtained from a sub-hologram corresponding to a *Central* view of the *Ring4K* hologram. We consider a deviation angle of $\pm 10.4^{\circ}$ from the *Central* viewpoint. As noticed from the obtained curve, the reconstruction quality is not degraded for adjacent views with a $\pm 2^{\circ}$ deviation. Then, it decreases sharply beyond this value.

### 5.3.3 Performance evaluation of the quality scalable scheme

**Comparison of the sorting criteria in terms of quality increase** In the first part of these experiments, we compare the performance of the three sorting criteria in terms of progressive increase of quality. Figure 5.10 presents the RD graphs obtained by these criteria for the *Ring4K* and *Piano4K* holograms. According to the obtained curves, the sorting function based on the distance $OC$ criterion allows the best scalable coding performance. Indeed, it achieves compression gains of $42.5\%$ and $50.4\%$ (Bjontegaard Delta (BD) rate) over the amplitude and angular dispersion criteria, respectively, for the reconstructed *TopLeft* view of *Ring4K*. Moreover, for low and medium bitrates, the amplitude sorting criterion outperforms the angular dispersion one.

To confirm these objective results, the *Piano4K* sub-hologram corresponding to the *BottomLeft* view is compressed from the Gabor expansion sorted following the three sorting criteria, using a bit rate of $25$Mb. Then, the numerical reconstruction corresponding to each criterion is given in Figure 5.11. The reconstructed sub-hologram based on the distance scalability (Figure 5.11.b) presents less distortion compared to the amplitude (Figure 5.11.c) and angular dispersion (Figure 5.11.d) ones. Thus, the atoms with the nearest light beams to the viewer's center are the most important for the reconstruction and should be transmitted



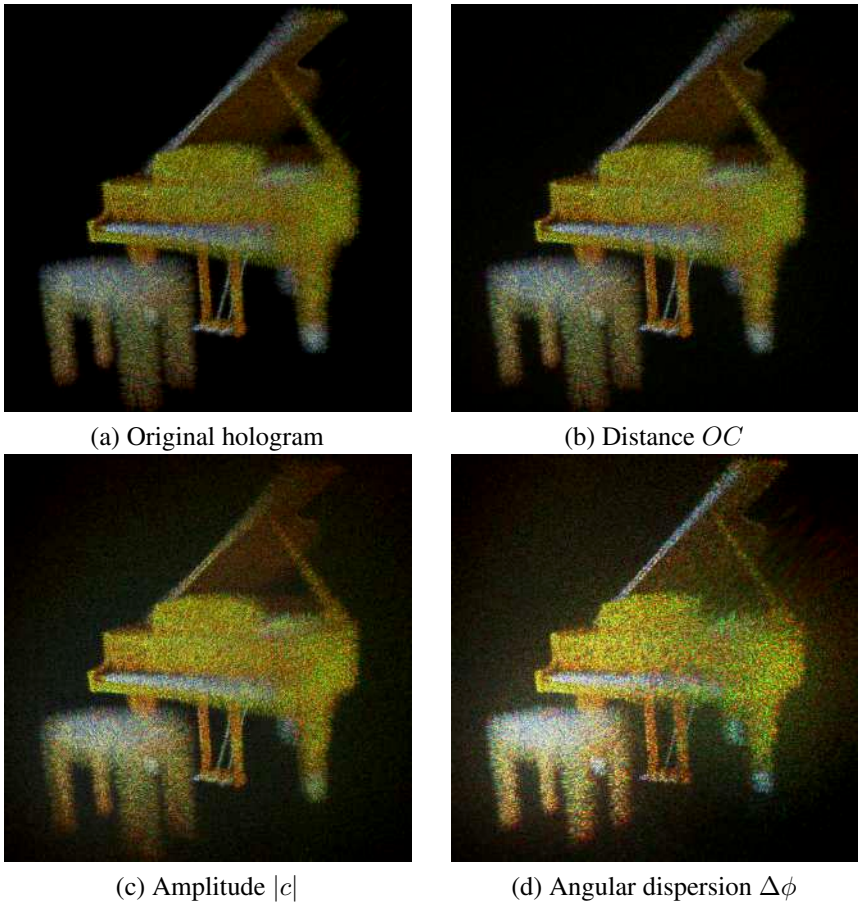

(a) *TopLeft* view (b) *Right* view

(c) *BottomLeft* view (d) *BottomRight* view

Fig. 5.10 RD graphs corresponding to numerical reconstructions of *Ring4K* (first row) and *Piano4K* (second row), compressed using: amplitude, distance and Angular dispersion criteria.

first. In the rest of the simulation tests, the quality scalability based on the proximity of the light beams to the observer's center, *i.e.* distance criterion, will be considered.

**RD performance** In the second part of experiments, we evaluate the performance of the proposed Quality-SHC (Q-SHC) scheme using the distance criterion. To this end, we compare our Q-SHC scheme to the Scalable High efficiency Video Coding (SHVC) [29] which is an easy, off-the-shelf solution, and a natural benchmark for our proposed coding scheme. The real and imaginary parts of the holograms are encoded using the reference software SHM12.4. The base layer is encoded with a QP of 50, whereas for the enhancement layers the QP values vary from 48 to 36 with a step of 2. Since the software only support *YUV 4:2:0* format for more than 2 layers, each hologram channel (R, G and B) is encoded separately after converting it to the *YUV 4:2:0* format (using null chrominance channels).



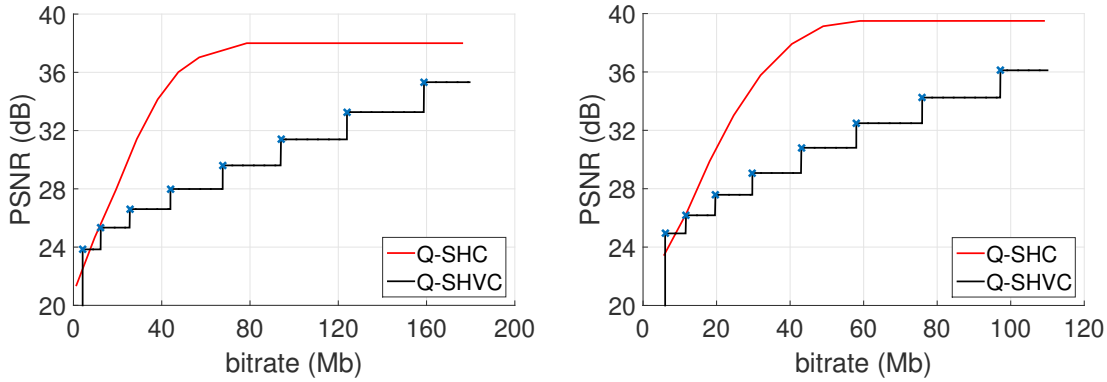
(a) Original hologram (b) Distance $OC$

(c) Amplitude $|c|$ (d) Angular dispersion $\Delta\phi$

Fig. 5.11 Numerical reconstructions of the *Piano4K* sub-hologram corresponding to the *BottomLeft* view: a bit rate of $25$Mb is used for the sub-holograms compressed from the Gabor expansion sorted following the three sorting criteria (b-d).

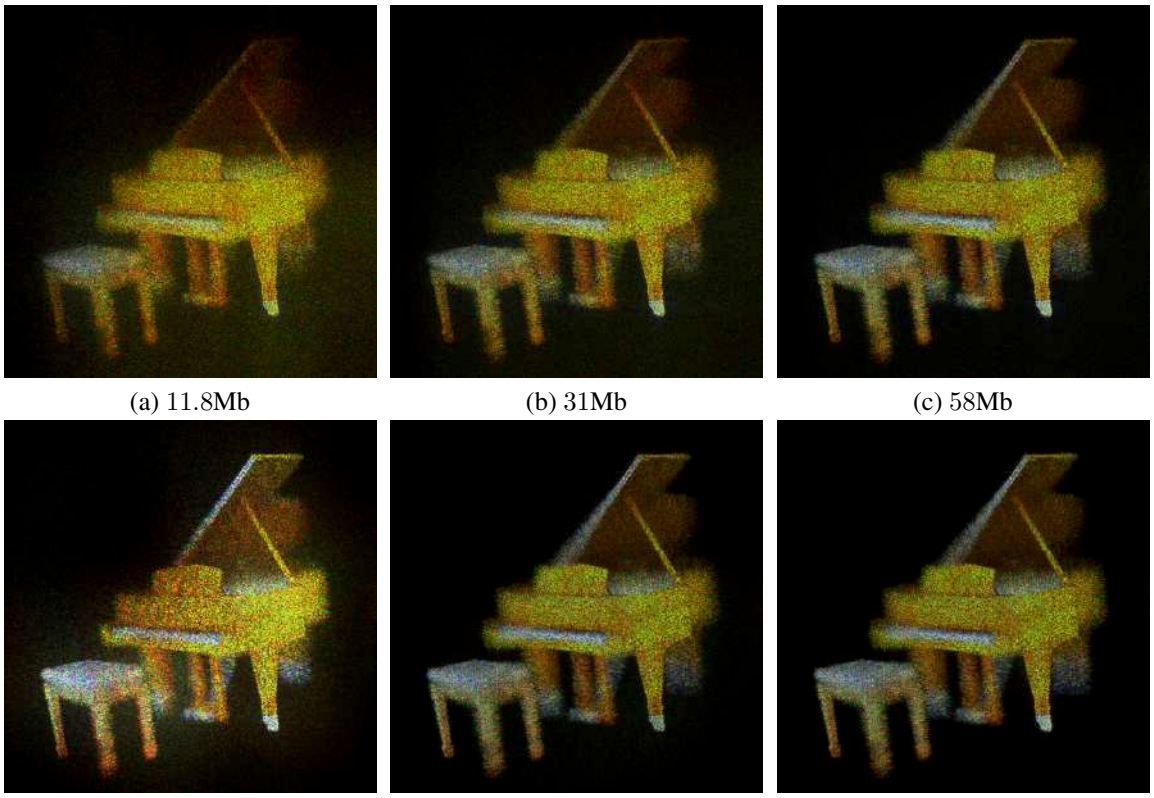
Figure 5.12 present the RD graphs obtained for the *Ring4K* and *Piano4K* holograms using Q-SHVC and Q-SHC schemes. As shown from the curves, the bitstream generated by the proposed approach allows a better progressivity in terms of reconstruction quality. Indeed, for the *BottomRight* view of *Piano4K*, the bit rate needed for Q-SHC to achieve a reconstruction quality of $32.5$dB is $23.8$Mb, whereas Q-SHVC attains the same quality for a bit rate of $58$Mb. Moreover, the maximum quality for the reconstructed *TopLeft* view of *Ring4K* is reached after $80$Mb for Q-SHC compared to more than $160$Mb for Q-SHVC.

To confirm the objective RD results, Figure 5.13 shows the numerical reconstructions corresponding to the *BottomRight* view of the compressed *Piano4K* hologram, decoded at different bit rates from the bitstreams generated by Q-SHVC and Q-SHC. As illustrated by the images, the visual quality increases more rapidly for Q-SHC compared to Q-SHVC, allowing then a better quality scalability.




(a) *View1* of the *Ring* (b) *View2* of the *Piano*

Fig. 5.12 Streaming simulations for a bandwidth of 10Mbps using Q-SHVC and VQ-SHC (our approach).



(a) 11.8Mb (b) 31Mb (c) 58Mb

(d) 11.8Mb (e) 31Mb (f) 58Mb

Fig. 5.13 Numerical reconstructions of the *BottomRight* view for the *Piano4K* hologram compressed using Q-SHVC (first row) and our approach (second row): the bitstream is truncated at different bit rates.

For Q-SHVC, the user has to wait until the whole base layer is received and decoded before having access to the first reconstructed view. However, the bitstream generated by our quality scalable coding scheme may be decoded at any bit rate from the first received atom. This fine-grain level of scalability is important to enable a rapid visualization of the sub-hologram by decoding the first received bits. Thus, an efficient quality scalable coding scheme is achieved for digital holograms.

## 5.4 Conclusion

In this chapter, we introduced our second contribution, which aims at designing a scalable holographic coding scheme dedicated to a view-dependent context. The proposed scheme is based on a Gabor decomposition of digital holograms, and consists on two levels of scalability: viewpoint and quality scalability.

The viewpoint scalability enables a significant data reduction by encoding only the information corresponding to the viewer position, *i.e.* sub-hologram. This was achieved thanks to the accurate space/frequency localization of Gabor wavelets, by adaptively selecting all Gabor atoms modeling the light beams passing through the viewing window.

The quality scalability is then developed for a progressive compression of the generated sub-holograms. The proposed approach consists of progressively encoding the most important Gabor atoms for the reconstructed view.

Experimental results revealed that our holographic scalable coding scheme clearly outperforms the scalable extension of HEVC in terms of RD performance and visual quality. Indeed, the numerical reconstructions of sub-holograms generated by the viewpoint scalable scheme provide the same reconstruction quality as the ones obtained from the entire compressed holograms, achieving remarkable compression gains. Moreover, the proposed quality scalable scheme permits a fine-grain progressivity with a rapid increase of reconstruction quality, compared to conventional quality scalable extension of HEVC.

# Part IV

# Digital hologram transmission

# Chapter 6

# Hologram streaming using viewpoint-quality scalable coding

The main objective of data compression is to alleviate the capabilities of storage mediums, and reduce the bandwidth requirements for a real-time transmission over communication networks. As illustrated by the example given in Chapter 1, downloading a $250K \times 250K$ raw monochromatic hologram using a Wi-Fi connection of $100$Mbps bandwidth would take about 1 hour and 23 minutes. Thus, it is obvious that today's wired/wireless communications channels will not be able to transmit high resolution holograms in time compatible with practical applications.

When it comes to transferring massive datasets over bandwidth-limited or noisy channels, streaming is often considered as a powerful tool to achieve efficient transmission. Indeed, it allows rapid and continuous access to the content while the data is being delivered to the end-user; *i.e.* before it is fully received. In general, conventional image/video streaming is achieved either by efficient scalable coding schemes or by adaptive encoding of the content at several quality levels in order to meet the bandwidth constraints. However, when considering digital holographic data, the first solution is not compatible with the wave propagation model of holographic patterns, and the second will lead to intractable amount of data to be stored.

To overcome the limitations imposed by traditional streaming solutions, we developed a holographic streaming framework in a server-client environment, using the Viewpoint-Quality Scalable Hologram Coding (VQ-SHC) scheme designed in Chapter 5. The remainder of this chapter is structured as follows. First, Section 6.1 provides the principle of holograms streaming, the shortcomings imposed by conventional image and video streaming approaches, and finally describes the proposed holographic streaming solution. Then, the results obtained from digital hologram streaming simulations are analyzed in Section 6.2.

## 6.1 Digital hologram streaming

The communication networks used in multimedia services has time-varying bandwidth, which means that the connection speed between servers and clients may vary frequently. This involves that the media encoding file size should fit the available network's bandwidth at any time in order to enable an online transfer and continuous access. This is the main purpose of streaming proposed as an alternative to file downloading. However, developing effective streaming architectures is a very challenging task, especially in case of bandwidth-limited channels and very large media files.

With the recent advances in data compression techniques and communication network evolution, immersive media streaming has been extensively studied, giving rise to efficient transmission systems for omnidirectional videos [35, 50, 109], light-field [119, 122] and volumetric media [153]. To the best of our knowledge, no effort has been carried out for designing digital hologram streaming solutions.

Accordingly, we provide some insights about the general principle of holographic streaming in Section 6.1.1. Then, we explain in Section 6.1.2 why conventional approaches deployed by image and video streaming architectures are incompatible with digital holograms. Finally, the proposed solution is described in Section 6.1.3.

### 6.1.1 Principle

Experientially, the working concept of media streaming means that when the user presses the playback button, the media content will be immediately available for the end-user. While the compressed file is being transferred, the media continues playing with more or less quality with respect to the network's bandwidth variation. In contrast, when using downloading, the user has to wait until the whole media file is transmitted before having access to it.

Technically, the main challenge in streaming systems is to continuously find the optimal trade-off between the accessible bandwidth and the quality of the received data. This depends on the effectiveness of compression algorithms used to encode contents, as well as the design of transmission facilities.

Streaming has been initially designed for video and audio signals, which can be streamed either live or on-demand depending on the target use-case [10]. On-demand streaming is provided by a means called *progressive streaming* or *progressive download*, where the file is saved to a hard disk and then played from that location. This concept has been also used for transmission of large images using JPEG 2000, allowing progressive visual display. Streaming digital holograms will be thus considered as a progressive downloading, since they are represented as still images.

### 6.1.2 Limitations imposed by image and video streaming solutions

The majority of conventional image and video streaming solutions are based on two coding paradigms: scalable coding and Multiple Description Coding (MDC) [95]. In the following paragraphs, we emphasize the limitations of these two paradigms for digital hologram streaming.

**Scalable coding:** has been widely deployed for image streaming [64, 65] and video streaming [8, 32, 45, 121, 183], since it adjusts the amount of data to be transmitted according to the network conditions. Thus, a practical system for holographic data streaming would benefit a lot from SHC, where a single bitstream contains a base layer, allowing an instantaneous access to a rough representation of the hologram, and a number of enhancement layers, each of which progressively increases the quality of the coded hologram. As stated in Chapter 3, a first implementation of SHC could be obtained by using scalable image or video codecs on holographic signals. However, as shown in Section 4.1.2, the compression efficiency of this solution is sub-optimal, since scalable codecs heavily rely on the natural image and video characteristics to achieve effective compression. Furthermore, since conventional scalable coding schemes do not allow viewpoint scalability for digital holograms, they must be fully compressed and transmitted. Given the huge amount of data to stream, the client will not be able to display the target view immediately and the user has to wait until the base layer is received and decoded. One more limitation of the layered coding is that each layer is crucial for decoding its successor. Therefore, for unreliable communication networks, the streaming will be interrupted whenever the base layer or a given enhancement layer is missing.

**Multiple Description Coding:** consists on encoding the original signal into two (or more) separate substreams referred to as *descriptions*. MDC has been developed as an alternative to cope with the packet losses that may be induced by layered coding [36], since each description can be decoded independently, being then useful for retrieving the signal. Moreover, it provides increasing quality as more descriptions are received. MDC has been extensively used for image transmission [96, 98, 164], and video streaming [97, 152, 159, 196]. Digital hologram streaming systems based on this paradigm may be realized by encoding the holographic signal into different bitstreams, one for each level of quality, resolution and viewpoint. These multiple descriptions generated by the coder are downloaded to the server, who then chooses the appropriate description to transmit based on the current connection speed and possibly on other parameters. However, given the considerable amounts of data contained in

holograms, such approach would be limited by the server computation and storage capabilities. Moreover, such solution would not mitigate the large time needed for the first visualization of a hologram.

More sophisticated streaming solutions have been developed such as the Dynamic Adaptive Streaming over HTTP also known as MPEG-DASH [127]. It was designed particularly for audio and video contents, and consists on splitting the audio/video into different small segments of short durations (some seconds), encoding each of them at several bit-rates or resolutions. Then, the client adaptively selects the segment with the highest possible bit rate that can be downloaded with the current network's bandwidth. Despite its high performance for video streaming, MPEG-DASH is not adapted for holographic image streaming.

### 6.1.3 Proposed solution

A typical use-case of holographic streaming would be in a server-client scenario: at the server side, we can consider a library of digital holograms with several parameters, which are already generated and stored. Then, a request is sent by the client to notify the server about the hologram to be visualized, the current user's viewpoint and the network's bandwidth.

As discussed in the previous section, current image and video streaming architectures are not suitable for transmitting high quality holograms with large resolution and FoV. In the following, we provide the overall description of our proposed solution.

Figure 6.1 shows the global server-client architecture of the proposed holographic streaming scheme, which is composed of an offline and online computation phases.

The offline phase corresponds to the generation and light beams-based sparse decomposition of the holographic data to be streamed. First, digital holograms are generated by numerical computations from 3D scene models, or acquired on a CCD connected to server computers. Then, an overcomplete set $A_N$ with $N$ Gabor atoms is produced by decomposing the target hologram $\mathbf{H}$ through a family of 2D complex Gabor wavelets. Finally, a sparse K-expansion is obtained using MP algorithm, and the set $A_K$ is stored in the server's RAM memory.

The online computation part corresponds to the proposed streaming framework based on the VQ-SHC scheme designed in Chapter 5. First, the client sends periodically a notification through back-channel to inform the server about the current observer's position $O$, which is determined using an eye-tracking system [170]. Based on this information, the Gabor atoms allowing viewpoint-quality scalable representation of the corresponding sub-hologram are extracted from the K-expansion. The obtained set $A_M$ is then split into different layers, where the atoms number of each one is estimated with respect to the available bandwidth.

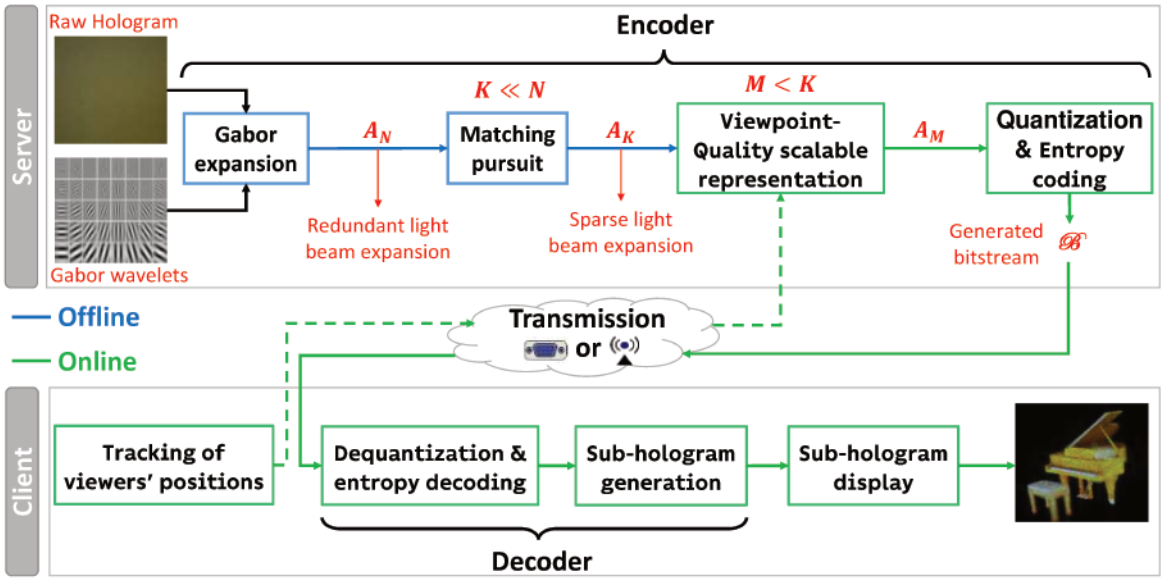



Fig. 6.1 Proposed server-client architecture for holographic streaming: *the blue blocks represent the offline computation, whereas the green ones correspond to the online calculation.*

In the next step, the atoms of each layer are quantized and entropy encoded using the MP encoder developed in Section 4.3. The generated embedded bitstream $\mathcal{B}$ is then progressively transmitted to client through forward channel. After decoding each received layer by the client, its atoms are weighted by their complex coefficients and summed to the hologram plane. Thanks to the fine-grain level of progressivity, the quality of the resulting sub-hologram is increased after each atom's contribution, and the 3D visualization is achieved by displaying it on the SLM.

For an immediate visualization of the target view, the base layer should be encoded with few Gabor atoms. When the user remains static, the following decoded layers contribute to refine the reconstruction quality until the highest one is reached. On the other hand, when the user moves from a position $O$ to $O'$, a new set $A_{M'}$ is generated by selecting the atoms emitting light beams into the window $V' \backslash V$.

For slow viewpoint's change, the overlapped region between two successive viewing windows is more important. Thus, the new view can be visualized from the decoded sub-hologram corresponding to current view, waiting for the atoms of next sub-hologram to be downloaded. As revealed by the experimental results obtained in previous chapter, a deviation angle of some few degrees can be tolerated without severely degrading the reconstruction quality of neighboring views.

However, when the user's position changes rapidly, the reconstruction quality decrease may be more important, since the number of new atoms to be encoded and transmitted

to client is higher. To anticipate any viewpoint switching lags caused by rapid observer's movements, the atoms contributing for the neighboring views are transmitted once a good quality is achieved for the current one. These atoms can be selected by predicting the future observer's positions. The impact of the viewer's velocity on the quality of experience will be experimentally studied in the next section from concrete streaming simulations.

It must be noted that the proposed streaming architecture can be also used in case of multi-observers, each visualizing the 3D scene from a specific point of view.

## 6.2 Experimental results

In this section, we evaluate the efficiency of the proposed holographic streaming approach in a server-client environment, by comparing its performance to conventional solutions based on image/video scalable codecs.

The holographic database as well as the observation parameters used for the experiments are described in Section 6.2.1. Then, the streaming simulations obtained by our designed architecture based on the VQ-SHC scheme are analyzed in Section 6.2.2.

### 6.2.1 Hologram and observer parameters

Two CGHs are used in the conducted experiments, including the *Piano8K* and *Dices8K* holograms computed from synthetic scenes representing a piano and three dices in front of a chessboard, respectively. These holograms have an 8K resolution ($8192 \times 8192$), a pixel pitch of $0.4\mu\mathrm{m}$, and a wide FoV of $53.1°$, allowing numerical reconstructions from different point of views.

For the conducted streaming simulations, we consider an observer having a horizontal trajectory with two phases: the first phase in which the viewer is static at a position $O_{left}$, and the second phase where the observer moves uniformly from $O_{left}$ to a position $O_{right}$. We denote by $\Delta t_{static}$ and $\Delta t_{moving}$ the duration of the static and moving phases, respectively. The considered values of these viewing parameters are summarized in Table 6.1 for both *Piano8K* and *Dices8K* holograms.

| | $\mathbf{\Delta t_{static}}$ (s) | $\mathbf{\Delta t_{moving}}$ (s) | $\mathbf{O_{left}}$ (cm) | $\mathbf{O_{right}}$ (cm) |
|---|---|---|---|---|
| **Piano8K** | 2.7 | 4.8 | (0, 0.16) | (0.33, 0.16) |
| **Dices8K** | 3.3 | 5.2 | (0, 0.16) | (0.33, 0.16) |

Table 6.1 Observer parameters.

### 6.2.2 Streaming simulations

**Time-Distortion (TD) performance evaluation** In the first part of experiments, we simulate the streaming of the sub-holograms corresponding to the viewer's trajectory defined in Table 6.1. To this end, we consider a time-invariant transmission bandwidth of $30$Mbps, which corresponds to a relatively high-speed wireless connection [81].

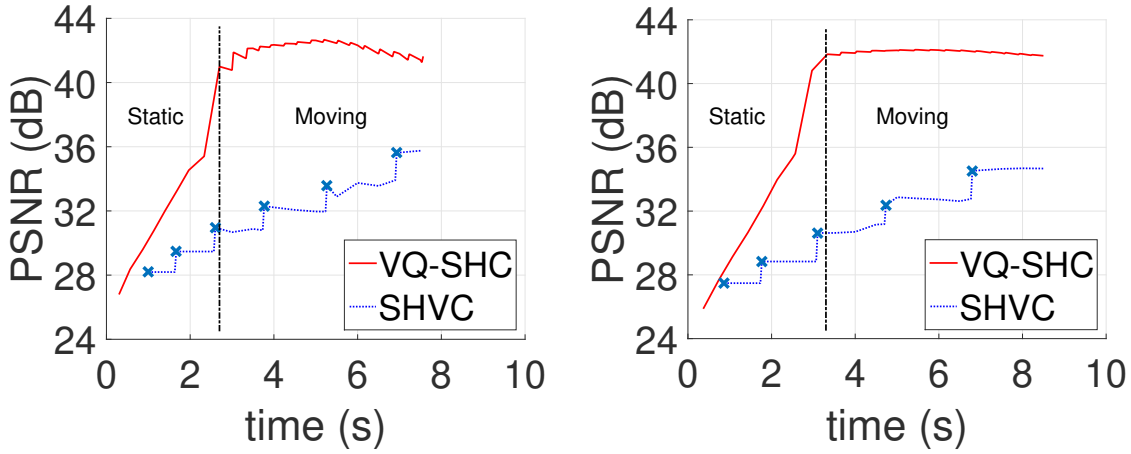
During the static phase, the bitstream encoding the sub-hologram of the left view is decoded progressively at different levels of quality. In our experiments, we consider 8 layers in this first phase. During the moving phase, once the observation center changes, the new bitstream generated according to VQ-SHC is transmitted and the corresponding sub-hologram is computed.

For the benchmark solution (SHVC), the reference software SHM$12.4$ is used to encode the real and imaginary parts of the original holograms. The base layer is encoded with a QP of 50, whereas for the enhancement layers the QP values vary from 48 to 40 with a step of 2. This choice is a convenient trade-off between the bitrate and distortion: for QP = 40, we achieve an acceptable image quality of $36$dB with a bitrate of $200$Mb. A smaller QP would lead to higher bitrate: $400$Mb for a QP of 34. Since the software only support *YUV 4:2:0* format for multi-layers, each hologram channel (R, G and B) is encoded separately as luminance (using null chrominance channels). It is judicious to note that the whole information about the holographic signal is encoded offline by SHVC. The decoding process is performed online and is independent of the user's movements.

The TD graphs of the streaming simulation for both holograms using SHVC and VQ-SHC are presented in Figure 6.2. At each time instant, we consider the viewpoint that can be reconstructed with each of the two methods, and we compute the PSNR with respect to the same viewpoint but reconstructed from uncompressed data. In the moments when a new layer is decoded for SHVC, the PSNR increases sharply, then it slowly fluctuates since the reconstructed viewpoint changes as the user moves. In any case, the user has to wait for the full reception of the base layer in order to start the visualization.

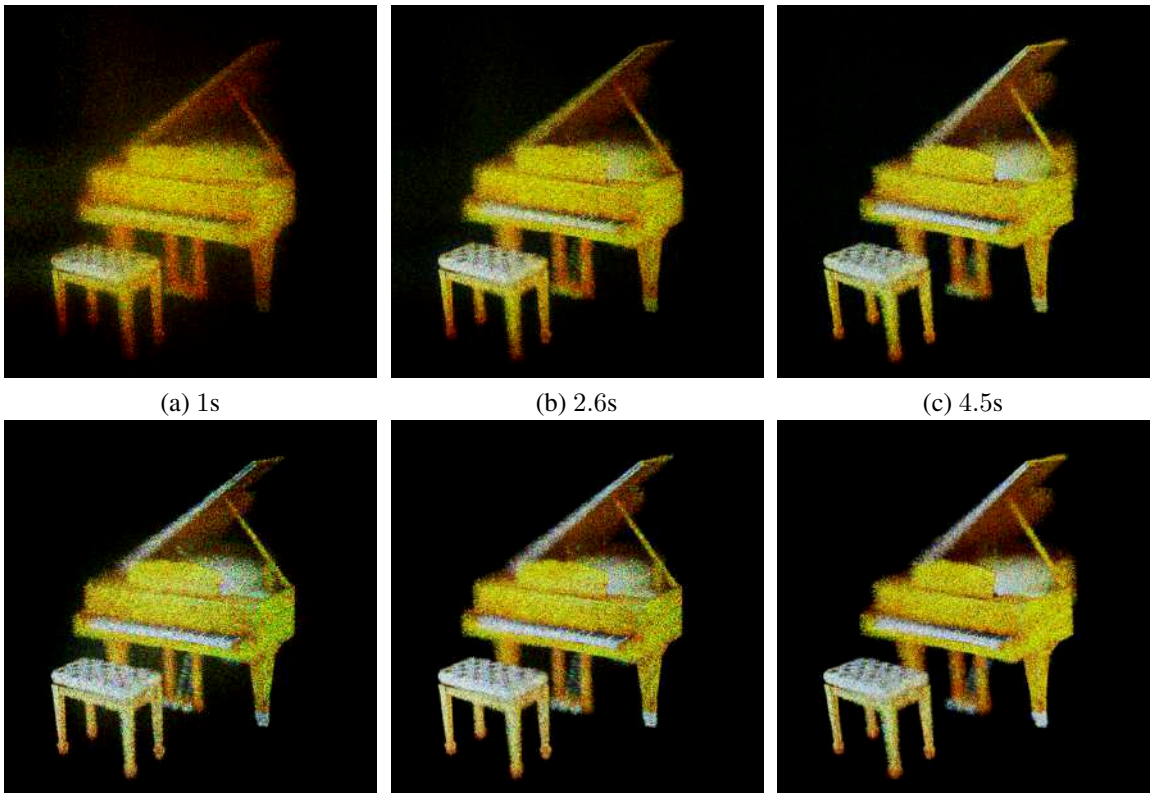
As illustrated by the curves, our approach allows a better performance in terms of scalability. Indeed, VQ-SHC achieves a PSNR of $30.6$dB for the *Piano8K* hologram after $1$s, whereas SHVC attains only $28.2$dB after the same duration. Also, it takes approximately $4$s for VQ-SHC to converge to a maximal reconstruction quality for the *Dices8K* hologram, whereas SHVC would need smaller QPs to achieve the same quality. We note that the small fluctuations in the PSNR values during the moving phase are due to the view changes.

These objective results are visually confirmed by the numerical reconstructions given in Figures 6.3, 6.4, and the videos provided in Appendix B. As shown by the reconstructions of the *Piano8K* and *Dices8K* holograms using SHVC, the shape and contrast of the scenes are




(a) *Piano8K* (b) *Dices8K*

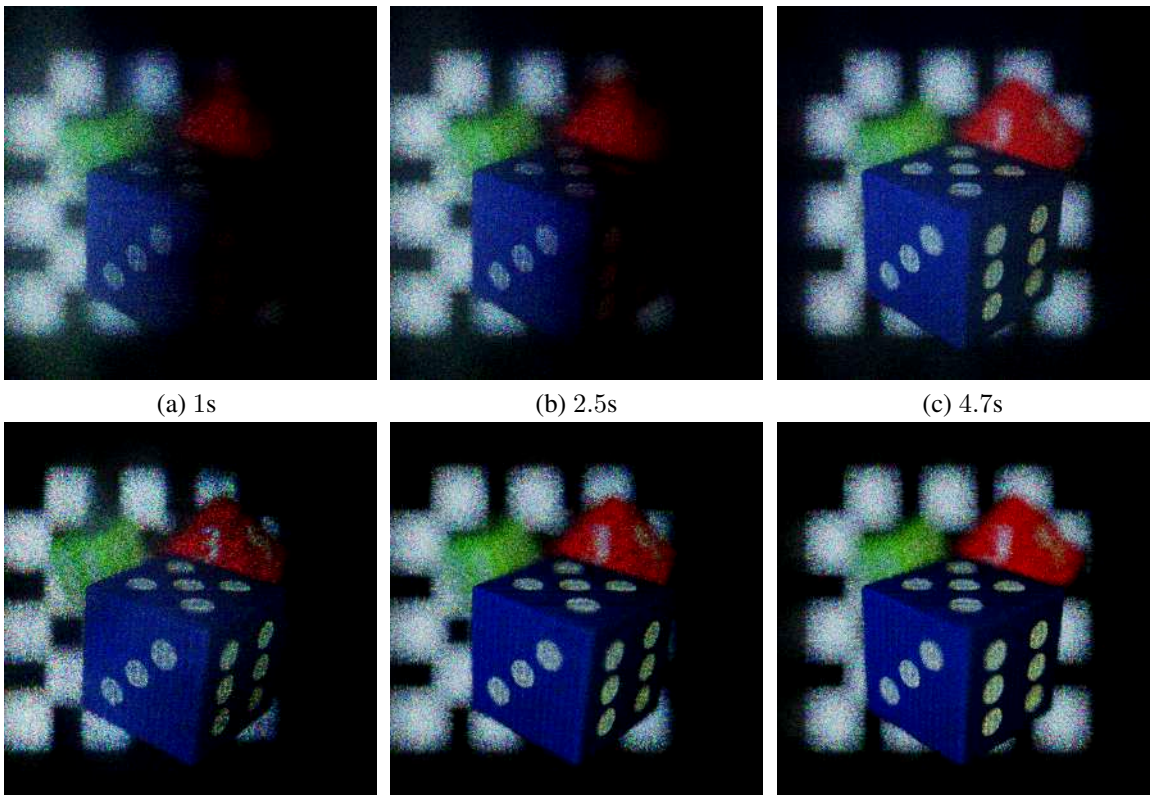
Fig. 6.2 TD curves for a simulated streaming of a moving viewer using SHVC and VQ-SHC (our approach) - The bandwidth is 30Mbps. (*The blue cross marks corresponds to the moments when a new layer is decoded for SHVC.*)



(a) 1s (b) 2.6s (c) 4.5s

(d) 1s (e) 2.6s (f) 4.5s

Fig. 6.3 Numerical reconstructions of *Piano8K* hologram compressed using Q-SHVC (first row) and our approach VQ-SHC (second row). The images are taken at different instants.



(a) 1s (b) 2.5s (c) 4.7s

(d) 1s (e) 2.5s (f) 4.7s

Fig. 6.4 Numerical reconstructions of *Dices8K* hologram compressed using Q-SHVC (first row) and our approach VQ-SHC (second row). The images are taken at different instants.

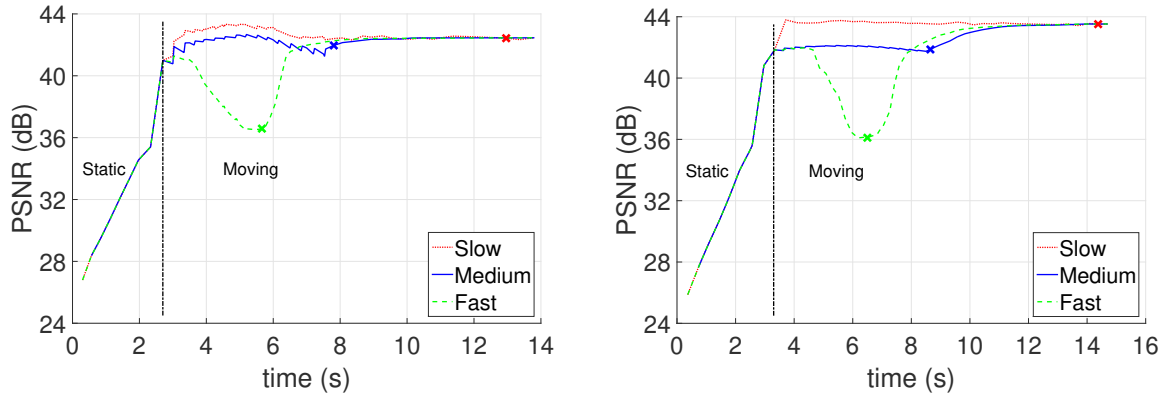
highly distorted after one second of transmission, which is the time needed for receiving the base layer. As the enhancement layers are decoded, the visual quality slightly improves but still exhibits some visible noise that affects the color information. On the other hand, these distortions are significantly reduced by VQ-SHC, which enables a good visual quality after only one second of transmission and converge to a visually lossless quality after $4$ seconds.

It must be noted that the proposed approach has the advantage of reducing the storage cost compared to conventional streaming solutions based on MDC, since the encoded information is not redundant. Moreover, decoding the atoms of a given layer is done independently of the previous ones, in contrast to scalable image and video streaming systems.

**Influence of the viewer's velocity** In the second test of streaming simulations, the influence of user's velocity on its quality of experience is studied. To this end, we consider three viewers' speeds: $V_s$, $V_m$ and $V_f$, corresponding to slow, medium and fast displacements, respectively. Table 6.2 summarizes the three speed values chosen for each hologram.

| | $V_s$ (mm/s) | $V_m$ (mm/s) | $V_f$ (mm/s) |
|---|---|---|---|
| *Piano8K* | 0.321 | 0.642 | 1.110 |
| *Dices8K* | 0.296 | 0.593 | 1.026 |

Table 6.2 Speeds corresponding to slow ($V_s$), medium ($V_m$) and fast ($V_f$) viewpoint changes.




(a) *Piano8K* (b) *Dices8K*

Fig. 6.5 Streaming simulation for three viewing speeds: $V_s$ (slow), $V_m$ (medium) and $V_f$ (fast) - The bandwidth is 30Mbps. (*The cross marks correspond to moments when the user reaches its last position. After this moments, the user remains static.*)

These values are computed by considering a user traversing the whole horizontal trajectory (3.276mm) in a given duration. As it can be noticed, the speed values are very small compared to real human scale due to the low resolution of the considered holograms. We assume that the streaming will finish when all the atoms of the K-expansion are downloaded. Moreover, once the user reaches its final position, he remains static until the end of streaming. The obtained TD graphs for both holograms are presented in Figure 6.5.

The TD curves reveal two important results:

$(i)$ During the moving phase, the reconstruction quality is degraded as the viewer's speed increases.

$(ii)$ At the end of the streaming, the reconstruction qualities converge to the same value, regardless the considered viewer's speed.

In the following, we provide some explanations of these phenomena.

In the moving phase, an average PSNR gain of 0.9dB and 1.5dB is achieved for slow displacements compared to medium ones, for *Piano8K* and *Dices8K* holograms, respectively. This result was expected, since for a slow viewpoint's change the user has more time to

download the atoms of the next view. For fast displacements, the quality of reconstructed views decreases sharply after a short movement duration ($0.8$s for *Piano8K* and $1.3$s for *Dices8K*).

Once the last viewpoint is attained, the remaining atoms are sent progressively. This justifies the fact that the quality increases a second time in case of medium and fast displacements, until reaching the same quality achieved by slow displacements.

## 6.3 Conclusion

In this chapter, we design a complete streaming framework specifically dedicated for holographic signals. The streaming architecture has been defined in a server-client environment, using the joint viewpoint-quality scalable coding scheme developed in the previous chapter. The proposed approach consists of two parts: an offline and online computation phases. During the offline computation phase, the hologram is generated and decomposed into a sparse Gabor expansion at the server side. In the online computation phase, the client notifies the server about the available bandwidth, the current user's viewpoint and possibly the prediction of its next viewpoint. Based on this information, the Gabor atoms are selected, sorted and encoded by the server into an embedded bitstream, which is transmitted to the client in order to enable a progressive streaming of the hologram following the observer's trajectory.

The experimental results reveal that the proposed streaming solution outperforms the conventional ones based on scalable video coding schemes such as SHVC. Indeed, the compression efficiency coupled with fine granularity of the viewpoint-quality scalable coding scheme enables a rapid first visualization of the holographic content, and achieves a better performance in terms of progressive increase of reconstruction quality. Thus, the proposed streaming framework is a first step towards an efficient progressive streaming of digital holograms with high resolutions.

# Chapter 7

# Low-complexity hologram streaming using viewpoint-quality scalable coding

An efficient holographic progressive streaming solution has been introduced in the previous chapter by combining the viewpoint and quality scalable coding schemes. Despite the high coding performance and fine granularity of the proposed approach, it implies an online selection, sorting and encoding of the Gabor atoms after each request from the client. Moreover, computing the sub-holograms of each user's viewpoint necessitates the calculation of the MP approximating linear combination. Thus, the computational complexity induced by the scalable encoding and sub-hologram generation may increase the latency during streaming and then impact the quality of experience, especially in case of multi-observers with frequent viewpoint changes.

To reduce the time between the moment when the server receives a notification, and the moment when the corresponding sub-hologram is displayed, we design a novel streaming architecture where the atoms selection, sorting and encoding are performed offline at the server side before starting the transfer. To achieve this, we propose a specific atom packetization to fully encode the Gabor expansion without discarding the viewpoint-quality scalability. Moreover, a GPU-based implementation is developed to alleviate the latency entailed by the sub-hologram generation.

The remainder of this chapter is organized as follows. First, Section 7.1 describes the proposed low-complexity streaming solution, including the design of the server-client architecture, the holographic encoder and decoder. Then, the experimental results provided in Section 7.2 evaluate the performance of the proposed approach in terms of compression efficiency and latency reduction.

## 7.1 Proposed streaming solution

Besides network's bandwidth limitations, one of the main causes that increases the streaming latency is the computational burden of the deployed codec. For instance, in live streaming systems where contents must be delivered online, the user has to wait for the data to be encoded, transmitted and decoded before having access to it. This gives rise to a frequent delay effect that affects the quality of experience and does not enable real-time visualization [16].

The VQ-SHC scheme used by the streaming framework developed in Chapter 6 has two main time-consuming operations, which are: Gabor atoms encoding and sub-holograms generation. In order to reduce the complexity induced by these operations, we propose a novel holographic streaming architecture where the atoms are grouped according to a block-based decomposition of the observer plane, and encoded offline. Then, the atoms of each block are assigned to different quality levels and encoded into packets.

Accordingly, the server-client structure of the designed low-complexity streaming framework is presented in Section 7.1.1. Then, the novel viewpoint-quality scalable holographic encoder is described in Section 7.1.2. For a given viewpoint, the adaptive selection process of relevant atoms' packets is explained in Section 7.1.3. Finally, Section 7.1.4 presents a fast implementation of sub-hologram generation using parallel calculation.

### 7.1.1 Server-client architecture

Figure 7.1 presents the architecture of the designed solution for a low-complexity hologram streaming in a client-server environment. The overall block-diagrams of the holographic encoder and decoder are given, as well as the transfers between the server and the client.

The inputs of the encoder are a discrete complex hologram $\mathbf{H}$ of resolution $N_x \times N_y$ and a family of 2D discrete Gabor wavelets. First of all, an overcomplete Gabor expansion $A_N$ is generated by computing the inner-products between the hologram $\mathbf{H}$ and the wavelets. To allow compression, MP algorithm is applied to the set $A_N$ yielding a sparse Gabor expansion $A_K$, where $K \ll N$. The proposed viewpoint scalable representation is obtained by splitting the atoms of $A_K$ following a block-based decomposition of the observer plane. Then, the atoms assigned to each block are grouped into different layers of quality, according to their importance for the reconstruction. The atoms' coefficients are quantized using a uniform scalar quantizer, and then losslessly compressed using Context-Adaptive Binary Arithmetic Coding (CABAC). Finally, the bitstream $\mathcal{B}$ encoding the full hologram is generated. Several markers can be inserted in $\mathcal{B}$ in order to enable the identification of the blocks' indexes and quality layers to be delivered. This can be done also by storing the partial bitstream

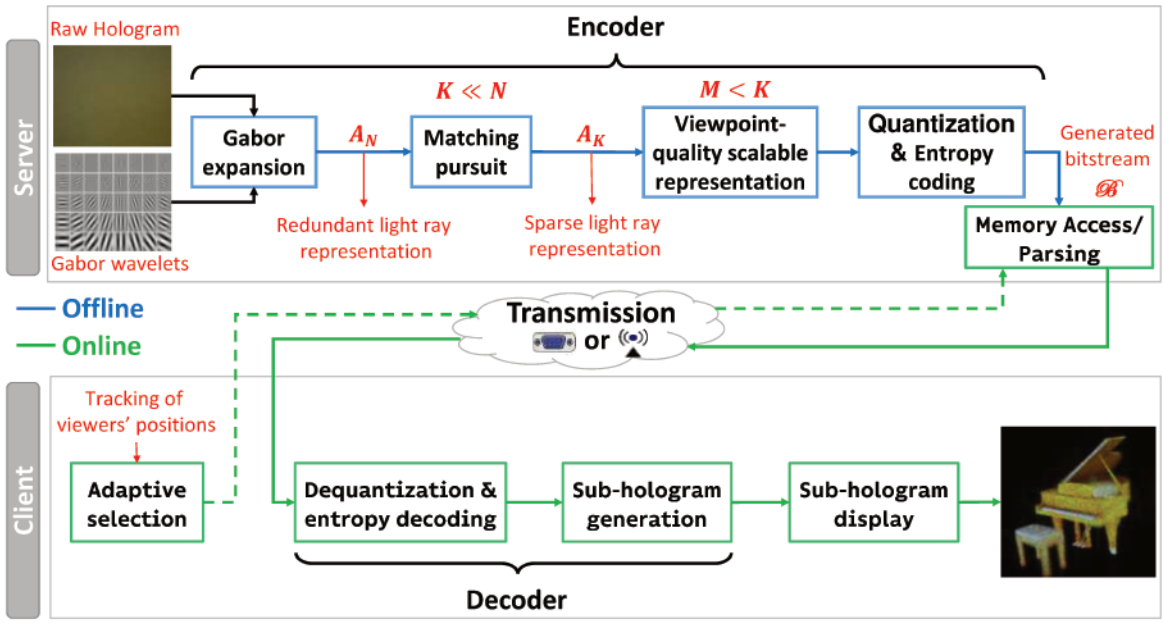


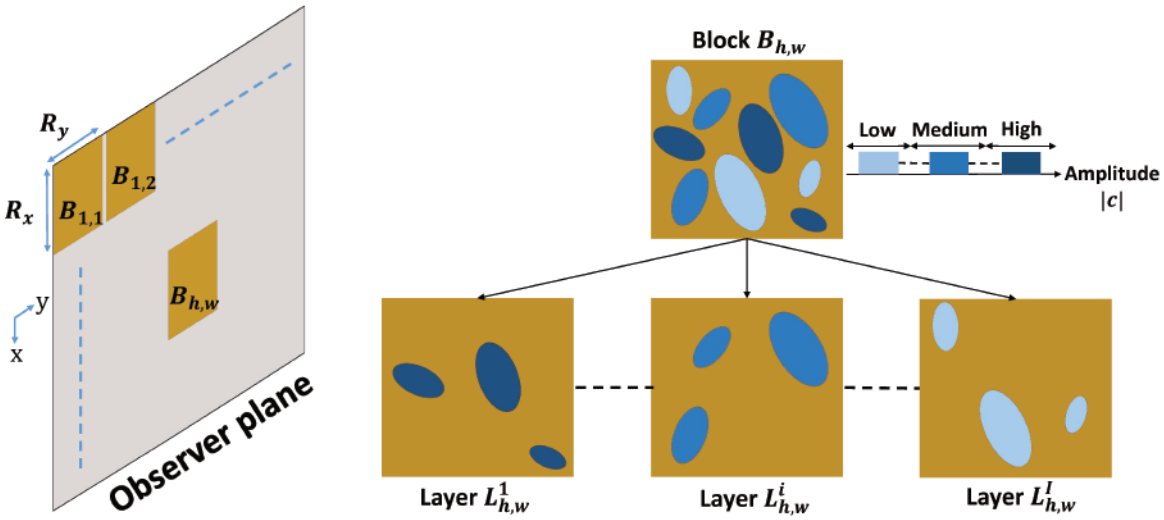

Fig. 7.1 Proposed server-client architecture for low-complexity holographic streaming: *the blue blocks represent the offline computation, whereas the green ones correspond to the online calculation.*

of each packet in a predefined space of the server's RAM memory. We note that all the aforementioned steps are done offline and only once by the server.

At the client side, the first step is to compute the user's position using an eye-tracking system. Based on this position, an adaptive selection module selects the blocks' and quality layers' indexes in a progressive order of importance, which are sent to the server via the back-channel. Correspondingly, the already generated bitstream is parsed and the relevant atoms' packets are sent on-the-fly to the client through the forward channel. An alternative to the parsing process is a direct access to the server's RAM memory. The decoding process starts by the dequantization and entropy decoding of the received packets. Then, the decoded atoms of each packet are weighted by their complex coefficients and summed to the hologram plane using a GPU-based implementation. Finally, the generated sub-hologram is visualized on the holographic display. We note that the decoding is done online and the reconstruction quality may be refined progressively as the atoms' packets are received.

### 7.1.2 Holographic encoder

**Viewpoint-quality scalable representation** To make visualization more fluid for end-users, we propose an offline encoding of the Gabor expansion obtained by MP. Indeed, at the server, the atoms are grouped into different packets in a way that may allow the client to have




(a) Tiling of the observer plane.

(b) Amplitude layers.

Fig. 7.2 Viewpoint-quality scalable representation of the Gabor expansion: *the observer plane is divided into different blocks, and each of them is split into several layers based on the atoms' amplitudes.*

a progressive view-based visualization with a low latency. To achieve this goal, we start by representing the atom set $A_K$ in the OP rather than the HP; *i.e.* every atom $a_k$ is represented by its diffraction pattern $P_k$ (*cf.* Section 5.2.2). Then, the OP is spatially divided into different blocks of resolution $R_x \times R_y$ (Figure 7.2a). A block is denoted by $B_{h,w}$, where $h$ and $w$ are the indexes along the $x$ and $y$ axis, respectively. The atoms assigned to a block $B_{h,w}$ are those having their diffraction patterns' centers inside this block. Therefore, the proposed representation provides a viewpoint scalability, *i.e.* a sub-hologram may be generated from the atoms corresponding to the blocks inscribed in the viewing window.

To enable scalability in terms of quality, it is important to know first the transmission order of the atom groups. To this end, the barycenter $G_{h,w}$ of the diffraction patterns is computed for each block. $G_{h,w}$ represents the centroid of spectral amplitude and is given by

$$G_{h,w} = \frac{\sum_{k=1}^{K^a_{h,w}} c_k C_k}{\sum_{k=1}^{N^a_{h,w}} c_k}, \tag{7.1}$$

where $K^a_{h,w}$ is the number of diffraction patterns in the block $B_{h,w}$. The importance of this measure to determine the most relevant atom groups will be emphasized in Section 7.1.3.

Moreover, sorting the atoms of each group according to their importance for the reconstruction is a key element for the quality scalability. Aiming at such functionality, we split the diffraction spectrums' patterns of each block $B_{h,w}$ into different layers $L^i_{h,w}$ based on a

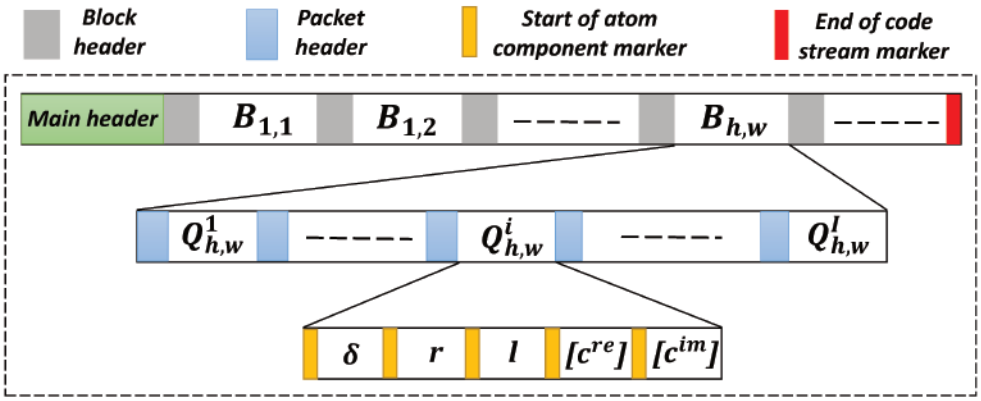


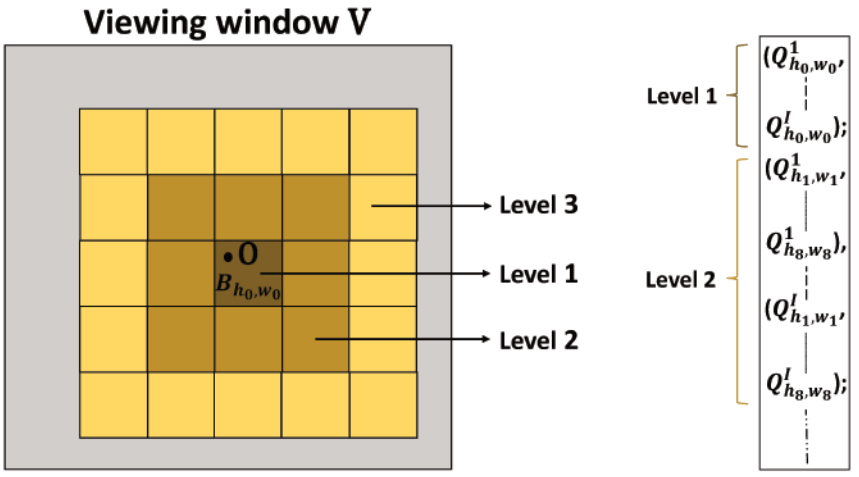

Fig. 7.3 Description of the bitstream structure.

threshold which take account of light intensity/amplitude (Figure 7.2b). Accordingly, for a given block $B_{h,w}$, we denote by $Q^i_{h,w}$ the packet of atoms having their amplitude between two thresholds $a_i$ and $a_{i+1}$.

**Bitstream generation** For transmission over communication channels, a compressed bitstream $\mathcal{B}$ is generated from the scalable representation given in the previous paragraph. Figure 7.3 illustrates the bitstream structure. First, the blocks are encoded in raster scan order. In each block $B_{h,w}$, the atoms' packets are encoded in a decreasing order of the amplitude threshold (i.e. from layer $L^1_{h,w}$ to $L^I_{h,w}$, where $I$ is the number of layers). Then, a partial bitstream is generated for each atom packet by encoding their atoms' components (i.e. the complex coefficients and the indexes). The real $c^{re}$ and imaginary $c^{im}$ of the atoms' coefficients are quantized using uniform scalar quantization. The atoms are sorted in a raster scan order and their positions are encoded differentially in $\delta$ (set of difference between consecutive atoms' positions). Finally, the atoms' components (quantized coefficients $[c^{re}]/[c^{im}]$, rotation indexes $r$, frequency dilation indexes $l$ and position indexes $\delta$) are losslessly encoded in the atoms positions' order using CABAC (*cf.* Section 4.3 for a detailed description of these encoding choices).

After each client's notification, the generated bitstream should be parsed to identify the blocks and packets corresponding to the requested sub-hologram. To allow this task, different marker headers are embedded in $\mathcal{B}$.

- *Main header*: is appended to the final bitstream and contain a marker segment that provides information about: start of bitstream, block's dimension, coding style (quantization levels, number of contexts used by the entropy coder).




(a) Block levels.

(b) Example of atom packets sorting.

Fig. 7.4 Sorting of blocks' and packets' indexes: *first, the blocks are sorted according to the distance between their barycenters and observer's center. The sorted blocks' are assigned to different levels, whitin each one the packets' are sorted in a decreasing amplitude order.*

- *Block header*: is appended to each block and provides the following information: beginning of the block part, number of layers, number of bytes encoding each atom packet.
- *Packet header*: is placed in front of every packet and contain a marker segment that identifies the start of the code stream for each atom's component.
- *End of code stream marker*: is added at the end of the bitstream to mark its termination.

The advantage of using a direct access memory to identify the partial bitstreams to be transmitted is to suppress the computational burden induced by parsing. However, this would require a higher storage capacity for the server's memory.

### 7.1.3 Adaptive selection

Let us consider a moving user who wants to visualize, at each instant, the encoded content on a holographic display. At the client side, the observation center $O$ is tracked after each movement. First, an adaptive selection module determines the blocks' and packets' indexes corresponding to the requested view. To enable progressive transmission, the selected indexes are sorted according to their importance for the reconstruction. As emphasized in Section 5.2.1, the most important criterion is the proximity of the emitted light beams to $O$.

This may be obtained by sorting the blocks' indexes in the increasing order of the distance $OG_{h,w}$ between the viewing center and the blocks' barycenters. The second priority criterion concerns the intensity of the diffracted light beams. To fill this criterion, the sorted blocks' are assigned to different levels (Figure 7.4a). Then, within each level, the packets' are sorted in a decreasing order of atoms amplitudes. An example of the sorted atoms' packets is illustrated by Figure 7.4b.

Once the adaptive selection and sorting are done, the client send the blocks' and packets' indexes to the server via the back-channel. Then, the bitstream $\mathcal{B}$ is parsed (or the server's RAM memory is directly browsed), and the partial bitstreams corresponding to the requested view are transmitted progressively to the client through the forward channel. The atoms' components of each received packet are entropy decoded and the target sub-hologram is generated and displayed. When an acceptable quality of reconstruction is reached, the adaptive selection module discards the packets' indexes corresponding to low-amplitude layers and selects instead those of the neighboring blocks with high amplitude. This would prevent from any viewpoint switching lags, caused by quick observer movements.

It is judicious to underline that the indexes selected for each observer's position must be stored by the client in a buffer. This is important to avoid multiple transmissions of packets common between overlapped views.

### 7.1.4 Holographic decoder

The holographic decoder is composed of two steps. First, the received atoms' packets are entropy decoded followed by the inverse quantization process. Then, the sub-hologram is computed from the decoded atoms of each packet. In the following, we proposed a GPU-based implementation in order to reduce the computational cost of the second step.

**Fast sub-hologram computation** The sub-hologram $\mathbf{H}^O$ corresponding to the observation center $O$ is generated from the decoded atoms using the approximating linear combination

$$\mathbf{H_O} = \sum_{h,w} \sum_{i} \mathbf{H}^O_{h,w,i} = \sum_{h,w} \sum_{i} \sum_{k=0}^{K^i_{h,w}} c_k g_{s[l_k],\theta[p_k]}, \tag{7.2}$$

where $\mathbf{H}^O_{h,w,i}$ denotes the sub-hologram generated from the atom packet $Q^i_{h,w}$, and $K^i_{h,w}$ is the number of atoms within this packet. Thus, the computational complexity of sub-hologram generation depends on $K^i_{h,w}$, the spatial resolution $(M_x, M_y)$ of the Gabor wavelets and also the distribution of the atoms' positions in the HP.

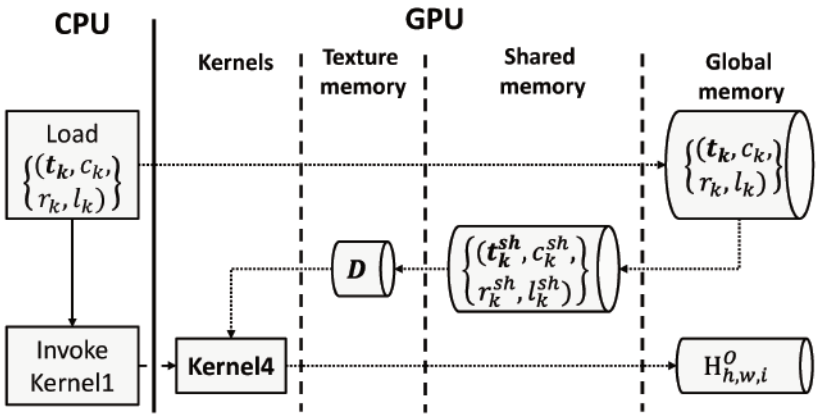



Fig. 7.5 Software structure of sub-hologram computation.

To reduce the decoder latency of our system, we propose a fast computation of sub-holograms using a GPU-based implementation in CUDA. Figure 7.5 shows the software structure for the computation of $\mathbf{H}^{O}_{h,w,i}$. First, the CPU loads the decoded atoms' components $\{\mathbf{t_k}, c_k, r_k, l_k\}_{0\leq k<K^i_{h,w}}$ into the GPU global memory ($\mathbf{t}$ is computed from $\delta$ using Equation 5.1). Then, CPU invokes kernel *Kernel4*, whose pseudo-code is given in Algorithm 3.

**Algorithm 3** *Kernel4* pseudo-code.

**Require:** $\{(\mathbf{t_k}, c_k, r_k, l_k)\}_{0\leq k<K^i_{h,w}}$
**Ensure:** $\mathbf{H}^{O}_{h,w,i}$
**for** $k \in \{0, .., K^i_{h,w} - 1\}$ in parallel **do**
  $(\mathbf{t^{sh}_k}, c^{sh}_k, r^{sh}_k, l^{sh}_k) \leftarrow (\mathbf{t_k}, c_k, r_k, l_k)$ ▷ Load $(\mathbf{t_k}, c_k, r_k, l_k)$ in shared memory
**end for**
Synchronize threads
**for all** $(x, y)$ in parallel **do**
  $r \leftarrow 0$ ▷ Initialize register value
  **for** $k \in \{0, .., K^i_{h,w} - 1\}$ **do**
    $X \leftarrow |x - t^{sh}_{x_k}|$
    $Y \leftarrow |y - t^{sh}_{y_k}|$
    $r \leftarrow r + c^{sh}_k \, g_{s[l^{sh}_k],\theta[r^{sh}_k]}(X, Y)$
  **end for**
  $\mathbf{H}^{O}_{h,w,i}(x, y) \leftarrow r$ ▷ Store output in global memory
**end for**

*Kernel4* computes the sub-hologram $\mathbf{H}^{O}_{h,w,i}$ corresponding to the atom packet $Q^i_{h,w}$. Since the computation of $\mathbf{H}^{O}_{h,w,i}$ involves one calculation per atom within the packet per sample of the output sub-hologram plane, all the threads executing *Kernel4* must access the set $\{(\mathbf{t_k}, c_k, r_k, l_k\}_{0\leq k<K^i_{h,w}}$ simultaneously. However, this data is stored in the global memory,

which has a long latency. Therefore, to reduce and optimize data transfers, *Kernel4* starts out by loading $\{(\mathbf{t_k}, c_k, r_k, l_k\}_{0\leq k<K^i_{h,w}}$ into the shared memory, which has a much shorter latency than global memory. Once this is done, threads must be synchronized to wait until all the data has been properly loaded into shared memory.

*Kernel4* then fetches **D** to compute the contribution of each Gabor atom within the packet $Q^i_{h,w}$. Since all the threads compute the contribution of the $k^{th}$ atom for different samples of the output plane simultaneously, consecutive threads are expected to fetch adjacent samples of **D**. Since the texture memory is optimized for 2D spatial locality memory accesses, The family of Gabor wavelets **D** is stored in this memory to achieve better performance.

Therefore, *Kernel4* fetches the Gabor-wavelets and performs a multiply-accumulate operation for each atom indexed by $k$. To limit the number of accesses to global memory, this multiply-accumulate operation is performed using an on-chip register, and the final result is stored only once into the global memory.

## 7.2 Experimental results

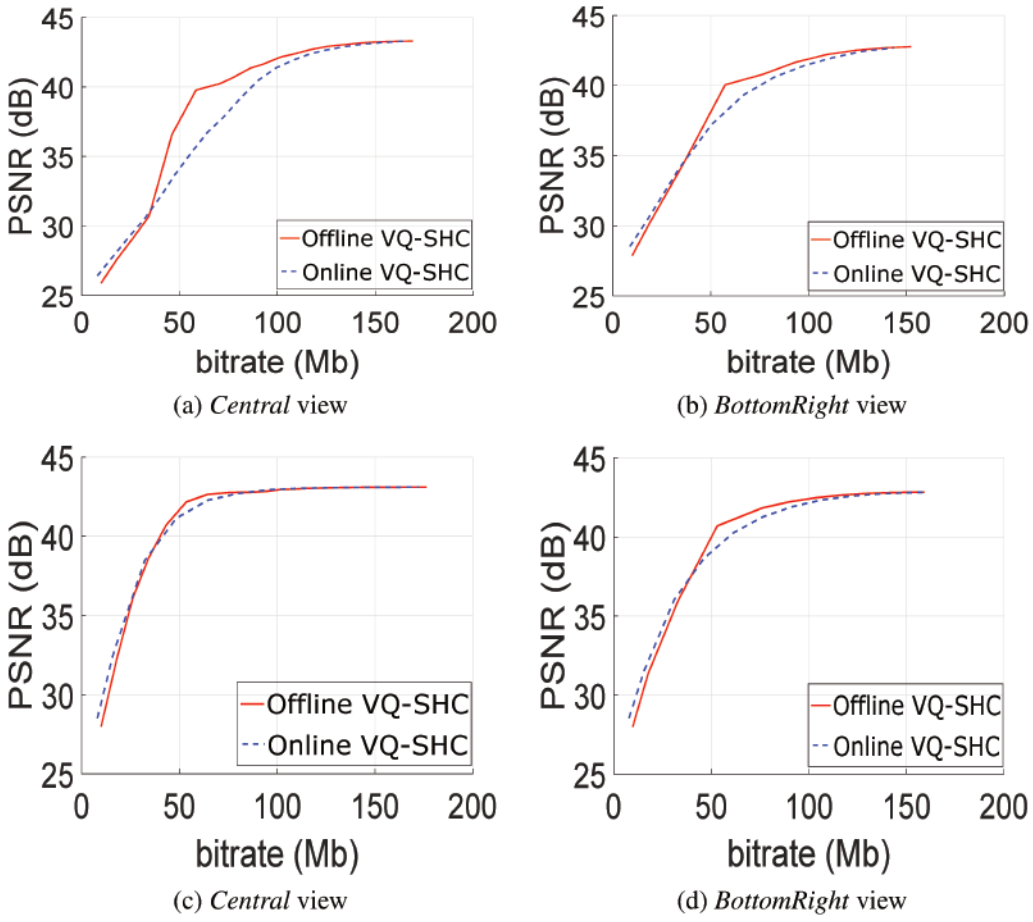
In this section, we evaluate the performance of the proposed streaming solution in terms of progressive compression and latency reduction. For notational purposes, our approach is denoted by *Offline VQ-SHC*, since the whole Gabor expansion is encoded offline using the VQ-SHC described in Section 7.1.2. We compare our method to the one we developed in the previous chapter, which will be denoted by *online VQ-SHC* since the atoms are encoded online after each client request using the VQ-SHC defined in Sections 5.1 and 5.2.

The main encoding choices for our designed low-complexity streaming framework based on the *Offline VQ-SHC* scheme are summarized below:

- The Gabor expansion $A_K$ is generated using VMP (*cf.* Section 4.2.1), where $K = \frac{N_x N_y}{4}$.

- The spatial resolution $(R_x, R_y)$ of the blocks $\{B_{h,w}\}$ is set to $(256, 256)$ and the number of amplitude layers I is set to $5$. A fine sampling of these parameters may enable a fine-grain scalability, however it would discard the correlations within atoms' positions and coefficients and then increase the overall encoding rate.

For the experiments, we use the *Piano8K* and *Dices8K* holograms previously described in Section 6.2.1. We consider two static users visualizing the hologram from two different viewpoints: *Central* and *BottomRight*, corresponding to the observation centers $O_1$ = (1.6mm, 1.6mm) and $O_2$ = (2.5mm, 2.5mm), respectively.




(a) *Central* view (b) *BottomRight* view

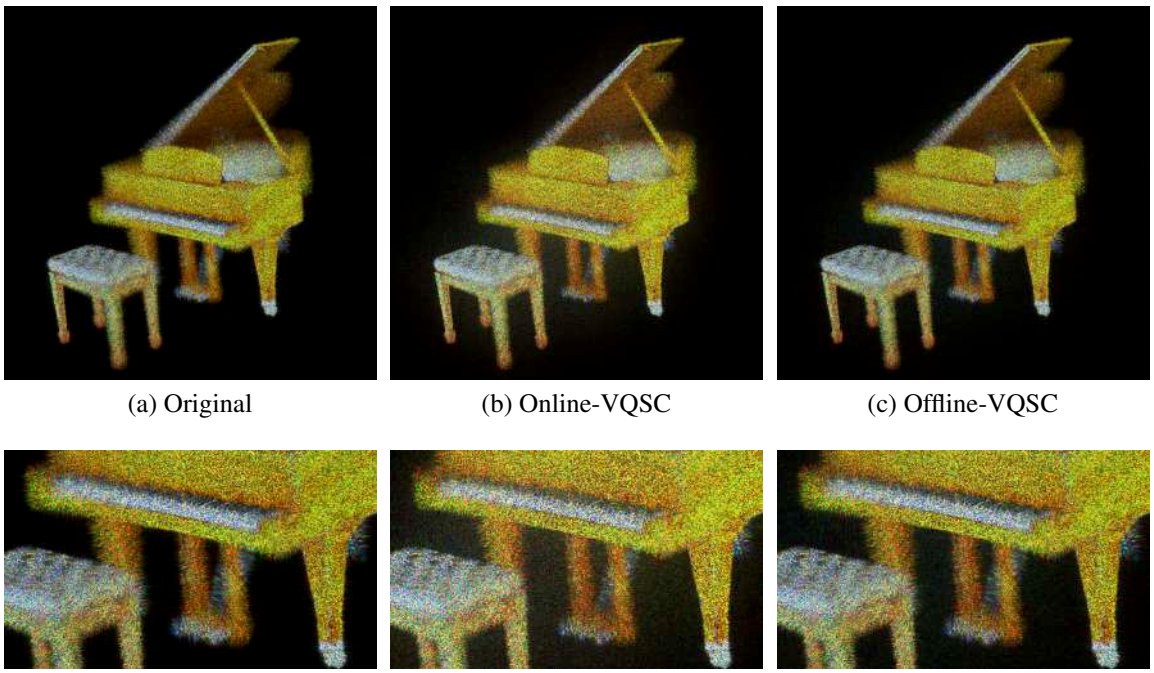
(c) *Central* view (d) *BottomRight* view

Fig. 7.6 RD curves obtained by *Online VQ-SHC* and *Offline VQ-SHC* (our approach) for the *Piano8K*(first row) and *Dices8K* (second row) holograms. The generated bitstream is progressively decoded and the PSNR is computed for the numerical reconstruction corresponding to each layer.

### 7.2.1 Compression performance

In the first part of experiments, the compression efficiency of the proposed approach is evaluated. Figure 7.6 illustrates the RD graphs obtained when streaming the *Piano8K* and *Dices8K* holograms, using *Online VQ-SHC* and *Offline VQ-SHC*.

As shown by the RD curves, the first decoded atoms (bitrate $\leq$ 35Mb) of the bitstream generated by *Online VQ-SHC* allow a slightly better reconstruction compared to those of the bitstream obtained from *Offline VQ-SHC*. This is explained by the fact that the fine-grain progressivity based on the distance criterion is discarded by the block-based representation of the proposed approach. The following atoms (bitrate $>$ 35Mb) decoded from the *Offline VQ-SHC* bitstream enable a better performance in terms of progressive compression. This



(a) Original (b) Online-VQSC (c) Offline-VQSC

Fig. 7.7 Numerical reconstructions the *Piano8K*'s *Central* view using: (a) original, (b) *Online VQ-SHC* and (c) *Offline VQ-SHC* (proposed approach). The considered bitrate is 60Mb.

is achieved thanks to the good trade-off provided by the proposed approach between the distance and amplitude criterion.

Figure 7.7 compares the numerical reconstructions of the *Piano8K* hologram obtained from *Online VQ-SHC* and *Offline VQ-SHC* to the original one. We consider a viewer placed at $O_1$ and a bitrate of 60Mb. The reconstructed view obtained from the proposed approach yields the best visual quality. Indeed, the front part of the piano (zoomed part) shows that the reconstructed image is less distorted for *Offline VQ-SHC* compared to *Online VQ-SHC*. Thus, the RD graphs results are confirmed with the visual ones.

### 7.2.2 Streaming latency evaluation

In the second part of experiments, the latency performance of the proposed streaming solution is evaluated. We denote by $L_{total}$ the total latency which corresponds to the delay occurring between the moment when the user's notification is sent to the sever and the moment when the requested view is displayed. For a view-dependent streaming system using our VQ-SHC schemes, $L_{total}$ is given by

$$L_{total} = L_{select} + L_{sort} + L_{encode} + L_{decode} + L_{subholo-generate}, \tag{7.3}$$

| **Latency** (ms) | **Method** | |
|---|---|---|
| | **Online VQ-SHC** | **Offline VQ-SHC** |
| $L_{encode}$ | 23.5 | 0 |
| $L_{decode}$ | 14.6 | 14.6 |
| $L_{subholo-generate}$ | 10.3 | 10.3 |

Table 7.1 Comparison of the streaming latency using *Online VQ-SHC* and *Offline VQ-SHC* (proposed method).

where $L_{select}$, $L_{sort}$, $L_{encode}$, $L_{decode}$ and $L_{subholo-generate}$ represent the latencies induced by the adaptive selection, sorting, encoding, decoding and sub-hologram generation, respectively.

According to the streaming architectures given in Figures 6.1 and 7.1, the proposed approach exhibits a lower latency for the adaptive selection and sorting module. Indeed, the *Offline VQ-SHC* scheme reduces the complexity of distance computation and sorting from $O(K)$ to $O(\frac{N_x \times N_y}{R_x \times R_y})$. Second, the latency of the encoding module (quantization + entropy coding) is null for the proposed method since all the Gabor expansion is encoded offline by the server. In contrary, the encoding latency of *Online VQ-SHC* may significantly increase the overall system latency in the case of high-resolution holograms or multi-users. At the client side, the latencies of entropy decoding and sub-hologram computation are the same for the offline and online methods because their modules are similars.

In the following, we compare only the latencies induced by the encoding, decoding and sub-hologram generation, since the computational burden of other process (adaptive selection, sorting and parsing/memory access) is negligible compared to these ones.

To evaluate these three latencies, we compute the execution time needed for atoms encoding, decoding and sub-hologram generation using the parallel computation described in Section 7.1.4. In the conducted test, we consider a packet of $7397$ atoms coressponding to the *Piano8K*'s sub-hologram for a *Central* view. The implementation is done in C++/CUDA on a PC equipped with an Intel Xeon e5-2620 v3 CPU operating at 2.40 GHz, a main memory of $32$GB, a GPU NVIDIA GeForce GTX 1070, under Microsoft Windows 10.

The latency values obtained in Table 7.1 show that the overall streaming latency is reduced by half approximately. This reduction may be more important in case of multi-viewers, since the number of atoms to encode would be higher. Moreover, we notice that the computational complexity of the decoder is $1.6$ times less than encoder.

By optimizing the entropy decoder implementation and considering a high-computational client device, the proposed approach may offer an efficient streaming of high-resolution holograms with a fluid visualization.

## 7.3 Conclusion

In this chapter, we described the design of a low-complexity progressive streaming framework to reduce the latency while transmitting large holograms. The novelty of the proposed server-client architecture consists on encoding the MP decomposition of holographic signals in an offline manner, independently of the observer's position. First, the server divides the observer plane into different blocks and the Gabor atoms are assigned to each block using the correspondence between Gabor wavelets and diffracted light beams. Then, each group of atoms is encoded in different packets according to their importance for the reconstruction. At the client side, the packets' indexes are determined after each user notification of the viewer's position. Accordingly, by applying a direct memory access or bitstream parsing, the relevant atoms' packets are transmitted progressively.

Compared to the streaming architecture developed in the previous chapter, where the atoms are encoded online during the transmission process, the experimental results reveal that our approach has better performance in terms of progressive compression and transmission latency. Thus, the proposed framework is a first step towards practical streaming of digital holograms with high resolutions.

# Part V

# Conclusion

# Chapter 8

# Conclusion

Among various emerging immersive media, digital holography is often considered as the most promising technology since it enables natural, faithful and comfortable 3D visualization to the naked eye. Unlike photographic images, digital holograms contain massive amount of data (Giga/Tera bytes), and exhibit irregular signal properties with low correlations. As a result, conventional image/video codecs deployed by today's communication networks may be unable to meet the bandwidth requirements for digital holographic data transmission in time compatible with interactive display systems.

As stated in Chapter 3, traditional image and video compression schemes heavily rely on the reduction of intra/inter-pixel block redundancies. Instead of this, the purpose of this thesis is to design specific coding schemes where the correlation between diffracted light beams is exploited. Such schemes may not lead to optimal compression gains when considering the hologram in its entirety, however they may provide a convenient way to interpret the 3D meaning of holographic signals in a view-dependent context. This light beam-based decomposition is the backbone of the contributions developed through this dissertation, which are composed of two main parts. The first part investigates the development of robust compression frameworks, which are able to reduce the overall size of digital holograms, and to allow a fine-grain scalable coding with respect to the directional degrees of freedom. The second part aims at designing efficient transmission architectures that may enable interactive 3D visualization, with a rapid and progressive increase of visual quality as well as a low latency.

To conclude this dissertation, a summarized compilation of the achieved work related to each chapter is presented in Section 8.1. Then, the potential improvements and future research directions related to our designed holographic codec are addressed in Section 8.2. Finally, the prospects and open issues raised by this thesis are discussed in Section 8.3.

## 8.1 Contributions summary

The first part of this work is devoted to digital holograms compression. To this end, we designed two novel coding schemes: one for full compression and the other for scalable compression in a view-dependent context.

Chapter 4 presented our first contribution, dedicated to digital holograms compression. Contrary to conventional pixel block-based decomposition, the introduced approach proposes to expand the input hologram into an overcomplete set of light beams locally emitted by its patterns. Such decomposition is enabled thanks to the duality existing between diffracted light beams and Gabor wavelets. To reduce the redundancy of the obtained Gabor expansion, Matching Pursuit algorithm is applied. For an effective compression, we studied the statistics and distribution of Gabor atoms components (coefficients and spatial-frequency indexes), and then designed a complete coding framework. The conducted experiments were performed on color digital holograms showing a better rate-distortion performance and visual reconstruction quality compared to classical image and video codecs (JPEG 2000 and intra-HEVC), especially for holograms with small pixel-pitches. The detailed description of this contribution has been published in the Optical Society's Applied Optics journal in 2018 [71], and received the *Editor's Pick* distinction.

Chapter 5 introduced our second contribution, which is devoted to scalable hologram compression in a view-dependent scenario. Based on the sparse Gabor expansion developed in the first contribution, we proposed a viewpoint-quality scalable hologram coding scheme by progressively encoding only the information corresponding to a given viewer (*i.e.* sub-hologram), instead of the whole hologram. The viewpoint scalability has been achieved thanks to the accurate space-frequency localization of Gabor wavelets, and consists on selecting only the Gabor atoms emitting light inside the observation window. The quality scalability is then obtained by sorting the set of selected atoms according to their importance for the reconstructed view. Experimental results revealed that the designed viewpoint scalable coding scheme enables significant compression gains compared to our previous contribution and state-of-the-art approaches. Moreover, the quality scalable scheme clearly outperforms SHVC (the scalable extension of HEVC) in terms of progressive increase of reconstruction quality. We presented the work and results related to the viewpoint scalability in the SPIE Photonics Europe, Digital Technologies for Imaging Applications conference in 2018 [72], and received the *Best Student Paper* award. One year after, we presented the outlines of the quality scalability at the IEEE International Conference on Image Processing (ICIP) [69].

The second part of this work is focused on digital hologram transmission in view-dependent scenarios. To this end, we designed and proposed two novel holographic stream-

ing architectures based on the joint viewpoint-quality scalable hologram coding scheme developed in the second contribution.

Chapter 6 exposed our third contribution, which aims at designing efficient progressive transmission solutions able to alleviate the bandwidth constraints when streaming high-quality holograms. The proposed streaming architecture is defined in a server-client environment using the aforementioned viewpoint-quality scalable hologram coding scheme. After each client notification about the observer's viewpoint and available network' bandwidth, the server selects, sorts and encodes Gabor atoms corresponding to the requested sub-hologram into different layers. In this way, the hologram is progressively streamed to the client following the user's trajectory. Hologram streaming has been simulated in concrete cases of moving user using a bandwidth corresponding to Wi-Fi connection. The comparison between our approach and conventional streaming solutions based on SHVC showed two main advantages of our method. The first one is that the content may be visualized immediately by decoding the first received atoms. The second strength of our proposed solution is to enable a rapid increase of reconstruction quality due to the coding efficiency of our scalable scheme and its fine-grain level of progressivity. This work was protected by a granted patent [68], and has been accepted for publication in The Springer's Annals of Telecommunications journal as an invited paper.

Chapter 7 presented our fourth contribution, which seeks to develop a low-complexity architecture for holographic streaming. Despite the effectiveness of the aforementioned streaming solution, it imposes an online encoding at the server which may increase the latency of the transmission chain. To overcome this hurdle, we propose a viewpoint-quality scalable coding scheme where the whole Gabor expansion is encoded offline. First, the observer plane is divided into spatial blocks. Then, the Gabor atoms are assigned to these blocks by exploiting the duality between Gabor wavelets and light beams. The atoms of each block are then classified into different layers according to their importance for the reconstruction and finally encoded in packets. At the decoder side, the atoms' packets are progressively decoded based on the viewer's position. Then, the corresponding sub-hologram is generated using parallel computations. Experimental results revealed that the designed approach reduces notably the streaming latency, without decreasing significantly the compression efficiency. We presented this work at the SPIE Optics + Photonics, Applications of Digital Image Processing conference in 2019 [70].

To summarize, we are convinced that considering light beam-based decomposition of digital holograms instead of conventional pixel block-based approach would play a major role for the design of future compression and transmission systems. In this thesis, we demonstrated the importance of Gabor wavelets and their excellent TF localization to interpret the optical

aspect related to the diffractive character of holograms from a signal processing point of view, providing then an efficient tool to analyze holographic patterns. Furthermore, the flexibility of light beam-based representation may enable various functionalities such as directional scalable coding and hologram edition.

## 8.2 Future works

As stated before, the work achieved in this thesis enables us to design a novel scalable codec specifically dedicated to digital holographic data, which enables efficient progressive transmission of high-resolution holograms with reduced latency. However, our contributions have some limitations which should be addressed in future works.

First of all, the optimization technique used by the proposed coding framework to reduce the redundancy of Gabor dictionary is based on Matching Pursuit. Theoretically, this technique does not provide the optimal representation in terms of sparsity. Thus, as possible improvements of our method, we propose to investigate the use of other optimization algorithms such as the Orthogonal Matching Pursuit [61, 154], or Basis Pursuit [39, 40] which minimizes the $L_1$-norm of coefficients occurring in the representation. In this way, we may obtain a more compact Gabor expansion of holographic signals leading to better compression gains. However, one has to verify that the entailed sparsity enhancement will not deteriorate the viewpoint scalable coding scheme performance.

Secondly, the proposed approach reached limited coding gains when considering full compression, especially for holograms with large pixel-pitches. This is due to the non-optimal energy compaction of Gabor wavelets transform. Alternatively, other TF analysis tools can be employed such as the Hilbert–Huang Transform (HHT) [110], in order to increase the overall size reduction of holograms. The fundamental part of HHT is the Empirical Mode Decomposition (EMD) method [73, 111, 162], which decomposes non-stationary signals into a complete basis of elementary components known as *Intrinsic Mode Functions*. Then, the TF representation is obtained by means of the Hilbert Spectral Analysis. The particularity of this transform is to adaptively decompose the analyzed signal according to its contained information instead of predefined transforms. Inspired by EMD and wavelet theory, an Empirical Wavelet Transform (EWT) paradigm has been introduced in [88]. Preliminary tests showed that EWT-based compression of digital holograms outperforms the Gabor wavelets-based one in some cases, when considering the hologram in its entirety. As a future work, EWT-based compression can be extended to the view-dependent context.

Last but not least, our proposed codec has been designed and evaluated only for still digital holograms. At the time being, very few work addressed holographic video compression [27,

53, 143, 171, 173]. As a future work, our coding framework could be extended to holographic video compression. Indeed, starting from the Gabor expansion of the first hologram frame, the TF domain may be segmented by regrouping light beams with similar direction of diffraction and color, where each group corresponds to a given portion of the 3D scene. Then, by approximating a rigid transformation of each portion, a linear transformation may be obtained and then applied on the atoms' indexes of the reference frame to predict the next one. Following this process, an efficient motion estimation/compensation mechanism can be built.

## 8.3 Prospects and open issues

In addition to the compression and transmission issues we tackle in this thesis, other hardware and software challenges remain unsolved [21, 167] and need in-depth studies.

The hardware issues are mainly related to the design of satisfactory digital holographic displays compatible with large full-color holograms. Indeed, the limited resolutions and pixel-pitches of current SLMs do not allow the display of life-size human holograms with wide FoV. This limitation may be tackled in AR and VR applications by using holographic HMDs: since the scene is restituted for only one viewpoint based on head movements, only a subset of the hologram must be displayed, thus, the small size of SLMs is no longer an issue. Moreover, today's holographic display technologies do not enable a proper subjective quality evaluation of holograms. Alternatively, the quality of the reconstructed compressed hologram is assessed on regular 2D displays [4].

One of the most important software open issues that needs further investigation is the lack of objective metrics for quality assesment of holographic content. Indeed, compressed digital holograms are evaluated either on the hologram plane or on the reconstruction one by using conventional image quality metrics such as PSNR, MSE and Structural Similarity Index (SSIM). However, such metrics are poorly correlated with the perceptual visual quality are not faithful because of the irregular signal properties of holographic patterns and the speckle noise induced in their optical and numerical reconstructions. Several techniques have been proposed in order to reduce the speckle effect and then enhance quality [17, 139, 148, 163, 185]. Moreover, some recent researches highlighted the dedicated design of new quality metrics [1–3], making a first step towards standard metrics specifically to holographic content.

Finally, digital holographic databases need to be enriched with more full-color still and dynamic holograms recorded from complex real scenes with large resolutions in order to enable better evaluation of holographic coding engines.

# Part VI

# References and Appendices

# References


[1] A. Ahar, A. Barri, and P. Schelkens. From sparse coding significance to perceptual quality: A new approach for image quality assessment. *IEEE Transactions on Image Processing*, 27(2):879–893, 2017.

[2] A. Ahar, T. Birnbaum, D. Blinder, A. Symeonidou, and P. Schelkens. Performance evaluation of sparseness significance ranking measure (ssrm) on holographic content. In *Laser Applications to Chemical, Security and Environmental Analysis*, pages JTu4A–10. Optical Society of America, 2018.

[3] A. Ahar, T. Birnbaum, C. Jaeh, and P. Schelkens. A new similarity measure for complex valued data. In *Digital Holography and Three-Dimensional Imaging*, pages Tu1A–6. Optical Society of America, 2017.

[4] A. Ahar, D. Blinder, T. Bruylants, C. Schretter, A. Munteanu, and P. Schelkens. Subjective quality assessment of numerically reconstructed compressed holograms. In *Applications of Digital Image Processing XXXVIII*, volume 9599, page 95990K. International Society for Optics and Photonics, 2015.

[5] N. Ahmed, T. Natarajan, and K. R. Rao. Discrete cosine transform. *IEEE transactions on Computers*, 100(1):90–93, 1974.

[6] J. Allen. Short term spectral analysis, synthesis, and modification by discrete fourier transform. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 25(3):235–238, 1977.

[7] T. Ando, K. Yamasaki, M. Okamoto, and E. Shimizu. Head-mounted display using a holographic optical element. In *Practical Holography XII*, volume 3293, pages 183–189. International Society for Optics and Photonics, 1998.

[8] T. André, M. Cagnazzo, M. Antonini, and M. Barlaud. Jpeg2000-compatible scalable scheme for wavelet-based video coding. *Journal on Image and Video Processing*, 2007(1):9–9, 2007.

[9] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies. Image coding using wavelet transform. *IEEE Transactions on image processing*, 1(2):205–220, 1992.

[10] D. Austerberry. *The technology of video and audio streaming*. Routledge, 2013.

[11] L. T. Bang, Z. Ali, P. D. Quang, J.-H. Park, and N. Kim. Compression of digital hologram for three-dimensional object using wavelet-bandelets transform. *Optics express*, 19(9):8019–8031, 2011.

[12] A. Batalla. *Multiple transforms for video coding*. PhD thesis, December 2015.

[13] F. Bergeaud and S. Mallat. Matching pursuit of images. In *Proceedings., International Conference on Image Processing*, volume 1, pages 53–56. IEEE, 1995.

[14] M. V. Bernardo, P. Fernandes, A. Arrifano, M. Antonini, E. Fonseca, P. T. Fiadeiro, A. M. Pinheiro, and M. Pereira. Holographic representation: Hologram plane vs. object plane. *Signal Processing: Image Communication*, 68:193–206, 2018.

[15] M. V. Bernardo, E. Fonseca, P. Fiadeiro, A. M. Pinheiro, and M. Pereira. A digital hologram compression scheme for representation on the object plane. In *Applications of Digital Image Processing XLI*, volume 10752, page 107520J. International Society for Optics and Photonics, 2018.

[16] V. D. Bhamidipati and S. Kilari. Effect of delay/delay variation on qoe in video streaming, 2010.

[17] V. Bianco, P. Memmolo, M. Leo, S. Montresor, C. Distante, M. Paturzo, P. Picart, B. Javidi, and P. Ferraro. Strategies for reducing speckle noise in digital holography. *Light: Science & Applications*, 7(1):48, 2018.

[18] T. Birnbaum, A. Ahar, D. Blinder, C. Schretter, T. Kozacki, and P. Schelkens. Wave atoms for digital hologram compression. *Applied optics*, 58(22):6193–6203, 2019.

[19] T. Birnbaum, A. Ahar, D. Blinder, C. Schretter, T. Kozacki, and P. Schelkens. Wave atoms for lossy compression of digital holograms. In *2019 Data Compression Conference (DCC)*, pages 398–407. IEEE, 2019.

[20] D. Blinder. *Efficient Representation, Generation and Compression of Digital Holograms*. PhD thesis, 2018.

[21] D. Blinder, A. Ahar, S. Bettens, T. Birnbaum, A. Symeonidou, H. Ottevaere, C. Schretter, and P. Schelkens. Signal processing challenges for digital holographic video display systems. *Signal Processing: Image Communication*, 70:114–130, 2019.

[22] D. Blinder, A. Ahar, A. Symeonidou, Y. Xing, T. Bruylants, C. Schreites, B. Pesquet-Popescu, F. Dufaux, A. Munteanu, and P. Schelkens. Open access database for experimental validations of holographic compression engines. In *2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*, pages 1–6. IEEE, 2015.

[23] D. Blinder, T. Bruylants, H. Ottevaere, A. Munteanu, and P. Schelkens. Jpeg 2000-based compression of fringe patterns for digital holographic microscopy. *Optical Engineering*, 53(12):123102, 2014.

[24] D. Blinder, T. Bruylants, E. Stijns, H. Ottevaere, and P. Schelkens. Wavelet coding of off-axis holographic images. In *Applications of Digital Image Processing XXXVI*, volume 8856, page 88561L. International Society for Optics and Photonics, 2013.

[25] D. Blinder and P. Schelkens. Integer fresnel transform for lossless hologram compression. In *2019 Data Compression Conference (DCC)*, pages 389–397. IEEE, 2019.

[26] D. Blinder, C. Schretter, H. Ottevaere, A. Munteanu, and P. Schelkens. Unitary transforms using time-frequency warping for digital holograms of deep scenes. *IEEE Transactions on Computational Imaging*, 4(2):206–218, 2018.

[27] D. Blinder, C. Schretter, and P. Schelkens. Global motion compensation for compressing holographic videos. *Optics express*, 26(20):25524–25533, 2018.

[28] B. G. Blundell and A. J. Schwarz. The classification of volumetric display systems: characteristics and predictability of the image space. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):66–75, 2002.

[29] J. M. Boyce, Y. Ye, J. Chen, and A. K. Ramasubramonian. Overview of shvc: Scalable extensions of the high efficiency video coding standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 26(1):20–34, 2015.

[30] G. B. Brandt. Image plane holography. *Applied optics*, 8(7):1421–1429, 1969.

[31] M. Burrows and D. J. Wheeler. A block-sorting lossless data compression algorithm. 1994.

[32] M. Cagnazzo, T. André, M. Antonini, and M. Barlaud. A model-based motion compensated video coder with jpeg2000 compatibility. In *2004 International Conference on Image Processing, 2004. ICIP'04.*, volume 4, pages 2255–2258. IEEE, 2004.

[33] M. Cagnazzo, M. Antonini, and M. Barlaud. Mutual information-based context quantization. *Signal Processing: Image Communication*, 25(1):64–74, 2010.

[34] A. R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo. Lossless image compression using integer to integer wavelet transforms. In *Proceedings of International Conference on Image Processing*, volume 1, pages 596–599. IEEE, 1997.

[35] J. Chakareski, R. Aksu, X. Corbillon, G. Simon, and V. Swaminathan. Viewport-driven rate-distortion optimized 360º video streaming. In *2018 IEEE International Conference on Communications (ICC)*, pages 1–7. IEEE, 2018.

[36] J. Chakareski, S. Han, and B. Girod. Layered coding vs. multiple descriptions for video streaming over multiple paths. *Multimedia Systems*, 10(4):275–285, 2005.

[37] N. Chamakhi, I. Bouzidi, A. O. Zaid, and F. Dufaux. Jpeg based compression of digital holograms. In *2018 7th European Workshop on Visual Information Processing (EUVIP)*, pages 1–6. IEEE, 2018.

[38] C.-L. Chang and B. Girod. Direction-adaptive discrete wavelet transform for image compression. *IEEE Transactions on Image Processing*, 16(5):1289–1302, 2007.

[39] S. Chen and D. Donoho. Basis pursuit. In *Proceedings of 1994 28th Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 41–44. IEEE, 1994.

[40] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM review*, 43(1):129–159, 2001.

[41] P. Cheremkhin and E. Kurbatova. Wavelet compression of off-axis digital holograms using real/imaginary and amplitude/phase parts. *Scientific reports*, 9(1):7561, 2019.

[42] P. A. Cheremkhin and E. A. Kurbatova. Numerical comparison of scalar and vector methods of digital hologram compression. In *Holography, Diffractive Optics, and Applications VII*, volume 10022, page 1002227. International Society for Optics and Photonics, 2016.

[43] P. A. Cheremkhin and E. A. Kurbatova. Compression of digital holograms using 1-level wavelet transforms, thresholding and quantization of wavelet coefficients. In *Digital Holography and Three-Dimensional Imaging*, pages W2A–38. Optical Society of America, 2017.

[44] P. A. Cheremkhin and E. A. Kurbatova. Quality of reconstruction of compressed off-axis digital holograms by frequency filtering and wavelets. *Applied optics*, 57(1):A55–A64, 2018.

[45] T.-L. Chin, T.-Y. Chen, C.-C. Huang, and T.-R. Hsiang. Scalable video streaming for multicast in wireless networks. In *2015 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, pages 182–187. IEEE, 2015.

[46] C. Christopoulos, A. Skodras, and T. Ebrahimi. The jpeg2000 still image coding system: an overview. *IEEE transactions on consumer electronics*, 46(4):1103–1127, 2000.

[47] T. Claasen and W. Mecklenbrauker. The wigner distribution—a tool for time-frequency signal analysis. *Philips J. Res*, 35(3):217–250, 1980.

[48] L. Cohen. *Time-frequency analysis*, volume 778. Prentice hall, 1995.

[49] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of computation*, 19(90):297–301, 1965.

[50] X. Corbillon, G. Simon, A. Devlic, and J. Chakareski. Viewport-adaptive navigable 360-degree video delivery. In *Communications (ICC), 2017 IEEE International Conference on*, pages 1–7. IEEE, 2017.

[51] R. Corda, A. Gilles, K.-J. Oh, A. Pinheiro, P. Schelkens, and C. Perra. An exploratory study towards objective quality evaluation of digital hologram coding tools. In *Applications of Digital Image Processing XLII*, volume 11137, page 111371D. International Society for Optics and Photonics, 2019.

[52] E. Darakis, M. Kowiel, R. Näsänen, and T. J. Naughton. Visually lossless compression of digital hologram sequences. In *Image Quality and System Performance VII*, volume 7529, page 752912. International Society for Optics and Photonics, 2010.

[53] E. Darakis and T. J. Naughton. Compression of digital hologram sequences using mpeg-4. In *Holography: Advances and Modern Trends*, volume 7358, page 735811. International Society for Optics and Photonics, 2009.

[54] E. Darakis, T. J. Naughton, and J. J. Soraghan. Compression defects in different reconstructions from phase-shifting digital holographic data. *Applied optics*, 46(21):4579–4586, 2007.

[55] E. Darakis and J. J. Soraghan. Compression of interference patterns with application to phase-shifting digital holography. *Applied optics*, 45(11):2437–2443, 2006.

[56] E. Darakis and J. J. Soraghan. Use of fresnelets for phase-shifting digital hologram compression. *IEEE transactions on image processing*, 15(12):3804–3811, 2006.

[57] E. Darakis and J. J. Soraghan. Reconstruction domain compression of phase-shifting digital holograms. *Applied Optics*, 46(3):351–356, 2007.

[58] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE transactions on information theory*, 36(5):961–1005, 1990.

[59] I. Daubechies. *Ten lectures on wavelets*, volume 61. Siam, 1992.

[60] G. Davis, S. Mallat, and M. Avellaneda. Adaptive greedy approximations. *Constructive approximation*, 13(1):57–98, 1997.

[61] G. M. Davis, S. G. Mallat, and Z. Zhang. Adaptive time-frequency decompositions with matching pursuit. In *Wavelet Applications*, volume 2242, pages 402–413. International Society for Optics and Photonics, 1994.

[62] A. H. Dekker. Kohonen neural networks for optimal colour quantization. *Network: Computation in Neural Systems*, 5(3):351–367, 1994.

[63] L. Demanet and L. Ying. Wave atoms and sparsity of oscillatory patterns. *Applied and Computational Harmonic Analysis*, 23(3):368–387, 2007.

[64] S. Deshpande and W. Zeng. Scalable streaming of jpeg2000 images using hypertext transfer protocol. In *Proceedings of the ninth ACM international conference on Multimedia*, pages 372–381. ACM, 2001.

[65] S. G. Deshpande and R. T. Thomas. Methods and systems for scalable streaming of images with client-side control, Aug. 21 2007. US Patent 7,260,614.

[66] N. A. Dodgson. Autostereoscopic 3d displays. *Computer*, 38(8):31–36, 2005.

[67] F. Dufaux, Y. Xing, B. Pesquet-Popescu, and P. Schelkens. Compression of digital holographic data: an overview. In *Applications of Digital Image Processing XXXVIII*, volume 9599, page 95990I. International Society for Optics and Photonics, 2015.

[68] A. El Rhammad, P. Gioia, A. Gilles, and M. Cagnazzo. Procédé et dispositif de transmission de données représentatives d'un hologramme numérique. Number 20190130_CD393FR, 2019.

[69] A. El Rhammad, P. Gioia, A. Gilles, and M. Cagnazzo. Scalable coding framework for a view-dependent streaming of digital holograms. In *2019 IEEE International Conference on Image Processing (ICIP)*, pages 146–150. IEEE, 2019.

[70] A. El Rhammad, P. Gioia, A. Gilles, and M. Cagnazzo. Towards practical hologram streaming using progressive coding. In *Applications of Digital Image Processing XLII*, volume 11137, page 111371E. International Society for Optics and Photonics, 2019.

[71] A. El Rhammad, P. Gioia, A. Gilles, M. Cagnazzo, and B. Pesquet-Popescu. Color digital hologram compression based on matching pursuit. *Applied optics*, 57(17):4930–4942, 2018.

[72] A. El Rhammad, P. Gioia, A. Gilles, M. Cagnazzo, and B. Pesquet-Popescu. View-dependent compression of digital hologram based on matching pursuit. In *Optics, Photonics, and Digital Technologies for Imaging Applications V*, volume 10679, page 106790L. International Society for Optics and Photonics, 2018.

[73] P. Flandrin, G. Rilling, and P. Goncalves. Empirical mode decomposition as a filter bank. *IEEE signal processing letters*, 11(2):112–114, 2004.

[74] J. M. Florence. Spatial light modulator with full complex light modulation capability, Sept. 15 1992. US Patent 5,148,157.

[75] A. Ford and A. Roberts. Colour space conversions. *Westminster University, London*, 1998:1–31, 1998.

[76] W. Freude and G. K. Grau. Rayleigh-sommerfeld and helmholtz-kirchhoff integrals: application to the scalar and vectorial theory of wave propagation and diffraction. *Journal of lightwave technology*, 13(1):24–32, 1995.

[77] D. Gabor. Theory of communication. part 1: The analysis of information. *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, 93(26):429–441, 1946.

[78] D. Gabor. A new microscopic principle, 1948.

[79] Q. Gao, J. Liu, X. Duan, T. Zhao, X. Li, and P. Liu. Compact see-through 3d head-mounted display based on wavefront modulation with holographic grating filter. *Optics express*, 25(7):8412–8424, 2017.

[80] M. R. Garey and D. S. Johnson. *Computers and intractability*, volume 29. wh freeman New York, 2002.

[81] V. Garg. *Wireless communications & networking*. Elsevier, 2010.

[82] D. Ge, X. Jiang, and Y. Ye. A note on the complexity of l p minimization. *Mathematical programming*, 129(2):285–299, 2011.

[83] R. Gerchberg and A. Saxton W. O. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35:237–250, 11 1971.

[84] A. Gersho and R. M. Gray. *Vector quantization and signal compression*, volume 159. Springer Science & Business Media, 2012.

[85] A. Gilles. *Fast hologram synthesis methods for realistic 3D visualization*. PhD thesis, Rennes, INSA, 2016.

[86] A. Gilles, P. Gioia, R. Cozot, and L. Morin. Computer generated hologram from multiview-plus-depth data considering specular reflections. In *2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, pages 1–6. IEEE, 2016.

[87] A. Gilles, P. Gioia, R. Cozot, and L. Morin. Hybrid approach for fast occlusion processing in computer-generated hologram calculation. *Applied optics*, 55(20):5459–5470, 2016.

[88] J. Gilles. Empirical wavelet transform. *IEEE transactions on signal processing*, 61(16):3999–4010, 2013.

[89] E. B. Goldstein. *Sensation and perception*. Cengage Learning, 2009.

[90] S. Golomb. Run-length encodings (corresp.). *IEEE transactions on information theory*, 12(3):399–401, 1966.

[91] J. W. Goodman. Foundations of scalar diffraction theory. *Introduction to Fourier Optics*, pages 31–62, 1968.

[92] J. W. Goodman. *Introduction to Fourier optics*. Roberts and Company Publishers, 2005.

[93] J. W. Goodman. *Speckle phenomena in optics: theory and applications*. Roberts and Company Publishers, 2007.

[94] J. W. Goodman and R. Lawrence. Digital image formation from electronically detected holograms. *Applied physics letters*, 11(3):77–79, 1967.

[95] V. K. Goyal. Multiple description coding: Compression meets the network. *IEEE Signal processing magazine*, 18(5):74–93, 2001.

[96] V. K. Goyal, J. Kovacevic, R. Arean, and M. Vetterli. Multiple description transform coding of images. In *International conference on image processing*, volume 1, pages 674–678, 1998.

[97] C. Greco, I. D. Nemoianu, M. Cagnazzo, and B. Pesquet-Popescu. A network coding scheduling for multiple description video streaming over wireless networks. In *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, pages 1915–1919. IEEE, 2012.

[98] T. Guionnet, C. Guillemot, and S. Pateux. Embedded multiple description coding for progressive image transmission over unreliable channels. In *Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205)*, volume 1, pages 94–97. IEEE, 2001.

[99] T. Haist and W. Osten. Holography using pixelated spatial light modulators—part 1: theory and basic considerations. *Journal of Micro/Nanolithography, MEMS, and MOEMS*, 14(4):041310, 2015.

[100] P. Hariharan and P. Hariharan. *Optical Holography: Principles, techniques and applications*. Cambridge University Press, 1996.

[101] M. Harris et al. Optimizing parallel reduction in cuda. *Nvidia developer technology*, 2(4):70, 2007.

[102] J. Hattay. *Wavelet-based lifting structures and blind source separation: applications to digital in-line holography*. PhD thesis, Rouen, 2016.

[103] R. Häussler, S. Reichelt, N. Leister, E. Zschau, R. Missbach, and A. Schwerdtner. Large real-time holographic displays: from prototypes to a consumer product. In *Stereoscopic Displays and Applications XX*, volume 7237, page 72370S. International Society for Optics and Photonics, 2009.

[104] R. Häussler, A. Schwerdtner, and N. Leister. Large holographic displays as an alternative to stereoscopic displays. In *Stereoscopic Displays and Applications XIX*, volume 6803, page 68030M. International Society for Optics and Photonics, 2008.

[105] W. Heisenberg. Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. In *Original Scientific Papers Wissenschaftliche Originalarbeiten*, pages 478–504. Springer, 1985.

[106] L. Hesselink, S. S. Orlov, and M. C. Bashaw. Holographic data storage systems. *Proceedings of the IEEE*, 92(8):1231–1280, 2004.

[107] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision*, 8(3):33–33, 2008.

[108] H. Hoshino, F. Okano, H. Isono, and I. Yuyama. Analysis of resolution limitation of integral photography. *JOSA A*, 15(8):2059–2065, 1998.

[109] M. Hosseini and V. Swaminathan. Adaptive 360 vr video streaming: Divide and conquer. In *Multimedia (ISM), 2016 IEEE International Symposium on*, pages 107–110. IEEE, 2016.

[110] N. E. Huang. *Hilbert-Huang transform and its applications*, volume 16. World Scientific, 2014.

[111] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971):903–995, 1998.

[112] D. A. Huffman. A method for the construction of minimum-redundancy codes. *Proceedings of the IRE*, 40(9):1098–1101, 1952.

[113] R. F. i Ventura, O. D. Escoda, and P. Vandergheynst. A matching pursuit full search algorithm for image approximations. *ITS Technical Report*, 2004.

[114] T. Ichikawa, T. Yoneyama, and Y. Sakamoto. Cgh calculation with the ray tracing method for the fourier transform optical system. *Optics express*, 21(26):32019–32031, 2013.

[115] S. Igarashi, K. Kakinuma, T. Nakamura, K. Ikeya, J. Arai, T. Mishina, K. Matsushima, and M. Yamaguchi. Computer-generated holograms of a life-size human captured from multi-viewpoint cameras. In *Digital Holography and Three-Dimensional Imaging*, pages Tu4A–4. Optical Society of America, 2019.

[116] B. Javidi and T. Nomura. Securing information by use of digital holography. *Optics letters*, 25(1):28–30, 2000.

[117] S. Jiao, Z. Jin, C. Chang, C. Zhou, W. Zou, and X. Li. Compression of phase-only holograms with jpeg standard and deep learning. *Applied Sciences*, 8(8):1258, 2018.

[118] R. K. Jones and D. N. Lee. Why two eyes are better than one: the two views of binocular vision. *Journal of Experimental Psychology: Human Perception and Performance*, 7(1):30, 1981.

[119] P. A. Kara, A. Cserkaszky, A. Barsi, M. G. Martini, and T. Balogh. Towards adaptive light field video streaming. *COMSOC MMTC Communications-Frontiers*, 2017.

[120] M. K. Kim. Principles and techniques of digital holographic microscopy. *SPIE reviews*, 1(1):018005, 2010.

[121] T. Kim. Scalable video streaming over the internet. 01 2005.

[122] P. T. Kovács, A. Zare, T. Balogh, R. Bregović, and A. Gotchev. Architectures and codecs for real-time light field streaming. *Journal of Imaging Science and Technology*, 61(1):10403–1, 2017.

[123] E. Kurbatova, P. Cheremkhin, N. Evtikhiev, V. Krasnov, and S. Starikov. Methods of compression of digital holograms. *Physics Procedia*, 73:328–332, 2015.

[124] D. Le Gall and A. Tabatabai. Sub-band coding of digital images using symmetric short kernel filters and arithmetic coding techniques. In *ICASSP-88., International Conference on Acoustics, Speech, and Signal Processing*, pages 761–764. IEEE, 1988.

[125] E. Le Pennec and S. Mallat. Bandelet image approximation and compression. *Multiscale Modeling & Simulation*, 4(3):992–1039, 2005.

[126] E. Le Pennec and S. Mallat. Sparse geometric image representations with bandelets. *IEEE transactions on image processing*, 14(4):423–438, 2005.

[127] S. Lederer, C. Müller, and C. Timmerer. Dynamic adaptive streaming over http dataset. In *Proceedings of the 3rd Multimedia Systems Conference*, pages 89–94. ACM, 2012.

[128] D.-H. Lee, J.-Y. Sim, C.-S. Kim, and S.-U. Lee. Viewing angle dependent coding of digital holograms. In *2011 19th European Signal Processing Conference*, pages 1367–1371. IEEE, 2011.

[129] E. N. Leith and J. Upatnieks. Reconstructed wavefronts and communication theory. *JOSA*, 52(10):1123–1130, 1962.

[130] D. Leseberg and C. Frère. Computer-generated holograms of 3-d objects composed of tilted planar segments. *Applied optics*, 27(14):3020–3024, 1988.

[131] M. Liebling, T. Blu, and M. Unser. Fresnelets: new multiresolution wavelet bases for digital holography. *IEEE Transactions on image processing*, 12(1):29–43, 2003.

[132] Y. Linde, A. Buzo, and R. Gray. An algorithm for vector quantizer design. *IEEE Transactions on communications*, 28(1):84–95, 1980.

[133] G. Lippmann. Epreuves reversibles donnant la sensation du relief. *J. Phys. Theor. Appl.*, 7(1):821–825, 1908.

[134] S. Lloyd. Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2):129–137, 1982.

[135] M. Lucente. The first 20 years of holographic video—and the next 20. In *SMPTE 2nd Annual International Conference on Stereoscopic 3D for Media and Entertainment*, pages 21–23, 2011.

[136] R. Maciol, Y. Yuan, and I. T. Nabney. Colour image coding with matching pursuit in the spatio-frequency domain. In *International Conference on Image Analysis and Processing*, pages 306–317. Springer, 2011.

[137] A. Maimone, A. Georgiou, and J. S. Kollin. Holographic near-eye displays for virtual and augmented reality. *ACM Transactions on Graphics (TOG)*, 36(4):85, 2017.

[138] S. G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on signal processing*, 41(12):3397–3415, 1993.

[139] J. Maycock, B. M. Hennelly, J. B. McDonald, Y. Frauel, A. Castro, B. Javidi, and T. J. Naughton. Reduction of speckle in digital holography by discrete fourier filtering. *JOSA A*, 24(6):1617–1622, 2007.

[140] G. A. Mills and I. Yamaguchi. Effects of quantization in phase-shifting digital holography. *Applied Optics*, 44(7):1216–1225, 2005.

[141] T. Mishina, M. Okui, and F. Okano. Calculation of holograms from elemental images captured by integral photography. *Applied optics*, 45(17):4026–4036, 2006.

[142] M. Moshinsky and C. Quesne. Linear canonical transformations and their unitary representations. *Journal of Mathematical Physics*, 12(8):1772–1780, 1971.

[143] R. K. Muhamad, D. Blinder, A. Symeonidou, T. Birnbaum, O. Watanabe, C. Schretter, and P. Schelkens. Exact global motion compensation for holographic video compression. *Applied Optics*, 58(34):G204–G217, 2019.

[144] S. Murata and N. Yasuda. Potential of digital holography in particle measurement. *Optics & Laser Technology*, 32(7-8):567–574, 2000.

[145] T. J. Naughton, Y. Frauel, B. Javidi, and E. Tajahuerce. Compression of digital holograms for three-dimensional object reconstruction and recognition. *Applied optics*, 41(20):4124–4132, 2002.

[146] D. Needell and J. A. Tropp. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Applied and computational harmonic analysis*, 26(3):301–321, 2009.

[147] H. Nishi, K. Matsushima, and S. Nakahara. Rendering of specular surfaces in polygon-based computer-generated holograms. *Applied optics*, 50(34):H245–H252, 2011.

[148] T. Nomura, M. Okamura, E. Nitanai, and T. Numata. Image quality improvement of digital holography by superposition of reconstructed images obtained by multiple wavelengths. *Applied optics*, 47(19):D38–D43, 2008.

[149] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand. Comparison of the coding efficiency of video coding standards—including high efficiency video coding (hevc). *IEEE Transactions on circuits and systems for video technology*, 22(12):1669–1684, 2012.

[150] N. Okada, T. Shimobaba, Y. Ichihashi, R. Oi, K. Yamamoto, T. Kakue, and T. Ito. Fast calculation of computer-generated hologram for rgb and depth images using wavefront recording plane method. *Photonics Letters of Poland*, 6(3):90–92, 2014.

[151] L. Onural. Diffraction from a wavelet point of view. *Optics letters*, 18(11):846–848, 1993.

[152] S. S. Panwar, K. W. Ross, and Y. Wang. On demand peer-to-peer video streaming with multiple description coding, Dec. 15 2009. US Patent 7,633,887.

[153] J. Park, P. A. Chou, and J.-N. Hwang. Rate-utility optimized streaming of volumetric media for augmented reality. *arXiv preprint arXiv:1804.09864*, 2018.

[154] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In *Proceedings of 27th Asilomar conference on signals, systems and computers*, pages 40–44. IEEE, 1993.

[155] K. Pearson. Vii. note on regression and inheritance in the case of two parents. *proceedings of the royal society of London*, 58(347-352):240–242, 1895.

[156] J. Peixeiro. *Holographic information coding*. PhD thesis, MSc Thesis, Instituto Superior Técnico, Lisbon, Portugal, 2016.

[157] J. Peixeiro, C. Brites, J. Ascenso, and F. Pereira. Digital holography: Benchmarking coding standards and representation formats. In *2016 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6. IEEE, 2016.

[158] J. P. Peixeiro, C. Brites, J. Ascenso, and F. Pereira. Holographic data coding: Benchmarking and extending hevc with adapted transforms. *IEEE Transactions on Multimedia*, 20(2):282–297, 2017.

[159] M. Pereira, M. Antonini, and M. Barlaud. Multiple description coding for internet video streaming. In *Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429)*, volume 3, pages III–281. IEEE, 2003.

[160] A. Rahmoune, P. Vandergheynst, and P. Frossard. Mp3d: Highly scalable video coding scheme based on matching pursuit. In *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages iii–133. IEEE, 2004.

[161] S. Reichelt, R. Haussler, N. Leister, G. Futterer, H. Stolle, and A. Schwerdtner. Holographic 3-d displays-electro-holography within the grasp of commercialization. In *Advances in Lasers and Electro Optics*. IntechOpen, 2010.

[162] G. Rilling, P. Flandrin, P. Goncalves, et al. On empirical mode decomposition and its algorithms. In *IEEE-EURASIP workshop on nonlinear signal and image processing*, volume 3, pages 8–11. NSIP-03, Grado (I), 2003.

[163] L. Rong, W. Xiao, F. Pan, S. Liu, and R. Li. Speckle noise reduction in digital holography by use of multiple polarization holograms. *Chinese Optics Letters*, 8(7):653–655, 2010.

[164] D. G. Sachs, A. Raghavan, and K. Ramchandran. Wireless image transmission using multiple-description-based concatenated codes. In *Image and Video Communications and Processing 2000*, volume 3974, pages 300–311. International Society for Optics and Photonics, 2000.

[165] A. Said, W. A. Pearlman, et al. A new, fast, and efficient image codec based on set partitioning in hierarchical trees. *IEEE Transactions on circuits and systems for video technology*, 6(3):243–250, 1996.

[166] P. Schelkens, D. Blinder, T. Birnbaum, A. Symeonidou, A. Ahar, and C. Schretter. Source coding of holographic data: challenges, algorithms and standardization efforts. In *Speckle 2018: VII International Conference on Speckle Metrology*, volume 10834, page 108340Y. International Society for Optics and Photonics, 2018.

[167] P. Schelkens, T. Ebrahimi, A. Gilles, P. Gioia, K.-J. Oh, F. Pereira, C. Perra, and A. M. Pinheiro. Jpeg pleno: Providing representation interoperability for holographic applications and devices. *ETRI Journal*, 41(1):93–108, 2019.

[168] U. Schnars, C. Falldorf, J. Watson, and W. Jüptner. Digital holography. In *Digital Holography and Wavefront Sensing*, pages 39–68. Springer, 2015.

[169] U. Schnars and W. P. Jüptner. Digital recording and numerical reconstruction of holograms. *Measurement science and technology*, 13(9):R85, 2002.

[170] A. Schwerdtner, N. Leister, R. Häussler, S. Reichelt, G. Fütterer, and A. Schwerdtner. 25.2: Eye-tracking solutions for real-time holographic 3-d display. In *SID Symposium Digest of Technical Papers*, volume 39, pages 345–347. Wiley Online Library, 2008.

[171] Y.-H. Seo, H.-J. Choi, J.-W. Bae, H.-J. Kang, S.-H. Lee, J.-S. Yoo, and D.-W. Kim. A new coding technique for digital holographic video using multi-view prediction. *IEICE TRANSACTIONS on Information and Systems*, 90(1):118–125, 2007.

[172] Y.-H. Seo, H.-J. Choi, and D.-W. Kim. Lossy coding technique for digital holographic signal. *Optical Engineering*, 45(6):065802, 2006.

[173] Y.-H. Seo, H.-J. Choi, and D.-W. Kim. 3d scanning-based compression technique for digital hologram video. *Signal Processing: Image Communication*, 22(2):144–156, 2007.

[174] O. G. Sezer. *Data-driven transform optimization for next generation multimedia applications*. PhD thesis, Georgia Institute of Technology, 2011.

[175] N. T. Shaked, B. Katz, and J. Rosen. Review of three-dimensional holographic imaging by multiple-viewpoint-projection based methods. *Appl. Opt.*, 48(34):H120–H136, 2009.

[176] A. E. Shortt, T. J. Naughton, and B. Javidi. Combined nonuniform quantization and lossless coding of digital holograms of three-dimensional objects. In *Three-Dimensional TV, Video, and Display II*, volume 5243, pages 81–88. International Society for Optics and Photonics, 2003.

[177] A. E. Shortt, T. J. Naughton, and B. Javidi. Nonuniform quantization compression techniques for digital holograms of three-dimensional objects. In *Optical Information Systems II*, volume 5557, pages 30–41. International Society for Optics and Photonics, 2004.

[178] A. E. Shortt, T. J. Naughton, and B. Javidi. Vector quantisation compression of digital holograms of three-dimensional objects. In *Opto-Ireland 2005: Photonic Engineering*, volume 5827, pages 265–273. International Society for Optics and Photonics, 2005.

[179] A. E. Shortt, T. J. Naughton, and B. Javidi. A companding approach for nonuniform quantization of digital holograms of three-dimensional objects. *Optics Express*, 14(12):5129–5134, 2006.

[180] A. E. Shortt, T. J. Naughton, and B. Javidi. Compression of digital holograms of three-dimensional objects using wavelets. *Optics Express*, 14(7):2625–2630, 2006.

[181] A. E. Shortt, T. J. Naughton, and B. Javidi. Histogram approaches for lossy compression of digital holograms of three-dimensional objects. *IEEE Transactions on Image Processing*, 16(6):1548–1556, 2007.

[182] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand. Overview of the high efficiency video coding (hevc) standard. *IEEE Transactions on circuits and systems for video technology*, 22(12):1649–1668, 2012.

[183] H. Sun, A. Vetro, and J. Xin. An overview of scalable video streaming. *Wireless Communications and Mobile Computing*, 7(2):159–172, 2007.

[184] W. Sweldens. The lifting scheme: A construction of second generation wavelets. *SIAM journal on mathematical analysis*, 29(2):511–546, 1998.

[185] A. Symeonidou, D. Blinder, A. Ahar, C. Schretter, A. Munteanu, and P. Schelkens. Speckle noise reduction for computer generated holograms of objects with diffuse surfaces. In *Optics, Photonics and Digital Technologies for Imaging Applications IV*, volume 9896, page 98960F. International Society for Optics and Photonics, 2016.

[186] A. Symeonidou, D. Blinder, A. Munteanu, and P. Schelkens. Computer-generated holograms by multiple wavefront recording plane method with occlusion culling. *Optics express*, 23(17):22149–22161, 2015.

[187] D. Taubman. High performance scalable image compression with ebcot. *IEEE Transactions on image processing*, 9(7):1158–1170, 2000.

[188] D. Taubman and M. Marcellin. *JPEG2000 image compression fundamentals, standards and practice: image compression fundamentals, standards and practice*, volume 642. Springer Science & Business Media, 2012.

[189] M. Unser. Splines: A perfect fit for signal and image processing. *IEEE Signal processing magazine*, 16(ARTICLE):22–38, 1999.

[190] v. A. Van der Schaaf and J. v. van Hateren. Modelling the power spectra of natural images: statistics and information. *Vision research*, 36(17):2759–2770, 1996.

[191] K. Viswanathan, P. Gioia, and L. Morin. Wavelet compression of digital holograms: towards a view-dependent framework. In *Applications of Digital Image Processing XXXVI*, volume 8856, page 88561N. International Society for Optics and Photonics, 2013.

[192] K. Viswanathan, P. Gioia, and L. Morin. Morlet wavelet transformed holograms for numerical adaptive view-based reconstruction. In *Optics and Photonics for Information Processing VIII*, volume 9216, page 92160G. International Society for Optics and Photonics, 2014.

[193] K. Viswanathan, P. Gioia, and L. Morin. A framework for view-dependent hologram representation and adaptive reconstruction. In *2015 IEEE International Conference on Image Processing (ICIP)*, pages 3334–3338. IEEE, 2015.

[194] C. Wagner, S. Seebacher, W. Osten, and W. Jüptner. Digital recording and numerical reconstruction of lensless fourier holograms in optical metrology. *Applied Optics*, 38(22):4812–4820, 1999.

[195] G. K. Wallace. The jpeg still picture compression standard. *IEEE transactions on consumer electronics*, 38(1):xviii–xxxiv, 1992.

[196] Y. Wang, A. R. Reibman, and S. Lin. Multiple description coding for video delivery. *Proceedings of the IEEE*, 93(1):57–70, 2004.

[197] J. Watkinson. *The MPEG handbook*. Routledge, 2012.

[198] T. A. Welch. A technique for high-performance data compression. *Computer*, (6):8–19, 1984.

[199] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra. Overview of the h. 264/avc video coding standard. *IEEE Transactions on circuits and systems for video technology*, 13(7):560–576, 2003.

[200] I. H. Witten, R. M. Neal, and J. G. Cleary. Arithmetic coding for data compression. *Communications of the ACM*, 30(6):520–540, 1987.

[201] Y. Xing, M. Kaaniche, B. Pesquet-Popescu, and F. Dufaux. Vector lifting scheme for phase-shifting holographic data compression. *Optical Engineering*, 53(11):112312, 2014.

[202] Y. Xing, M. Kaaniche, B. Pesquet-Popescu, and F. Dufaux. Adaptive nonseparable vector lifting scheme for digital holographic data compression. *Applied optics*, 54(1):A98–A109, 2015.

[203] Y. Xing, M. Kaaniche, B. Pesquet-Popescu, and F. Dufaux. *Digital holographic data representation and compression*. Academic Press, 2015.

[204] Y. Xing, B. Pesquet-Popescu, and F. Dufaux. Compression of computer generated phase-shifting hologram sequence using avc and hevc. In *Applications of Digital Image Processing XXXVI*, volume 8856, page 88561M. International Society for Optics and Photonics, 2013.

[205] Y. Xing, B. Pesquet-Popescu, and F. Dufaux. Vector quantization for computer generated phase-shifting holograms. In *2013 Asilomar Conference on Signals, Systems and Computers*, pages 709–713. IEEE, 2013.

[206] Y. Xing, B. Pesquet-Popescu, and F. Dufaux. Comparative study of scalar and vector quantization on different phase-shifting digital holographic data representations. In *2014 3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON)*, pages 1–4. IEEE, 2014.

[207] I. Yamaguchi and T. Zhang. Phase-shifting digital holography. *Optics letters*, 22(16):1268–1270, 1997.

[208] F. Yaraş, H. Kang, and L. Onural. State of the art in holographic displays: a survey. *Journal of display technology*, 6(10):443–454, 2010.

[209] H.-J. Yeom, H.-J. Kim, S.-B. Kim, H. Zhang, B. Li, Y.-M. Ji, S.-H. Kim, and J.-H. Park. 3d holographic head mounted display using holographic optical elements with astigmatism aberration compensation. *Optics express*, 23(25):32025–32034, 2015.

[210] H. Yoshikawa and T. Yamaguchi. Computer-generated holograms for 3d display. *Chinese optics letters*, 7(12):1079–1082, 2009.

[211] H. Yu, Z. Zhang, and J. Zhong. A jpeg-based enhanced compression algorithm of digital holograms. In *Holography, Diffractive Optics, and Applications V*, volume 8556, page 85560K. International Society for Optics and Photonics, 2012.

[212] X. Yu and D. Hu. A sparse representation image denoising method based on orthogonal matching pursuit. *Telkomnika*, 13(4):1330, 2015.

[213] J. Ziv and A. Lempel. A universal algorithm for sequential data compression. *IEEE Transactions on information theory*, 23(3):337–343, 1977.

# Appendix A

# Publications, patent and distinctions

The main contributions of this thesis have already been published, including two international journal publications, three international conferences, one national conference and one patent.

## A.1 Scientific journals

**[J1] Anas El Rhammad,** Patrick Gioia, Antonin Gilles, Marco Cagnazzo and Béatrice Pesquet-Popescu, "Color digital hologram compression based on matching pursuit", in *Applied optics, The Optical Society*, Vol. 57, No. 17, pages 4930–4942, 2018.

*Abstract* - With the recent widespread interest for head-mounted displays applied to virtual or augmented reality, holography has been considered as an appealing technique for a revolutionary and natural 3D visualization system. However, due to the tremendous amount of data required by holograms and to the very different properties of holographic data compared to common imagery, compression of digital holograms is a highly challenging topic for researchers. In this study, we introduce a novel approach for color hologram compression based on matching pursuit and using an overcomplete Gabor's dictionary. A detailed framework, together with a GPU-implementation, from hologram decomposition to bitstream generation is studied and the results are discussed and compared to existing hologram compression algorithms.

**[J2] Anas El Rhammad,** Patrick Gioia, Antonin Gilles and Marco Cagnazzo, "Progressive hologram transmission using a view-dependent scalable compression scheme", in *Annals of Telecommunications, Springer*, 2019.

*Abstract* Over the last few years, holography has been emerging as an alternative to stereoscopic imaging since it provides users with the most realistic and comfortable 3D experience. However, high quality holograms enabling a free-viewpoint visualization contain

tremendous amount of data. Therefore, a user willing to access to a remote hologram repository would face high downloading time, even with high speed networks. To reduce transmission time, a joint viewpoint-quality scalable compression scheme is proposed. At the encoder side, the hologram is first decomposed into a sparse set of diffracted light rays using Matching Pursuit over a Gabor atoms dictionary. Then, the atoms corresponding to a given user's viewpoint are selected to form a sub-hologram. Finally, the pruned atoms are sorted and encoded according to their importance for the reconstructed view. The proposed approach allows a progressive decoding of the sub-hologram from the first received atom. Streaming simulations for a moving user reveal that our approach outperforms conventional scalable codecs such as scalable H.265 and enables a practical streaming with a better quality of experience.

## A.2 National/International Conferences

**[C1] Anas El Rhammad,** Patrick Gioia, Antonin Gilles, Marco Cagnazzo and Béatrice Pesquet-Popescu, "View-dependent compression of digital holograms based on matching pursuit", in *SPIE Photonics Europe*, Strasbourg, France, 2018.

*Abstract* - In this paper we investigate the suitability of Gabor Wavelets for an adaptive partial reconstruction of holograms based on the viewer position. Matching Pursuit is used for a sparse light rays decomposition of holographic patterns. At the decoding stage, sub-holograms are generated by selecting the diffracted rays corresponding to a specific area of visualization. The use of sub-holograms has been suggested in the literature as an alternative to full compression, by degrading a hologram with respect to the directional degrees of freedom. We present our approach in a complete framework for color digital holograms compression and explain, in details, how it can be efficiently exploited in the context of holographic Head-Mounted Displays. Among other aspects, encoding, adaptive reconstruction and selective degradation are studied.

**[C2] Anas El Rhammad,** Patrick Gioia, Antonin Gilles and Marco Cagnazzo, "Compression adaptative d'hologrammes numériques en utilisant la technique du "Matching Pursuit"", in *COmpression et REprésentation des Signaux Audiovisuels (CORESA)*, Poitiers, France, 2018.

*Abstract* - L'objectif de cet article est d'étudier la pertinence des ondelettes de Gabor pour la représentation ainsi que la compression des hologrammes numériques en fonction de la position de l'observateur. Le "Matching Pursuit" est utilisé pour une décomposition parcimonieuse des rayons lumineux émis par l'hologramme. Un sous-hologramme est généré

en sélectionnant les rayons diffractés vers le champ de visualisation. Dans la littérature, l'utilisation des sous-hologrammes a été proposée comme une alternative à la compression des données massives contenues dans un signal holographique. Notre approche est présentée dans un schéma complet de transmission d'hologrammes en couleur qui peut notamment être utilisé dans le cadre des visiocasques.

**[C3] Anas El Rhammad,** Patrick Gioia, Antonin Gilles and Marco Cagnazzo, "Scalable coding framework for a view-dependent streaming of digital holograms", in *IEEE International Conference on Image Processing (ICIP)*, Taipei, Taiwan, 2019.

*Abstract* - Unlike conventional images and videos, digital holograms contain large amounts of data with very low redundancy. Consequently, current communication networks may be not able to meet the bandwidth requirements for hologram transmission in reasonable time. To enable practical streaming of holographic contents, we propose a progressive coding method that combines quality scalability with viewpoint scalability. From a Gabor wavelets decomposition of the hologram, the server starts by selecting the coefficients corresponding to the user's viewpoint. Then, the selected coefficients are encoded progressively according to their importance for the reconstructed view. Experimental results reveal that our approach outperforms conventional scalable codecs and enables the streaming of holographic data with a better quality of experience.

**[C4] Anas El Rhammad,** Patrick Gioia, Antonin Gilles and Marco Cagnazzo, "Towards practical hologram streaming using progressive coding", in *SPIE Optics + Photonics*, San Diego, USA, 2019.

*Abstract* - Digital holography is an emerging technology for 3D visualization which is expected to dethrone conventional stereoscopic devices in the future. Aside from their specific signal properties, high quality holograms with broad viewing angles contain massive amount of data. For a reasonable transmission time, efficient scalable compression schemes are needed to bridge the gap between the overwhelming volume of data and the limited bandwidth of the communication channels. The viewpoint scalability is a powerful property since it allows to encode and transmit only the information corresponding to the observer's view. However, this approach imposes an online encoding at the server which may increase the latency of the transmission chain. To overcome this hurdle, we propose a scalable compression framework based on Gabor-wavelets decomposition, where the whole hologram is encoded offline. First, the observer plane is divided into spatial blocks. Then, the Gabor atoms are assigned to these blocks by exploiting the duality between Gabor wavelets and light rays. The atoms of each block are then classified into different layers according to their importance for the reconstruction and encoded in packets. At the decoder side, the atoms'

packets are progressively decoded based on the viewer's position. Then, the corresponding sub-hologram is generated using a GPU implementation. Results show that our approach enables a practical progressive streaming of digital holograms with a low latency.

## A.3 Patent

**[P1] Anas El Rhammad,** Patrick Gioia, Antonin Gilles and Marco Cagnazzo, "Procédé et dispositif de transmission de données représentatives d'un hologramme numérique", France, $20190130\_CD393FR$, 2019.

## A.4 Distinctions

- **Best Student Paper award** in *SPIE Photonics Europe* 2018 for our paper entitled "View-dependent compression of digital holograms based on matching pursuit".
- **Editors' Pick honor** from the *Optical Society* in 2018, recognizing our journal paper "Color digital hologram compression based on matching pursuit".
- **Amongst the top 5 papers** at CORESA 2018 for the work on view-dependent compression of digital holograms using Matching Pursuit (proposed for an extended version as an invited paper in the Annals of Telecommunications journal, the extension was accepted).

# Appendix B

# Videos of streaming simulations

The videos corresponding to the streaming of the *Piano8K* and *Dices8K* holograms can be downloaded from the link below:

"https://www.researchgate.net/publication/338117265_holo_streaming"

**Titre :** Compression efficace pour une transmission scalable des hologrammes numériques

**Mots clés :** Holographie, Hologramme numérique, Compression, Transmission, Scalabilité.

**Résumé :** Contrairement à la stéréoscopie, l'holographie permet une visualisation 3D la plus naturelle et confortable possible. En revanche, les hologrammes numériques contiennent une grande quantité de données avec peu de corrélations.
Dans la première partie de ce travail, nous avons introduit deux méthodes de compression des hologrammes numériques. Tout d'abord, nous avons décomposé l'hologramme en faisceaux lumineux en utilisant les ondelettes de Gabor. Pour des fins de compression, nous avons compacté l'expansion obtenue en utilisant l'algorithme du «Matching Pursuit». Ensuite, nous avons conçu un schéma de codage spécifique aux coefficients et indices des atomes de Gabor. L'approche proposée a atteint une meilleure performance par rapport aux codeurs classiques.
Deuxièmement, en exploitant la dualité entre les ondelettes de Gabor et les faisceaux lumineux émis par l'hologramme, nous avons développé un schéma de codage scalable en termes de point de vue et qualité. En effet, seuls les atomes de Gabor qui émettent la lumière dans la fenêtre de l'observateur considéré sont sélectionnés, triés et puis encodés. Le débit binaire a été considérablement réduit, sans dégrader la qualité de reconstruction par rapport à celle obtenue en encodant la totalité de l'hologramme.
Dans la deuxième partie de ce travail, nous avons conçu deux architectures serveur-client pour une transmission progressive d'hologrammes en fonction de la position d'observateur à l'aide d'un codage scalable. Dans la première solution, un flux binaire scalable à grain fin est généré en ligne par le serveur, après chaque notification du client concernant la position de l'utilisateur. Les résultats expérimentaux ont révélé que la méthode proposée permet une visualisation rapide en décodant les premiers atomes reçus en plus d'une augmentation progressive de la qualité.
Enfin, pour réduire la latence causée par la charge de calcul au moment du codage, nous avons proposé une deuxième solution où l'intégralité de l'expansion de Gabor est encodée hors ligne par le serveur, puis décodée en ligne suivant la trajectoire de l'observateur. Pour permettre un codage scalable, nous avons regroupé les atomes de Gabor à la suite d'une décomposition par blocs du plan de l'observateur. Ensuite, les atomes de chaque bloc sont affectés à différents niveaux de qualité et codés par paquets. Les tests de simulation ont montré que l'architecture proposée permet une transmission à faible latence sans augmenter significativement le coût d'encodage.

**Title :** Efficient compression for scalable transmission of digital holograms

**Keywords :** Holography, Digital hologram, Compression, Transmission, Scalability.

**Abstract :** Contrary to conventional stereoscopy, holography provides the most natural and comfortable 3D visualization. However, digital holograms contain massive amount of data with very few correlations.
In the first part of this work, we introduced two methods for digital holograms compression. First, we proposed a redundant light beams-based decomposition of holograms using the Gabor wavelets. For compression purposes, we sparsified the obtained expansion using the Matching Pursuit algorithm. Then, we designed a specific encoder framework for the coefficients and indexes of Gabor atoms. The proposed approach achieved better compression performance compared to the state of the art methods.
Second, by exploiting the duality between Gabor wavelets and diffracted light beams, we developed a viewpoint-quality scalable coding scheme. Indeed, for a given observer's position, only the Gabor atoms that emit light into the viewer's window are selected, sorted and then encoded. The bit rate has been significantly reduced, without degrading the reconstruction quality obtained by encoding the whole hologram.
In the second part of this work, we designed two server-client architectures for a view-dependent progressive transmission of holograms using scalable coding. In the first solution, a fine-grain scalable bitstream is generated online by the server, after each client notification about the user's position. Experimental results reveal that this method enables a rapid visualization by decoding the first received atoms in addition to a progressive increase of quality.
Finally, to reduce the latency caused by the computational burden of encoding, we proposed a second solution where the whole Gabor expansion is encoded offline by the server, and then decoded online with respect to the viewer's trajectory. To enable a scalable compression, we grouped the Gabor atoms following a block-based decomposition of the observer plane. Then, the atoms of each block are assigned to different quality levels and encoded in packets. Simulations tests show that the proposed architecture allows a low-latency transmission without significantly increasing the encoding rate.

**Institut Polytechnique de Paris**
91120 Palaiseau, France