UNIVERSITÉ DU QUÉBEC

INRS Institut national de la recherche scientifique

Centre Énergie Matériaux Télécommunications

Orthogonality-based Beam Shaping for Three-dimensional Surface Imaging and Long-wavelength Infrared Single-pixel Imaging Thermometry

Par

Cheng Jiang

Thèse présenté(e) pour l'obtention

du grade de Philosophiae Doctor (Ph.D.)

en sciences de l'énergie et des matériaux

Jury d'évaluation

Président du jury et examinateur interne	Professeur Tsuneyuki Ozaki INRS-ÉMT
Examinateur externe	Professeur Beiwen Li Iowa State University
Examinateur externe	Professeur Michael F. Becker University of Texas at Austin
Directeur de recherche	Professeur Jinyang Liang INRS-ÉMT

Dedication

To my beloved grandfather, Dongyang Yu.

ACKNOWLEDGEMENTS

I would like to express my gratitude to many people without whom I cannot finish this dissertation. First and foremost, my sincere thanks go to my supervisor, Dr. Jinyang Liang. His unwavering guidance, continuous availability, and encouraging words, not only propelled my research but also greatly influenced the shaping of this dissertation. Dr. Patrick Kilcullen also deserves special mention for offering crucial insights, steadfast support, and engaging in invaluable discussions that benefitted my research.

I extend my appreciation to Dr. Tsuneyuki Ozaki, Dr. Beiwen Li, and Dr. Michael F. Becker for their insightful feedback and their role as members of my dissertation committee.

I also would like to thank my laboratory colleagues Dr. Xianglei Liu, Yingming Lai, Alejandra Gomez Ramirez, Xiaoyue Yan, Dr. Miguel Marquez, Beniwal Madhu, Dr. Jingdan Liu, Miao Liu, Siqi Wang, for their assistance and helpful suggestions.

In addition, I would like to appreciate for the unwavering support and inspiration provided by my friends and colleagues. This journey has been enriched and influenced by each one of you. My special thanks go to Cheng Lu, Qiaochu Wang, Dr. Jingxi Li, Dr. Yuhao Wu, Dr. Penghui Ji, Yi Wang, Kaiyuan Wang, Yuqi (Yuki) Zhou, Sze Yui (Gary) Fung, Jie Zheng, Haotian Yu, Dr. Xin Chai, Dr. Qingzhe Zhang, Dr. Li Shi, Dr. Yuting Lei, Dr. Daling Cui, Dr. Junliang Dong, Pei You, Qiwen Qiu, Huan Zheng, Weijia Li, Yuxiao Zhang, Chengcheng Tang, Dr. Yang Liu, Dr. Faying Li, Dr. Fan Yang, Dr. Yong Wang, Chen Wang, and Wanting He.

Lastly, my deepest thanks go to my family. This work is dedicated to my parents and grandparents, whose emotional support has been my pillars of strength throughout this journey.

RÉSUMÉ

Les techniques de mise en forme de faisceau basées sur l'orthogonalité, qui préservent les entités de simplicité dans le calcul, la décomposition efficace du signal et la stabilité avec des informations non corrélées, sont essentielles dans divers domaines et applications. En particulier, cette thèse se concentre sur la mise en forme des profils d'intensité basés sur l'orthogonalité des franges sinusoïdales et des matrices S cycliques, qui sont intégrales pour la profilométrie de projection de franges (FPP) et l'imagerie à pixel unique (SPI). L'imagerie de surface tridimensionnelle (3D) à haute vitesse de la FPP répond aux secteurs industriels, médicaux et de divertissement. La demande croissante pour un retour d'information 3D en temps réel et précis dans des dynamiques rapides, y compris l'inspection des collisions et les évaluations de la qualité des produits, nécessite des améliorations continues de la technologie FPP pour assurer une imagerie 3D robuste et rapide. Pendant ce temps, le SPI offre une flexibilité et un rapport qualité-prix, devenant essentiel pour des tâches d'imagerie spécialisées, par exemple, l'imagerie non visible. Au milieu de la pandémie, il y a eu un intérêt croissant pour des solutions d'imagerie thermométrique, ce qui stimule l'exploration de l'application du SPI dans la plage spectrale infrarouge à longue longueur d'onde (LWIR).

Cependant, dans les systèmes FPP existants, la limitation de la vitesse de projection a empêché la frange sinusoïdale d'atteindre le niveau kilohertz (kHz). La bande passante limitée de la transmission de données a empêché la caméra de diffuser des données en continu, ce qui a donc entraîné des difficultés dans l'acquisition, le traitement et l'affichage d'images 3D au niveau kHz lors de la survenue d'événements dynamiques (c'est-à-dire en temps réel). Les méthodes existantes luttent pour atteindre un large champ de vision (FOV) à des acquisitions de niveau kilohertz (kHz) en raison d'un compromis entre les taux de lecture des capteurs et le nombre de pixels activés. De plus, elles ne peuvent souvent pas simultanément fournir robustesse, tolérance aux variations de réflectance et flexibilité des distances de travail réglables avec des FOV de niveau mètre carré à des taux vidéo.

Pour relever les défis de la FPP, nous avons développé une profilométrie d'illumination à bande limitée à grande vitesse (BLIP) en trois configurations. La première configuration, utilisant une seule caméra avec une interface CoaXPress (CI), permet la reconstruction d'informations de surface 3D en temps réel à 1 kHz. La deuxième configuration, utilisant deux caméras avec une CI, utilise une acquisition interlacée temporellement (TIA) pour améliorer l'imagerie 3D à plus de 1000 images par seconde sur un FOV allant jusqu'à 180 × 130 mm2. La troisième configuration

est la BLIP multi-échelle (MS). Soutenue par la synergie de la projection d'intensité à deux niveaux, la projection de franges multi-fréquences et une méthode itérative pour la compensation de la distorsion, la MS-BLIP peut discerner avec précision des objets 3D spatialement séparés avec une réflectance très variable. La CI-BLIP a été appliquée à la mécanique des fluides en imageant la dynamique d'un drapeau, ce qui a permis d'observer la propagation des ondes, le déphasage induit par la gravité et le mouvement de battement asymétrique. Pendant ce temps, la TIA-BLIP a permis la visualisation 3D de la vibration du verre induite par le son. En outre, avec un FOV allant jusqu'à 1,7 m × 1,1 m et une distance de travail allant jusqu'à 2,8 m, la MS-BLIP est appliquée à la capture de mouvements corporels humains complets à un taux vidéo.

Contrairement à la BLIP, qui repose fondamentalement sur une illumination active, le SPI peut être utilisé en mode de détection passive, le rendant plus adapté aux applications avec des considérations de sécurité. Les méthodes traditionnelles de mesure de la température peuvent être invasives ou présenter des préoccupations de sécurité, telles que les thermomètres infrarouges qui utilisent une illumination active. Bien que la thermographie bidimensionnelle (2D) offre une alternative non invasive, elle est plus coûteuse et son large champ de vision ne cible pas entièrement les canthi internes, qui sont cruciaux pour des lectures précises de la température humaine. Cela entraîne une redondance dans l'acquisition de données et pourrait entraîner des inexactitudes en raison du mélange avec l'arrière-plan. Lorsqu'il est configuré pour se concentrer sur les canthi internes, le SPI peut réduire les ressources nécessaires pour la thermographie 2D et éliminer le mélange avec l'arrière-plan. Cependant, l'application du SPI à l'imagerie thermométrique LWIR a été entravée en raison de plages spectrales qui dépassent la fenêtre de transmission du verre de projection et les spécifications de revêtement optimisées des encodeurs spatiaux tels que les dispositifs à micromiroir numérique (DMD), entravant le dépistage efficace de la température corporelle humaine.

En réponse à ce défi, nous développons une thermométrie d'imagerie infrarouge à pixel unique (SPIRIT) avec détection passive dans la gamme spectrale LWIR. Utilisant des matrices S cycliques, l'agrégation diagonale est développée pour générer les motifs de masquage sur un masque physique, ce qui facilite le balayage unidimensionnel (1D). Tirant parti de la douceur 2D de la matrice de signaux encodée, un balayage linéaire compatible avec la détection comprimée et une interpolation de données sont mis en œuvre pour une reconstruction d'image efficace. SPIRIT permet la première imagerie thermique des canthi internes par SPI avec une taille de trame de 11 × 13 pixels. Avec une analyse statistique, SPIRIT a effectué le dépistage de la

température de sujets humains et a étudié les fluctuations de la température corporelle et l'impact de l'utilisation prolongée de lunettes.

Mots-clés: Mise en forme du faisceau basée sur l'orthogonalité; franges sinusoïdales; matrices S cycliques; profilométrie par projection de franges; imagerie à pixel unique; imagerie de surface en trois dimensions; thermométrie d'imagerie infrarouge à longue longueur d'onde; projection haute vitesse; dépistage de la température corporelle humaine.

ABSTRACT

Orthogonality-based beam shaping techniques, which preserve the entities of simplicity in computation, efficient signal decomposition, and stability with uncorrelated information, are critical in various fields and applications. In particular, this dissertation focuses on shaping the orthogonality-based intensity profiles of sinusoidal fringes and cyclic S-matrices, which are integral for fringe projection profilometry (FPP) and single-pixel imaging (SPI). FPP's high-speed three-dimensional (3D) surface imaging caters to industrial, medical, and entertainment sectors. The growing demand for precise real-time 3D feedback in rapid dynamics, including collision inspection and product quality assessments, necessitates ongoing enhancements in FPP technology to ensure robust, high-speed 3D imaging. Meanwhile, SPI provides flexibility and cost-effectiveness, becoming crucial for specialized imaging tasks, e.g., non-visible imaging. Amidst the pandemic, there has been a growing interest in solutions for imaging thermometry, which drives exploration of SPI's application in the long-wavelength infrared (LWIR) spectral range.

However, in existing FPP systems, the limitation of projection speed has prevented the sinusoidal fringe to reach the kilohertz (kHz) level. The limited bandwidth of the data transmission has prevented the camera from streaming data continuously, which thus has brought difficulties in kHz-level image acquisition, processing, and display of 3D information during the occurrence of dynamic events (i.e., in real time). Existing methods grapple with achieving a wide field of view (FOV) at kilohertz (kHz)-level acquisitions due to a trade-off between sensor readout rates and activated pixel numbers. Additionally, they often can't simultaneously provide robustness, tolerance to reflectance variations, and flexibility of tunable working distances with meter-square-level FOVs at video rates.

To address FPP's challenges, we have developed high-speed band-limited illumination profilometry (BLIP) in three configurations. The first configuration, employing a single camera with a CoaXPress interface (CI), enables real-time 3D surface information reconstruction at 1 kHz. The second configuration, employing two cameras with a CI, uses temporally interlaced acquisition (TIA) to improve the 3D imaging over 1000 frames per second on an FOV of up to 180×130 mm^2. The third configuration is multi-scale (MS) BLIP. Supported by the synergy of dual-level intensity projection, multi-frequency fringe projection, and an iterative method for distortion compensation, MS-BLIP can accurately discern spatially separated 3D objects with highly varying reflectance. CI-BLIP has been applied to fluid mechanics by imaging dynamics of a flag, which allowed observation of the wave propagation, gravity-induced phase mismatch, and asymmetric

flapping motion. Meanwhile, TIA-BLIP has empowered the 3D visualization of glass vibration induced by sound. In addition, with an FOV of up to 1.7 m × 1.1 m and a working distance of up to 2.8 m, MS-BLIP is applied to capturing full human-body movements at video rate.

Unlike BLIP, which fundamentally relies on active illumination, SPI can be utilized in a passive detection mode, making it more suitable for applications with safety considerations. Traditional temperature measurement methods can be invasive or carry safety concerns, such as infrared thermometers that use active illumination. While two-dimensional (2D) thermography offers a non-invasive alternative, it's more expensive and its wide field of view doesn't fully target the inner canthi, which are crucial for accurate human temperature readings. This brings redundancy in data acquisition and could result in inaccuracies due to background blending. When configured to focus on the inner canthi, SPI can reduce the resources needed for 2D thermography and eliminate background blending. However, SPI's application to LWIR imaging thermometry has been impeded due to spectral ranges that exceed the transmittance window of the prorective glass and the optimized coatings specifications of spatial encoders like digital micromirror devices (DMDs), hindering efficient human body temperature screening.

In response to this challenge, we develop single-pixel infrared imaging thermometry (SPIRIT) with passive detection at the spectral range of LWIR. Utilizing cyclic S-matrices, diagonal aggregation is developed to generate the masking patterns onto a physical mask, which facilitates one-dimensional (1D) scanning. Leveraging the 2D smoothness of the encoded signal matrix, a compressed-sensing-compatible linear scan and data interpolation are implemented for efficient imaging reconstruction. SPIRIT enables the first-time thermal imaging of the inner canthi by SPI with a frame size of 11 × 13 pixels. With statistical analysis, SPIRIT has performed temperature screening of human subjects and investigated body temperature fluctuations and the impact of prolonged eyeglass usage.

Keywords : Orthogonality-based beam shaping; sinusoidal fringes; cyclic S-matrices; fringe projection profilometry; single-pixel imaging; three-dimensional surface imaging; long-wavelength infrared imaging thermometry; high-speed projection; human body temperature screening.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBRIATIONS

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
mm	Millimeter
μm	Micrometer
ms	Millisecond
μs	Microsecond
Hz	Hertz
kHz	Kilohertz
fps	Frames per second
kfps	Thousands fps
CCD	Charge-coupled device
CMOS	Complementary metal-oxide-semiconductor
FOV	Field of view
DOF	Depth of field
ESF	Edge spread function
LSF	Line spread function
Gbps	Gigabit per second
BLIP	Band-limited illumination profilometry
CI-BLIP	BLIP with CoaXPress interface
TIA-BLIP	BLIP with temporally interlaced acquisition
MS-BLIP	Multi-scale BLIP
SPIRIT	Single-pixel infrared imaging thermometry
SNR	Signal to noise ratio
SWIR	Short-wavelength infrared
MIR	Mid-wavelength infrared
LWIR	Long-wavelength infrared
SLM	Spatial light modulator
DMD	Digital micro-mirror device
FPP	Fringe projection profilometry
FTP	Fourier transform profilometry

PSFPP	Phase-shifting FPP
SLP	Structed-light profilometry
ED	Error diffusion
SPI	Single pixel imaging
THz	Terahertz
MSE	mean squared error
GPU	graphics processing unit
CPU	Central Processing Unit
RAM	Random access memory
MZT	Motorized z-translation stage
LA	Lock-in amplifier
LED	Light-emitting diode
TIR	Total internal reflection

LIST OF PUBLICATIONS

Book Chapter:

[1] **C. Jiang**, Y. Li, S. Feng, Y. Hu, W. Yin, J. Qian, C. Zuo, J. Liang, “Fringe projection profilometry”, in *Coded Optical Imaging (Chapter 14)*, Springer Nature (2024).

Peer-reviewed Journal Publications (*equal contribution):

[1] **C. Jiang**, P. Kilcullen, Y. Lai, S. Wang, T. Ozaki, and J. Liang, “Multi-scale band-limited illumination profilometry for robust three-dimensional surface imaging at video rate”, Optics Express 30, 19824-19838 (2022).

[2] **C. Jiang***, P. Kilcullen*, Y. Lai, T. Ozaki, J. Liang, “High-speed dual-view band-limited illumination profilometry using temporally interlaced acquisition”, Photonics Research 8, 1808-1817 (2020).

[3] **C. Jiang***, P. Kilcullen*, X. Liu, J. Gribben, A. Boate, T. Ozaki, J. Liang, “Real-time High-speed Three-dimensional Surface Imaging using CoaXPress-Interfaced Band-Limited Illumination Profilometry”, Optics Letters 45, 964-967 (2020).

[4] M. Marquez, Y. Lai, X. Liu, **C. Jiang**, S. Zhang, H. Arguello, J. Liang, “Deep-Learning Supervised Snapshot Compressive Imaging Enabled by an End-to-End Adaptive Neural Network“, IEEE Journal of Selected Topics in Signal Processing 16, 688-699 (2022).

[5] X. Liu, A. Skripka, Y. Lai, **C. Jiang**, J. Liu, F. Vetrone, J. Liang, "Single-shot phosphorescence lifetime imaging thermometry for real-time wide-field dynamic temperature sensing", Nature Communications 12, 6401 (2021)

[6] X. Liu, J. Monteiro, I. Albuquerque, Y. Lai, **C. Jiang**, S. Zhang, T. H. Falk, J. Liang, "Single-shot real-time compressed ultrahigh-speed imaging enabled by a snapshot-to-video autoencoder", Photonics Research 9, 2464-2474 (2021)

[7] P. Kilcullen, **C. Jiang**, T. Ozaki, J. Liang, “Camera-free three-dimensional dual photography”, Optics Express 28, 29377-29389 (2020)

Conference proceedings:

[1] **C Jiang**, P Kilcullen, Y Lai, S Wang, T Ozaki, J Liang, “Video-rate multi-scale band-limited illumination profilometry with cubic-meter-level measurement volume”, in Photonics North (PN), 2023

[2] **C Jiang**, P Kilcullen, Y Lai, S Wang, T Ozaki, J Liang, " Dynamic three-dimensional imaging with cubic-meter-level measurement volume using multi-scale band-limited illumination profilometry", Proceedings of SPIE 12435, Emerging Digital Micromirror Device Based Systems and Applications XV, San Francisco, United States (2023).

[3] **C. Jiang** and J. Liang, "High-speed band-limited illumination profilometry (BLIP)," in Imaging and Applied Optics Congress, 2022

[4] **C. Jiang**, P. Kilcullen, X. Liu, Y. Lai, T. Ozaki and J. Liang, "High-speed band-limited illumination profilometry", Photonics North (PN), 2021.

[5] **C. Jiang**, P. Kilcullen, X. Liu, Y. Lai, T. Ozaki, J. Liang, " High-speed three-dimensional surface measurement using band-limited illumination profilometry (BLIP)", Proceedings of SPIE 11698, Emerging Digital Micromirror Device Based Systems and Applications XIII (116980V), San Francisco, United States (2021).

[6] **C. Jiang**, P. Kilcullen, X. Liu, T. Ozaki, J. Liang, "Three-dimensional structured light profilometry using a bandwidth-limited projector", Proceedings of SPIE 10932, Emerging Digital Micromirror Device Based Systems and Applications XI (109320K), San Francisco, United States (2019).

Patents:

[1] J Liang, **C Jiang**, P Kilcullen, "A method and a system for 3D imaging", US Patent App. 17/310,489, 2022.

[2] J Liang, **C Jiang**, P Kilcullen, "Method and system for high-speed dual-view band-limited illumination profilometry", US Patent App. 17/444,307, 2022

1 INTRODUCTION

1.1 Motivation

Orthogonality, in terms of geometry, indicates two lines (or planes) are perpendicular to each other and intersect at a right angle. In terms of vectors and functions, orthogonality refers to a dot product (inner product) being zero. The attractiveness of orthogonality, in general, is that orthogonal entities don't "interfere" with each other. In data analysis and signal processing, orthogonal components contain uncorrelated information, which can be useful for separating signals from noise or for compressing information (Oppenheim *et al.*, 1997). Furthermore, in numerical computations and algorithms, using orthogonal vectors or functions can lead to simplicity in computations and more stable results, reducing the numerical errors (Axler, 2015). In addition, Orthogonal functions or vectors can be used as a basis to represent other functions or vectors, which can benefit an efficient data decomposition (Stein & Shakarchi, 2011). Leveraging these uniqueness, orthogonality-based beam shaping techniques have shown their great potential by utilizatingorthogonal functions to tailor light beam profiles.

Hermite-Gaussian (HG) modes, for example, which remain orthogonal across the plane perpendicular to propagation, serve as propagation-invariant modes in parabolic-index multi-mode fibers and closely echo the communication modes of square apertures in free-space propagation (Flaes *et al.*, 2018; Rodenburg, 2015). Concurrently, Laguerre-Gaussian (LG) modes, exhibiting similar orthogonality, are pivotal in stimulated emission depletion (STED) microscopy, enabling super-resolution imaging (Hell & Wichmann, 1994).

Specifically, this dissertation focuses on the beam shaping techniques based on sinusoidal functions and Hadamard matrices, which are extensively studied for beam shaping over decades (Horadam, 2012; Saxena *et al.*, 2015). Sinusoidal beam shaping involves modulating the beam profile such that it has a sinusoidal intensity variation across its transverse section. The ability to decompose a complex waveform into orthogonal sinusoidal components can be very beneficial for signal analysis, which has led to the wide applications such as structured illumination microscopy (Gustafsson, 2000), spatial frequency domain imaging (Gioux *et al.*, 2019) and fringe projection profilometry (FPP) (Xu & Zhang, 2020). In addition, the Hadamard matrix is a square matrix with elements being either +1 or -1 (Horadam, 2012), the rows (or columns) of which is mutually orthogonal. Hadamard matrix has found applications in areas like coding theory (Ping *et al.*, 2003), signal processing (Yarlagadda & Hershey, 2012) and quantum computing (Leung *et al.*, 2000) due to its simplicity in computation.

Owing to the phase information of the sinusoidal fringe pattern's insensitivity to reflectivity changes on an object's surface, this sinusoidal fringe beam shaping is especially crucial for FPP in three-dimensional (3D) surface imaging (Li *et al.*, 2014). Meanwhile, the continuous amplitude modulation of the sinusoidal beam profile offers high-resolution in pixel-to-pixel correspondence between the projector and camera, paving the way for advanced sub-pixel imaging accuracy (Zuo *et al.*, 2018).

The sinusoidal fringes employed in FPP are commonly generated by using digital micromirror devices (DMDs) (Liang, 2020). Each micromirror on the DMD can be independently tilted to either +12° or –12° from its surface normal to generate binary patterns at up to tens of kilohertz. Despite being a binary amplitude spatial light modulator, by controlling the spatially dependent reflectance via temporal dithering of each micromirror, the DMD can produce grayscale sinusoidal patterns (Hornbeck, 1997). The DMD-based FPP, flexible in system development and accurate in 3D measurements (Su & Zhang, 2010), has been implemented in numerous scientific and industrial applications (Deetjen *et al.*, 2017; Li & Zhang, 2018).

Despite these advantages, most existing systems are insufficient to reach a kilohertz (kHz)-level rate for sinusoidal fringe pattern generation, which is highly demanded for high-speed 3D surface imaging. The projection rate of fringe patterns based on the conventional dithering method is clamped at hundreds of hertz (Zhang, 2018). To improve the projection speed, binary defocusing techniques (Hyun *et al.*, 2017; Li *et al.*, 2014; Li & Zhang, 2017) have been developed to produce a quasi-sinusoidal pattern by slightly defocusing a single binary DMD pattern. Nonetheless, the image is generated at a plane unconjugate to the DMD, which compromises the depth-sensing range and is less convenient to operate with fringe patterns of different frequencies. Meanwhile, the DMD's uneven surface (Parthiban *et al.*, 2016) could induce image distortion to the defocused sinusoidal patterns at the unconjugate plane, which may decrease measurement accuracy, especially under coherent illumination.

Additionally, the Hadamard matrix plays a pivotal role in single-pixel imaging (SPI) (Gibson *et al.*, 2020). The inherent orthogonal characteristics of the Hadamard matrix ensure efficient inverse computations and bestow a robust resistance to noise, thereby optimizing signal acquisition and reconstruction (Sloane & Harwit, 1979). Experiments using Hadamard matrices require paired measurements due to their negative entries. To overcome this, cyclic S-matrices, which are derivated from Hadamard matrices and only have elements of {0,+1}—are fundamental masking patterns used in SPI (Zhang *et al.*, 2017). Cyclic S-matrices become the choice of multiple SPI systems driven by several advantages. Firstly, S-matrices inherently optimize

multiplex advantage for {0, 1}-valued encoding designs (Wehlburg *et al.*, 2001; Yue *et al.*, 2018), coupled with the benefit of possessing easily computable analytical inverses. In addition, the inherent cyclic property facilitates the amalgamation of multiple masking patterns into adjacent positions, rendering them apt for efficient placement on a physical substrate. Furthermore, this property also fosters highly efficient data processing, as the storage of merely the initial encoding is ample for computing both the set of encoding patterns and the matrix transform employed in image reconstruction.

Akin to FPP, the intensity modulation in SPI typically employs spatial light modulators, e.g., DMDs (Edgar *et al.*, 2015; Gibson *et al.*, 2020; Hahamovich *et al.*, 2021). Leveraging DMDs in SPI allows for the rapid display of masking matrices, modulating the target's information and ensuring high-speed spatial data recovery (Kilcullen *et al.*, 2022). By eliminating the need for a pixel-array sensor together with compressive sensing, SPI offers a cost-effective solution. This advantage has expanded SPI's use in non-visible imaging, including short-wavelength infrared imaging (Tao *et al.*, 2022), mid-wavelength infrared imaging (Wang *et al.*, 2023), terahertz imaging (Stantchev *et al.*, 2020), and photoacoustic imaging (Liang *et al.*, 2014).

SPI boasts unique capabilities in long-wavelength infrared (LWIR) imaging, notably in applications like human body temperature screening—a crucial tool in the wake of pandemics. Traditional temperature measurement methods are either inconvenient due to their invasive or semi-invasive nature (Buono *et al.*, 2007). Even non-invasive approaches, such as thermal guns, carry safety concerns stemming from their reliance on active illumination (Furstenberg *et al.*, 2008). To mitigate these concerns, two-dimensional (2D) thermography emerged, although it typically comes with a higher cost (Sarawade & Charniya, 2018). A drawback of 2D thermography is its broad field of view (FOV), which does not fully allocate the active pixels targeting the inner canthi—the critical area for accurate human body temperature measurement. (Ferrari *et al.*, 2021). The temperature readout with a relatively small amount of pixels can lead to inaccuracies due to background blending (Avdelidis & Moropoulou, 2004). SPI, when adeptly configured to concentrate solely on the inner canthi, could effectively reduce the redundancy in expense and active resources required for 2D thermography, all while eliminating background blending.

Yet, despite these merits, the application of SPI in human body thermometry remains uncharted. A significant barrier is the limited efficacy of DMDs within the long-wavelength infrared (LWIR) domain due to the low transmittance of the protection glass as well as the low reflectivity of mirror coatings in this spectral range. This region overlaps with the blackbody radiation signature of objects at human body temperatures (Vatansever & Hamblin, 2012). Thus, it

becomes essential to probe alternative modulation mechanisms for harnessing SPI's capabilities in thermal imaging and infrared thermography.

1.2 Objectives

This dissertation aims to provide solutions to overcome challenges associated with sinusoidal beam shaping in FPP and explores the adoption of SPI for LWIR-based thermometric imaging.

First, we propose band-limited illumination profilometry (BLIP), which generates a greyscale sinusoidal fringe pattern on the conjugate plane of DMD from a single binary pattern by a $4f$ imaging system with an optical low-pass filter at the Fourier plane. BLIP's universality permits its melding with varying configurations, resulting in diverse demonstrations.

Aiming for kHz-level real-time acquisition, processing, and display of 3D data during dynamic occurrences, we advocate for the CoaXPress-interfaced BLIP (CI-BLIP). This method employs band-limited illumination to project four sinusoidal fringe patterns and one calibration frame onto the objects at 5 kHz. Augmented by the CoaXPress interface, the rapid image acquisition facilitates a data transmission rate peaking at 25 Gbps, which allows continuous streaming of fringe images to the host computer. Additionally, employing a graphic processing unit (GPU)-based image reconstruction will be used for on-line 3D image reconstruction and display. Overall, CI-BLIP will enable real-time 3D surface information reconstruction at 1 kHz for a comprehensive 60 Hz for a full capture frame. Leveraging these capabilities, CI-BLIP will be used to study the fluid dynamics of a non-rigid flag in 3D, potentially setting the stage for groundbreaking research into high-speed nonlinear flapping dynamics (Argentina & Mahadevan, 2005).

To enhance the imaging capability with kHz-level 3D imaging speed, we envision a dual-view BLIP amalgamated with temporally interlaced acquisition (TIA). This dual-view strategy minimizes the requisite projection patterns by sidelining the calibration frame. TIA's design captures alternating sinusoidal frames with each camera, thereby optimizing data acquisition bandwidth. This optimized bandwidth can either elevate acquisition speed without pixel count alteration or retain the same imaging rate while expanding pixel activation. Coupled with an innovative coordinate-centric image reconstruction algorithm, TIA-BLIP strives to enhance the imaging FOV without sacrificing 3D imaging speed, or alternatively, to boost the 3D imaging speed without diminishing the FOV. We anticipate harnessing TIA-BLIP to analyze dynamic glass behaviors under external forces during unique safety test evaluations.

Furthermore, we seek to endow BLIP with the ability to capture isolated 3D objects, exhibit adaptability across broad reflectance variances, and offer adjustable working distances for extensive FOVs at video speeds. Thus, we propose Multi-scale (MS) BLIP. incorporates dual-intensity projections to navigate reflectance variations in 3D objects. It utilizes a multi-frequency fringe projection technique to discern isolated 3D entities with discernable depth discontinuities. An iterative approach will fine-tune distortion compensations, enhancing the reconstruction fidelity. Our aspirations for MS-BLIP include 3D digitization of intricate art pieces and 3D documentation of human physical activity.

Finally, to apply SPI for precise and efficient human body temperature measurements with passive detection, we propose single-pixel infrared imaging thermometry (SPIRIT). This method focuses on capturing thermal radiation from pivotal regions representing human body temperature, specifically targeting the inner canthi. Rather than employing a DMD, we use a physical mask. For optimal scanning of the encoding matrices, one-dimensional (1D) scanning is favored. To streamline the assembly of masking patterns and reduce scanning procedures, we suggest an aggregated masking assembly combined with data interpolation. We employ diagonally aggregated cyclic S-matrices to create a compressed set of masking patterns. This design allows for a unidirectional scan that sequentially conducts spatial encoding. These masking patterns are then restructured to cater to the dual sub-FOVs centered on the inner canthi. Following this, another lens collects the encoded thermal radiation, directing it to the single-pixel detector. To elevate the signal-to-noise ratio (SNR), both a lock-in amplifier and an optical chopper are incorporated into the system. With calibration, the recovered image undergoes a transformation into a temperature map. We envisage deploying SPIRIT for body temperature screenings and periodic temperature monitoring, aiming to chronicle temperature fluctuations over a day's span.

1.3 Organization

This dissertation is structured as follows:

In Chapter 2, we provide an illustration of our proposed methodologies, complemented by a foundational overview of the principles of FPP and SPI.

In Chapter 3, we introduce the prototype and demonstrations of CI-BLIP, marking the first-time realization of 1kHz 3D imaging capturing fluid-flag interactions. Chapter 4 highlights the enhanced capabilities of TIA-BLIP, boasting kHz-level 3D imaging with an extended FOV. Chapter 5 unfolds the robust, expansive, and video-rate 3D imaging capabilities of MS-BLIP.

In Chapter 6, we present the SPIRIT prototype, accompanied by demonstrations of its utility in human body temperature screening and continuous monitoring.

Chapter 7 encapsulates the conclusions drawn from our research, along with potential future directions.

Finally, Chapter 8 presents a summary of the entire work in French.

1.4 References

Argentina M & Mahadevan L (2005) Fluid-flow-induced flutter of a flag. *Proceedings of the National Academy of Sciences* 102(6):1829-1834.

Avdelidis N & Moropoulou A (2004) Applications of infrared thermography for the investigation of historic structures. *Journal of Cultural Heritage* 5(1):119-127.

Axler S (2015) *Linear algebra done right.* Springer,

Buono MJ, Jechort A, Marques R, Smith C & Welch J (2007) Comparison of infrared versus contact thermometry for measuring skin temperature during exercise in the heat. *Physiological measurement* 28(8):855.

Deetjen ME, Biewener AA & Lentink D (2017) High-speed surface reconstruction of a flying bird using structured light. *Journal of Experimental Biology* 220(11):1956-1961.

Edgar M, Gibson GM, Bowman RW, Sun B, Radwell N, Mitchell KJ, Welsh SS & Padgett MJ (2015) Simultaneous real-time visible and infrared video with single-pixel detectors. *Scientific reports* 5(1):1-8.

Ferrari C, Berlincioni L, Bertini M & Del Bimbo A (2021) Inner eye canthus localization for human body temperature screening. *2020 25th International Conference on Pattern Recognition (ICPR).* IEEE, p 8833-8840.

Flaes DEB, Stopka J, Turtaev S, De Boer JF, Tyc T & Čižmár T (2018) Robustness of light-transport processes to bending deformations in graded-index multimode waveguides. *Physical review letters* 120(23):233901.

Furstenberg R, Kendziora C, Stepnowski J, Stepnowski S, Rake M, Papantonakis M, Nguyen V, Hubler G & McGill R (2008) Stand-off detection of trace explosives via resonant infrared photothermal imaging. *Applied Physics Letters* 93(22).

Gibson GM, Johnson SD & Padgett MJ (2020) Single-pixel imaging 12 years on: a review. *Optics express* 28(19):28190-28208.

Gioux S, Mazhar A & Cuccia DJ (2019) Spatial frequency domain imaging in 2019: principles, applications, and perspectives. *Journal of biomedical optics* 24(7):071613-071613.

Gustafsson MG (2000) Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of microscopy* 198(2):82-87.

Hahamovich E, Monin S, Hazan Y & Rosenthal A (2021) Single pixel imaging at megahertz switching rates via cyclic Hadamard masks. *Nature Communications* 12(1):4516.

Hell SW & Wichmann J (1994) Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics letters* 19(11):780-782.

Horadam KJ (2012) *Hadamard matrices and their applications.* Princeton university press,

Hornbeck LJ (1997) Digital light processing for high-brightness high-resolution applications. *Projection Displays III.* International Society for Optics and Photonics, p 27-40.

Hyun J-S, Li B & Zhang S (2017) High-speed high-accuracy three-dimensional shape measurement using digital binary defocusing method versus sinusoidal method. *Optical Engineering* 56(7):074102.

Kilcullen P, Ozaki T & Liang J (2022) Compressed ultrahigh-speed single-pixel imaging by swept aggregate patterns. *Nature Communications* 13(1):7879.

Leung DW, Chuang IL, Yamaguchi F & Yamamoto Y (2000) Efficient implementation of coupled logic gates for quantum computation. *Physical Review A* 61(4):042310.

Li B, Wang Y, Dai J, Lohry W & Zhang S (2014) Some recent advances on superfast 3D shape measurement with digital binary defocusing techniques. *Optics and Lasers in Engineering* 54:236-246.

Li B & Zhang S (2017) Microscopic structured light 3D profilometry: Binary defocusing technique vs. sinusoidal fringe projection. *Optics and Lasers in Engineering* 96:117-123.

Li B & Zhang S (2018) Novel method for measuring a dense 3d strain map of robotic flapping wings. *Measurement Science and Technology* 29(4):045402.

Liang J (2020) Punching holes in light: recent progress in single-shot coded-aperture optical imaging. *Reports on Progress in Physics* 83(11):116101.

Liang J, Gao L, Li C & Wang LV (2014) Spatially Fourier-encoded photoacoustic microscopy using a digital micromirror device. *Optics letters* 39(3):430-433.

Oppenheim AV, Willsky AS, Nawab SH & Ding J-J (1997) *Signals and systems.* Prentice hall Upper Saddle River, NJ,

Parthiban V, Kohn Jr RN, Liang J & Becker MF (2016) Experimental demonstration of precise holograms using complex light modulation. *Emerging Digital Micromirror Device Based Systems and Applications VIII.* International Society for Optics and Photonics, p 97610M.

Ping L, Leung W & Wu KY (2003) Low-rate turbo-Hadamard codes. *IEEE Transactions on Information Theory* 49(12):3213-3224.

Rodenburg B (2015) *Communicating with transverse modes of light.* University of Rochester,

Sarawade AA & Charniya NN (2018) Infrared thermography and its applications: a review. *2018 3rd International conference on communication and electronics systems (ICCES).* IEEE, p 280-285.

Saxena M, Eluru G & Gorthi SS (2015) Structured illumination microscopy. *Advances in Optics and Photonics* 7(2):241-275.

Sloane NJ & Harwit M (1979) Hadamard Transform Optics. (Academic Press New York).

Stantchev RI, Yu X, Blu T & Pickwell-MacPherson E (2020) Real-time terahertz imaging with a single-pixel detector. *Nature communications* 11(1):2535.

Stein EM & Shakarchi R (2011) *Fourier analysis: an introduction.* Princeton University Press,

Su X & Zhang Q (2010) Dynamic 3-D shape measurement method: a review. *Optics and Lasers in Engineering* 48(2):191-204.

Tao C, Zhu H, Zhang Y, Luo S, Ling Q, Zhang B, Yu Z, Tao X, Chen D & Li Q (2022) Shortwave infrared single-pixel spectral imaging based on a GSST phase-change metasurface. *Optics Express* 30(19):33697-33707.

Vatansever F & Hamblin MR (2012) Far infrared radiation (FIR): its biological effects and medical applications. *Photonics Lasers Med* 4:255-266.

Wang Y, Huang K, Fang J, Yan M, Wu E & Zeng H (2023) Mid-infrared single-pixel imaging at the single-photon level. *Nature Communications* 14(1):1073.

Wehlburg CM, Wehlburg JC, Gentry SM & Smith JL (2001) Optimization and characterization of an imaging Hadamard spectrometer. *Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VII.* SPIE, p 506-515.

Xu J & Zhang S (2020) Status, challenges, and future perspectives of fringe projection profilometry. *Optics and Lasers in Engineering* 135:106193.

Yarlagadda RK & Hershey JE (2012) *Hadamard matrix analysis and synthesis: with applications to communications and signal/image processing.* Springer Science & Business Media,

Yue J, Han J, Li L & Bai L-f (2018) Denoising analysis of spatial pixel multiplex coded spectrometer with Hadamard H-matrix. *Optics Communications* 407:355-360.

Zhang S (2018) High-speed 3D shape measurement with structured light methods: A review. *Optics and Lasers in Engineering* 106:119-131.

Zhang Z, Wang X, Zheng G & Zhong J (2017) Hadamard single-pixel imaging versus Fourier single-pixel imaging. *Optics Express* 25(16):19619-19639.

Zuo C, Feng S, Huang L, Tao T, Yin W & Chen Q (2018) Phase shifting algorithms for fringe projection profilometry: A review. *Optics and Lasers in Engineering* 109:23-59.

2 METHODS

In this chapter, we outline the methodologies of two orthogonality-based beam shaping techniques: sinusoidal fringe intensity modulation as band-limited illumination profilometry (BLIP) and cyclic S-matrices spatial encoding as single-pixel infrared imaging thermometry (SPIRIT). Alongside, we provide foundational insights into their respective imaging systems: fringe projection profilometry (FPP) using sinusoidal fringe illumination and the single-pixel imaging (SPI) system that leverages cyclic S-matrices encoding.

2.1 Band-limited illumination profilometry (BLIP)

2.1.1 Basic principle of fringe projection profilometry

Geometry of the triangulation method:

Figure 2.1 shows the schematic diagram of a typical 3D shape measurement system based on FPP. The projector illuminates the object with fringe patterns, which are deformed by the geometry of the object's surface. A camera captures the deformed structured images from another perspective. In such a system, the correspondence between the camera pixel C and the phase line coordinate A is established by analyzing the deformation of captured structured images with known features (e.g., a single phase line) projected by the projector. Once the correspondence is obtained by fringe analysis (detailed in Section 2.1.3), with the system calibration information, the world coordinates (x, y, z) of object point P can be reconstructed using triangulation.

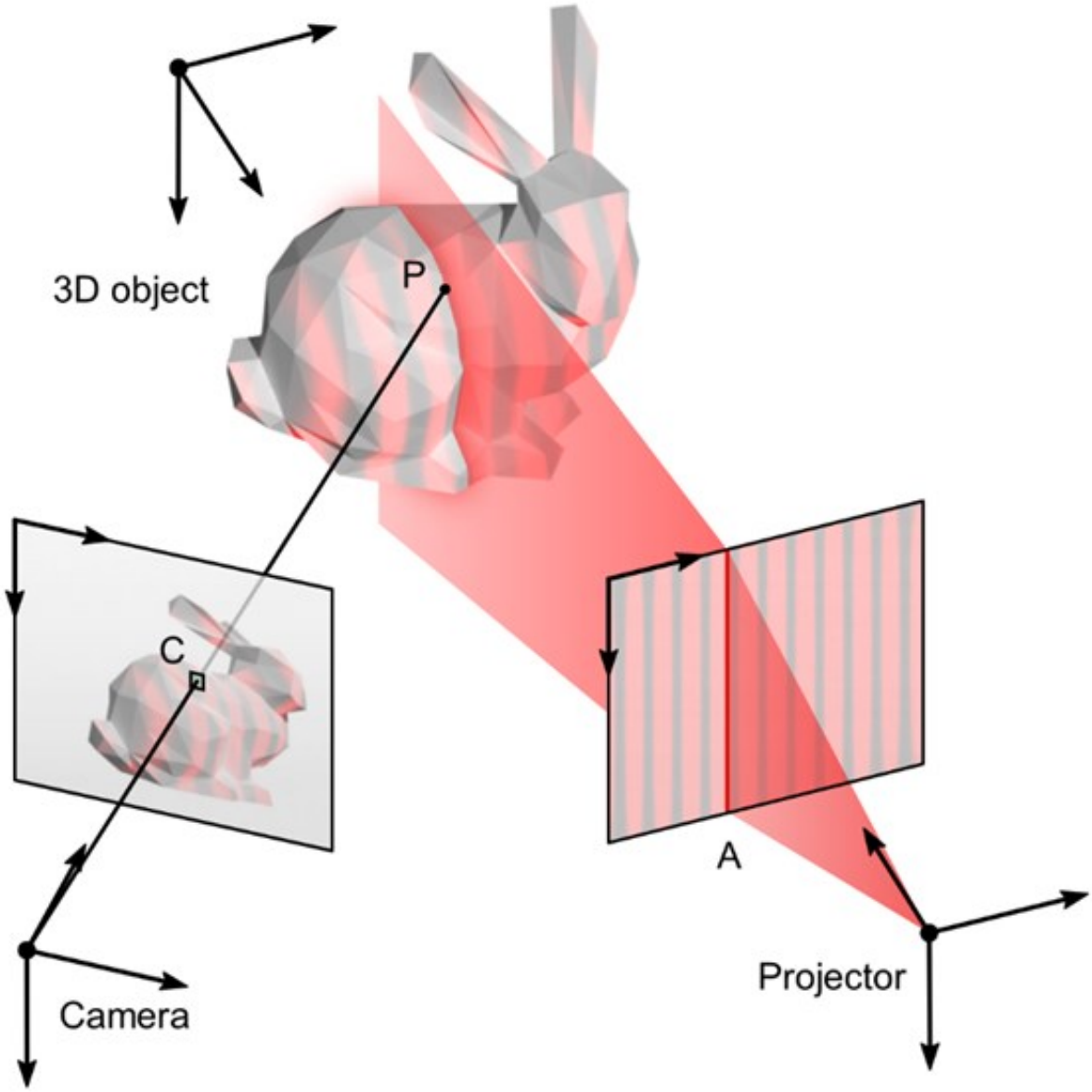

Figure 2.1 Schematic diagram of a typical FPP system.

System modeling:

The transformation from the world coordinate system to the camera/projector coordinate system can be mathematically described by the pinhole model, which provides the solution for the triangulation method. For a 3D point (x, y, z) in the world coordinate system, in a perfect imaging condition, the pinhole model describes its corresponding pixel on the camera sensor, (u_c, v_c), to be

$$s_{\mathrm{c}}\begin{bmatrix}u_{\mathrm{c}}\\ v_{\mathrm{c}}\\ 1\end{bmatrix}=A_{\mathrm{c}}[R_{\mathrm{c}}\ T_{\mathrm{c}}]\begin{bmatrix}x\\ y\\ z\\ 1\end{bmatrix}=\begin{bmatrix}f_{cu} & 0 & u_{\mathrm{cp}}\\ 0 & f_{cv} & v_{\mathrm{cp}}\\ 0 & 0 & 1\end{bmatrix}[R_{\mathrm{c}}\ T_{\mathrm{c}}]\begin{bmatrix}x\\ y\\ z\\ 1\end{bmatrix}. \tag{2.1}$$

Overall, this operation can be expressed by $(u_{\mathrm{c}}, v_{\mathrm{c}}) = \mathrm{Proj}_{\mathrm{c}}\ (x, y, z)$. Akin to the pinhole camera model, the projective behavior of the projector can also be described based on the corresponding projector pixel $(u_{\mathrm{p}}, v_{\mathrm{p}})$ as

$$s_{\mathrm{p}}\begin{bmatrix}u_{\mathrm{p}}\\ v_{\mathrm{p}}\\ 1\end{bmatrix}=A_{\mathrm{p}}\big[R_{\mathrm{p}}\ T_{\mathrm{p}}\big]\begin{bmatrix}x\\ y\\ z\\ 1\end{bmatrix}=\begin{bmatrix}f_{\mathrm{p}u} & 0 & u_{\mathrm{pp}}\\ 0 & f_{\mathrm{p}v} & v_{\mathrm{pp}}\\ 0 & 0 & 1\end{bmatrix}\big[R_{\mathrm{p}}\ T_{\mathrm{p}}\big]\begin{bmatrix}x\\ y\\ z\\ 1\end{bmatrix}. \tag{2.2}$$

Similarly, the projection of the 3D point to the projector is modeled by $(u_{\mathrm{p}}, v_{\mathrm{p}}) = \mathrm{Proj}_{\mathrm{p}}\ (x, y, z)$. In Eqs. (2.1)–(2.2), the matrix A_{c} and A_{p} contain the intrinsic parameters of the camera and projector: (f_{cu}, f_{cv}) and $(f_{\mathrm{p}u}, f_{\mathrm{p}v})$ describe the effective focal lengths along the axes of the camera and projector; $(u_{\mathrm{cp}}, v_{\mathrm{cp}})$ and $(u_{\mathrm{pp}}, v_{\mathrm{pp}})$ are the coordinates of the principal points of the camera and projector, separately. Defined as the extrinsic parameters of the camera and projector, R_{c} and R_{p} are 3×3 matrices accounting for rotation. T_{c} and T_{p} are 3×1 matrices for translation. Finally, s_{c} and s_{p} are scalar factors for numerical extraction of $(u_{\mathrm{c}}, v_{\mathrm{c}})$ and $(u_{\mathrm{p}}, v_{\mathrm{p}})$.

The intrinsic and extrinsic parameters can be extracted by system calibration, which will be described in detail in Section ***system calibration***. Here, we define $A_{\mathrm{c}}[R_{\mathrm{c}}\ T_{\mathrm{c}}] = \begin{bmatrix}a_{11} & \cdots & a_{14}\\ \vdots & \ddots & \vdots\\ a_{31} & \cdots & a_{34}\end{bmatrix}$ and $A_{\mathrm{p}}\big[R_{\mathrm{p}}\ T_{\mathrm{p}}\big] = \begin{bmatrix}b_{11} & \cdots & b_{14}\\ \vdots & \ddots & \vdots\\ b_{31} & \cdots & b_{34}\end{bmatrix}$. Once the correspondence between the camera pixel $(u_{\mathrm{c}}, v_{\mathrm{c}})$ and one of the projector phase line's coordinates (i.e., u_{p} or v_{p}) is established, the world coordinates of 3D point P (i.e., $[x, y, z]^T$) can be calculated as (Jiang ***et al.***, 2022)

$$\begin{aligned}\begin{bmatrix}x\\ y\\ z\end{bmatrix}&=\mathrm{Tri}\big(u_{\mathrm{c}}, v_{\mathrm{c}}, u_{\mathrm{p}}\big)\\ &=\begin{bmatrix}a_{11}-u_{\mathrm{c}}a_{31} & a_{12}-u_{\mathrm{c}}a_{32} & a_{13}-u_{\mathrm{c}}a_{33}\\ a_{21}-v_{\mathrm{c}}a_{31} & a_{22}-v_{\mathrm{c}}a_{32} & a_{23}-v_{\mathrm{c}}a_{33}\\ b_{11}-u_{\mathrm{p}}b_{31} & b_{12}-u_{\mathrm{p}}b_{32} & b_{13}-u_{\mathrm{p}}b_{33}\end{bmatrix}^{-1}\begin{bmatrix}u_{\mathrm{c}}a_{34}-a_{14}\\ v_{\mathrm{c}}a_{34}-a_{24}\\ u_{\mathrm{p}}b_{34}-b_{14}\end{bmatrix}.\end{aligned} \tag{2.3}$$

System calibration:

Accurate system calibration, which calculates the intrinsic and extrinsic parameters of the camera and projector, is crucial to 3D shape measurements [i.e., by solving Eq. (2.3)]. In a typical process, a flat calibration plane with known feature points (e.g., checkerboard) is used for camera

calibration. The flat checkerboard positioned with different poses is imaged by the camera. All captured checkerboard images are processed using the open-source Matlab toolbox to extract the grid corners [Figs. 2.2(a) and (b)] and calculate the camera's calibration parameters (Bouguet, 2004).

The calibration of the projector is conventionally difficult and complex because the projector cannot directly capture images like a camera. Zhang and Huang (Zhang & Huang, 2006) developed a method that is now extensively adopted. This method enables the projector to capture images like a camera so that the complicated projector calibration problem becomes a well-established camera calibration problem. In particular, a phase-shifting method (detailed in Section 2.1.3) is implemented to establish the correspondence between the camera pixels to the projector pixels, which is then used to transform the camera-captured checkerboard images into projector-captured images. As displayed in Fig. 2.2(c), both vertical and horizontal sets of fringe patterns with their corresponding center line patterns are projected onto the checkerboard plane. Estimation of the absolute phase from these images is carried out using the four-step phase-shifted algorithm. Then, the absolute phase maps extracted for both the horizontal and vertical directions are used to determine a pixel-wise mapping from a camera-captured image of the checkerboard plane [Fig. 2.2(d)] to a correctly altered image representing the view of the checkerboard plane from the perspective of the projector [Fig. 2.2(e)]. Finally, the same MATLAB toolbox used in camera calibration is implemented to compute the corresponding calibration parameters of the projector.

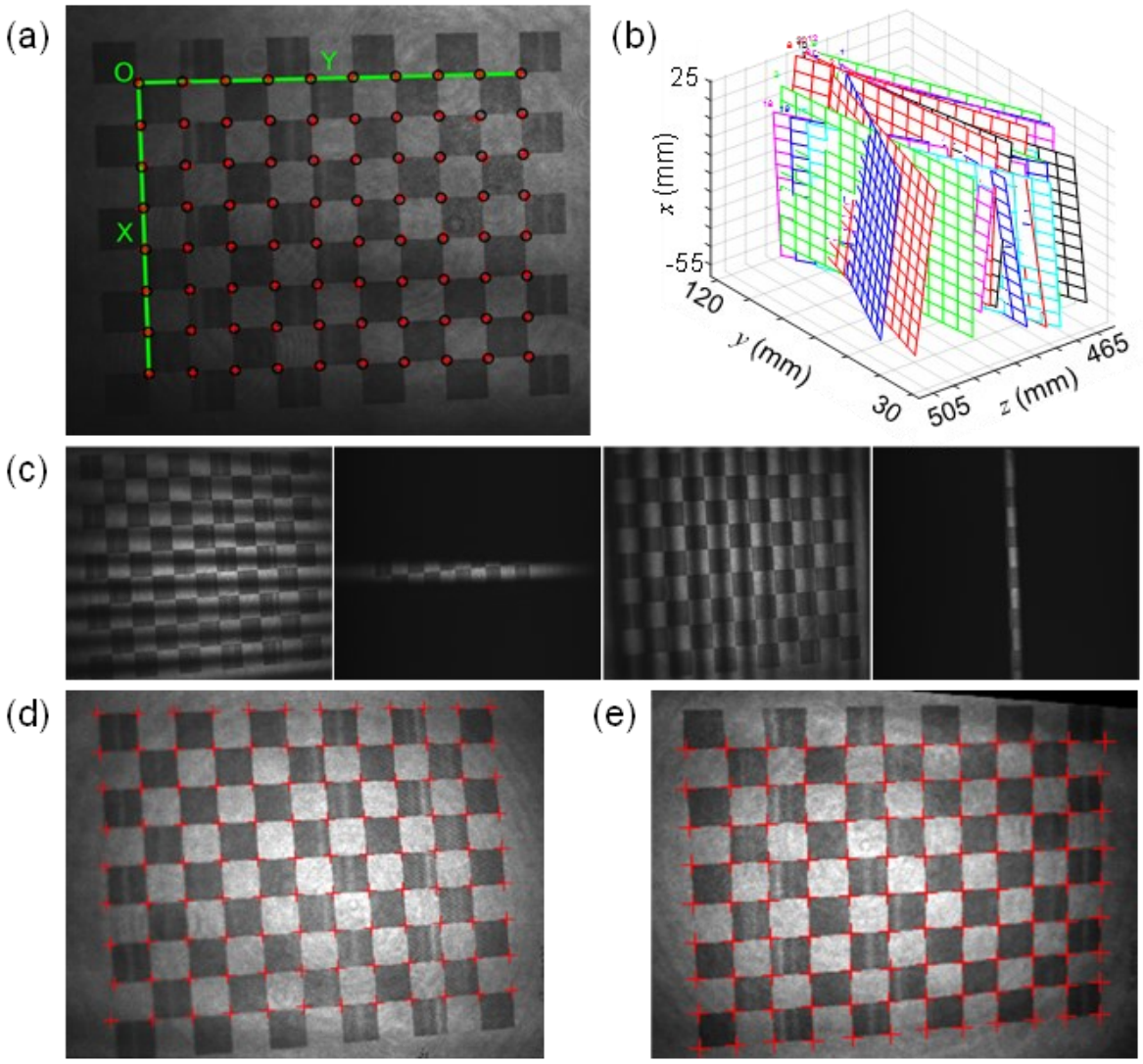

Figure 2.2 System calibration.

(a) Grid corners extraction of a selected camera-captured image. "O" represents the selected corner point; "X" and "Y" represent the 2D coordinates; red points represent the software-extracted grid corners. (b) Camera calibration results show all the positions of the checkerboard poses. (c) Images of a selected horizontal fringe pattern, the horizontal centerline, a selected vertical fringe pattern, and the vertical centerline. (d) Camera image with selected grids. (e) "Projector image" with the corresponding grids.

2.1.2 Band-limited illumination

To facilitate a high-speed sinusoidal fringe projection with high quality and stability, we introduced the band-limited illumination. This approach utilizes a $4f$ imaging system to transform a single binary pattern displayed on the DMD into a greyscale sinusoidal fringe pattern on its conjugate plane. Crucially, this transformation is facilitated by an optical low-pass filter positioned at the Fourier plane of the imaging system.

System setup:

The schematic of the band-limited illumination is shown in Fig. 2.3(a). After expansion and collimation of the light source (laser with a wavelength of λ), the illumination beam is directed to a 0.45" DMD (AJD-4500, Ajile Light Industries) at an incident angle of ~24° to its surface normal. A band-limited 4f imaging system that consists of two lenses [Lens1 and Lens2 in Fig. 2.3(a), with focal lengths of f_1 and f_2] and one pinhole converts the binary patterns loaded on the DMD to grayscale fringes at the intermediate image plane. Then, a camera lens projects these fringe patterns onto a 3D object.

Operating principle:

An error diffusion (ED) algorithm (Chung *et al.*, 2011)—a halftoning technique—is used to design the binary patterns that are loaded onto the DMD. The spatial-frequency components of beam profiles from the DMD consist of the target sinusoidal fringe spectrum as well as a high-pass filtered error spectrum (Knox, 1992). ED acts as a high-pass filter, exhibiting 'blue noise' characteristics in the spatial frequency domain due to its quasi-random distribution of binary values. This blue noise causes minimal distortion in low to mid spatial frequencies but introduces significant errors at higher frequencies. This suggests that ED is ideal for producing band-limited target images.

The sinusoidal pattern, with a period of p_{f} DMD pixels [top image of Fig. 2.3(b)] is used as the input image. The grayscale pattern is processed by the ED algorithm to produce one binary DMD pattern pixel by pixel from left to right in a row and then from top to bottom row by row. For a specific DMD pixel, its binary value is determined by a dynamic threshold that depends on the gradient magnitude of the current pixel in the grayscale image. Then, the binarization error of the processed pixel in the grayscale image is diffused to the neighboring pixels based on specific weights, which results in the intensity change of the surrounding pixels accordingly. By raster processing, the binary DMD pattern is obtained [bottom image of Fig. 2.3(b)].

The pinhole works as a low-pass filter to control the optical bandwidth. Its diameter, determined by the system bandwidth, is calculated as

$$D = \frac{\lambda f_1}{P_{\mathrm{p}}} \quad , \tag{2.4}$$

where P_{p} denotes the fringe period. For example, the projected fringe pattern has a period of 60 DMD pixels, while each DMD pixel has a dimension of 10.8 μm, yielding a $P_{\mathrm{p}} = 648$ μm. Thus, the

required pinhole diameter is D = 124.26 μm. In the experiment, a 150-μm-diameter pinhole is used to ensure all spatial frequency content of the sinusoidal fringe pattern passing through the system.

The generated binary DMD patterns possess blue noise characteristics in the spatial frequency domain (Liang *et al.*, 2012), which is manifested by precisely matched imaging content to the corresponding grayscale pattern within the system's bandwidth [Fig. 2.3(c)]. To filter high-spatial-frequency noise, we place the pinhole as an optical low-pass filter at the Fourier plane in the $4f$ imaging system (Liang *et al.*, 2010). Examined at the intermediate imaging plane, the resultant beam profiles are high-quality grayscale sinusoidal fringes [Fig. 2.3(d)], with a root-mean-square error of 5.71% with respect to the target pattern.

As a result, in BLIP, regardless of their frequencies, sinusoidal fringe patterns are always generated on the image plane of the DMD. Compared to binary defocusing, this configuration eliminates the distortion brought by uneven DMD surfaces and maintains the contrast of projected fringe patterns. It is also inherently compatible with multi-frequency fringe projection, varying working distances, and different fields of view.

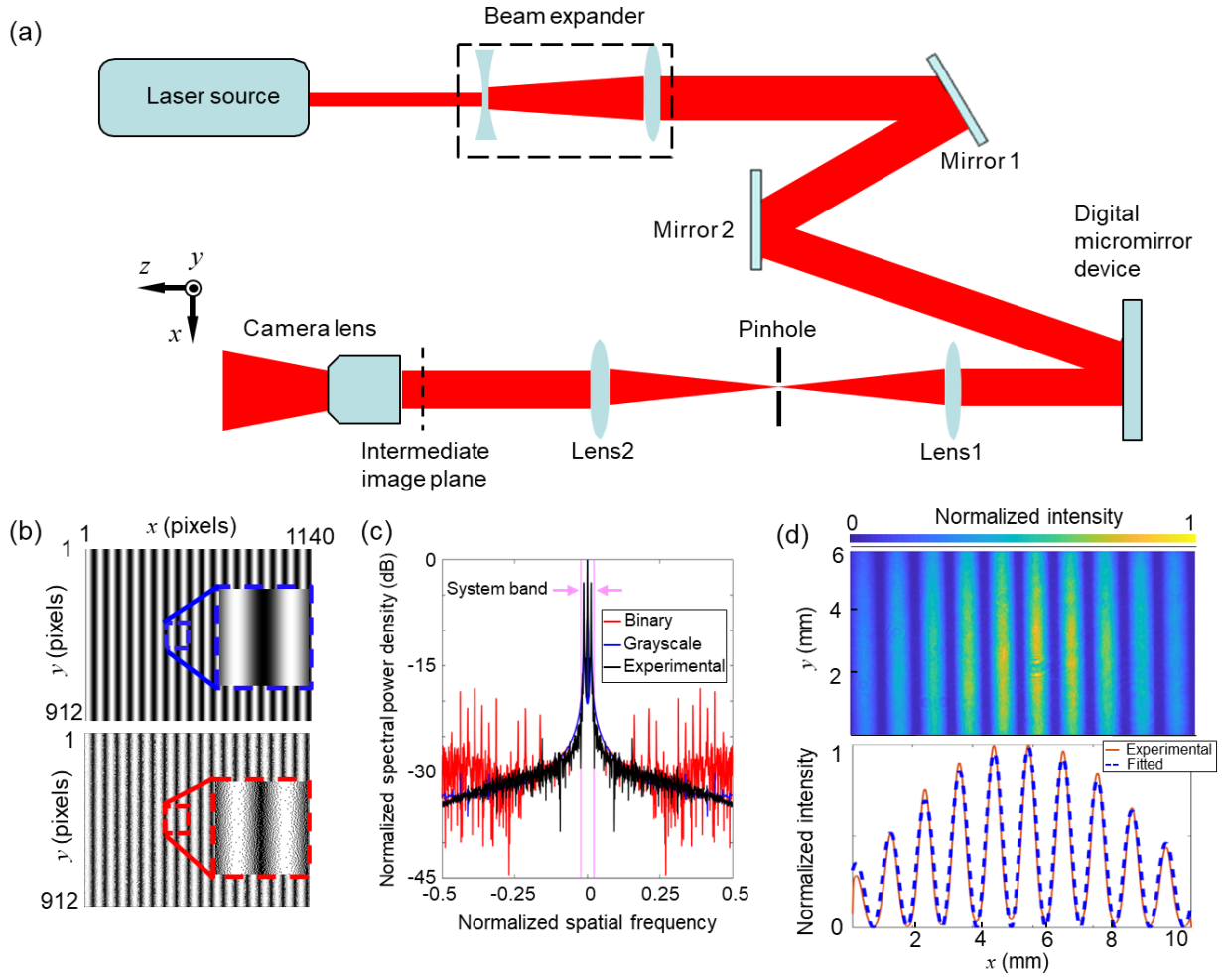

Figure 2.3 Band-limited illumination.

(a)System schematic. (b) Target fringe pattern (top) and its corresponding binary DMD pattern (bottom). Close-up views show the grayscale and binary characteristics of the target pattern and the DMD pattern, respectively. (c) Comparison of the central cross-sections of the spectral power density of the target sinusoidal pattern, the corresponding binary DMD pattern, and the experimentally captured pattern. (d) Top panel: Grayscale sinusoidal pattern generated experimentally by band-limited illumination. Bottom panel: The averaged cross-section from the top panel with a fitted result.

2.1.3 Sinusoidal fringe analysis

After the fringe projection, the projected patterns are deformed by the geometry of the object's surface. The deformed structured patterns are then captured by the camera. Based on Eq. (2.3), the 3D reconstruction of an object's surface relies on the determination of (u_c, v_c, u_p). Thus, the following step is to extract the correspondence between the camera pixel (u_c, v_c) and the projector coordinate u_p. BLIP extracts the phase for 3D reconstruction. The fringe analysis of sinusoidal patterns involves two steps: wrapped phase extraction and phase unwrapping.

Wrapped phase extraction:

There are two major techniques for wrapped phase extraction: Fourier transform profilometry (FTP) and phase-shifting fringe projection profilometry (PSFPP). Theoretically, FTP can use one fringe pattern to calculate the pixel-wise phase value with the assistance of a spatial carrier (Zhang *et al.*, 2019; Zuo *et al.*, 2018b). The phase is extracted by applying a properly designed band-pass filter in the spatial frequency domain. The small number of projected patterns (typically, less than three) of FTP makes it suitable for the 3D shape measurement of dynamic surfaces (Mao *et al.*, 2007; Su & Chen, 2001). The FTP method works well with a smooth object surface profile and low fringe noise. However, in general, FTP still struggles to handle complex surface reconstruction and high noise. To eliminate the weakness, PSFPP has been developed to capture more fringe patterns with phase shifting (Huang & Zhang, 2006; Zhang & Yau, 2006; Zuo *et al.*, 2018a). The multiple-shot nature of PSFPP makes it more robust and can achieve pixel-wise phase measurement with higher resolution and accuracy. Furthermore, the PSFPP measurements are more resilient to non-uniform background intensity and fringe modulation.

BLIP employs PSFPP as a universal scheme for wrapped phase extraction. The intensity of the camera pixel (u_c, v_c) in the captured deformed structure images can be expressed as

$$I_k(u_c, v_c) = I_b(u_c, v_c) + I_{va}(u_c, v_c)\cos[\varphi(u_c, v_c) - 2\pi k/K]. \tag{2.5}$$

Here, $I_k(u_c, v_c)$ represents the intensity in the k^{th} deformed structure image, K is the number of phase-shifting steps, and $k \in [0, K-1]$, $I_b(u_c, v_c)$, $I_{va}(u_c, v_c)$, and $\varphi(u_c, v_c)$ represent the background, intensity variation, and depth-dependent phase, respectively. Based on the phase-shifting algorithm, the relative phase value at each pixel can be calculated as

$$\varphi(u_c, v_c) = \tan^{-1}\left(\frac{\sum_{k=0}^{K-1} I_k(u_c, v_c)\sin\left(\frac{2\pi k}{K}\right)}{\sum_{k=0}^{K-1} I_k(u_c, v_c)\cos\left(\frac{2\pi k}{K}\right)}\right). \tag{2.6}$$

Equation (2.6) extracts the wrapped phase map in the range of $(-\pi, +\pi]$. Since there are three unknown factors I_b, I_{va} and φ, three phase-shifted fringe patterns are sufficient to calculate the wrapped phase value.

Phase unwrapping:

From Eq. (2.6), it can be seen that the wrapped phase map is not available for converting the phase value to a specific projector's column coordinate pixel by pixel. Thus, the phase obtained from fringe analysis techniques is required to be unwrapped before 3D coordinates conversion (Judge & Bryanston-Cross, 1994). This process aims to determine the fringe order $n(u_c, v_c)$—an

integer number for each pixel—to remove the 2π discontinuities. The unwrapped phase value can be calculated by

$$\tilde{\varphi}(u_{\mathrm{c}}, v_{\mathrm{c}}) = \varphi(u_{\mathrm{c}}, v_{\mathrm{c}}) + n(u_{\mathrm{c}}, v_{\mathrm{c}}) \times 2\pi. \tag{2.7}$$

Depending on the specific needs of various applications, the geometric-constraint-based phase unwrapping (discussed in Section 4.2.3), temporal phase unwrapping (as elaborated in Section 5.2.2), and spatial phase unwrapping were utilized for the three distinct demonstrations of BLIP. A comprehensive explanation of spatial phase unwrapping is provided below.

A spatial phase unwrapping method unwraps the phase by referring to the phase values of neighboring points on a single-wrapped phase map. It assumes the surface geometry is smooth, which means that there are no more than 2π phase changes between two adjacent points along the path (Bone, 1991; Ghiglia & Pritt, 1998). Using this principle, spatial phase unwrapping is carried out from one pixel to its neighboring pixels and through the entire image. Supposing that $(u_{\mathrm{c}1}, v_{\mathrm{c}1})$ and $(u_{\mathrm{c}2}, v_{\mathrm{c}2})$ are two adjacent pixels on the path, then the fringe order $n(u_{\mathrm{c}2}, v_{\mathrm{c}2})$ is calculated by

$$n(u_{\mathrm{c}2}, v_{\mathrm{c}2}) = \begin{cases} n(u_{\mathrm{c}1}, v_{\mathrm{c}1}) - 1, & \text{if } \Delta\varphi = 2\pi \\ n(u_{\mathrm{c}1}, v_{\mathrm{c}1}) + 1, & \text{if } \Delta\varphi = -2\pi \\ n(u_{\mathrm{c}1}, v_{\mathrm{c}1}), & \text{otherwise} \end{cases}, \tag{2.8}$$

where $\Delta\varphi = \varphi(u_{\mathrm{c}2}, v_{\mathrm{c}2}) - \varphi(u_{\mathrm{c}1}, v_{\mathrm{c}1})$.

Due to the noise and the sharp depth change, the wrapped phase difference $\Delta\varphi$ may not exhibit 2π steps, which limits the practical robustness of this approach. Several modified spatial phase unwrapping algorithms have been developed to enhance the unwrapping accuracy (Chen *et al.*, 2013; Flynn, 1997; Ghiglia & Romero, 1994). They identify and separate the objects' surface into high-quality areas (with smooth phase variance) and low-quality areas (with more abrupt phase changes). Phase unwrapping is then carried out from higher-quality phase points to lower-quality phase points to reduce the probability of incorrectly unwrapped phase points.

After obtaining the unwrapped phase map, an additional image, which is captured by illumination with a center line on the projector, is used to provide a phase offset value φ_0 corresponding to the projector's center (Jiang *et al.*, 2020; Zhang & Huang, 2006). Then, the horizontal projector coordinate u_{p} can be obtained using the absolute phase value $\tilde{\varphi}(u_{\mathrm{c}}, v_{\mathrm{c}})$ as

$$u_{\mathrm{p}} = (\tilde{\varphi}(u_{\mathrm{c}}, v_{\mathrm{c}}) - \varphi_0)\frac{P_{\mathrm{p}}}{2\pi} + \frac{1}{2}(N_u - 1), \tag{2.9}$$

where N_u is the width of the projector (in pixels). Once u_p is determined with respect to specific camera pixel $(u_\mathrm{c}, v_\mathrm{c})$, the 3D coordinate (x, y, z) can be extracted by Eq. (2.3).

2.1.4 Quantifications of BLIP

Depth resolution:

To quantify the depth resolution of BLIP, two stacked planar surfaces offset by ~9° [Fig. 2.4(a)] are imaged. In this configuration, the depth difference between the two surfaces along the x axis increased monotonously, starting from a negative maximum at the right edge, to zero at the center, and reaching a positive maximum at the left edge. The reconstructed image of the 3D object [Fig. 2.4(b)] allowed analyzing cross-sections of depth at different x positions. The depth difference (denoted by z_d) of the two surfaces is calculated, which closely agrees to the ground truth [Fig. 2.4(c)]. In addition, standard deviations of the measured depths were used as the system's noise level. The depth resolution was defined as when z_d equals two times the system's noise level. The depth resolutions mentioned in Chapter 3-5 are calculated based on the same method as described above.

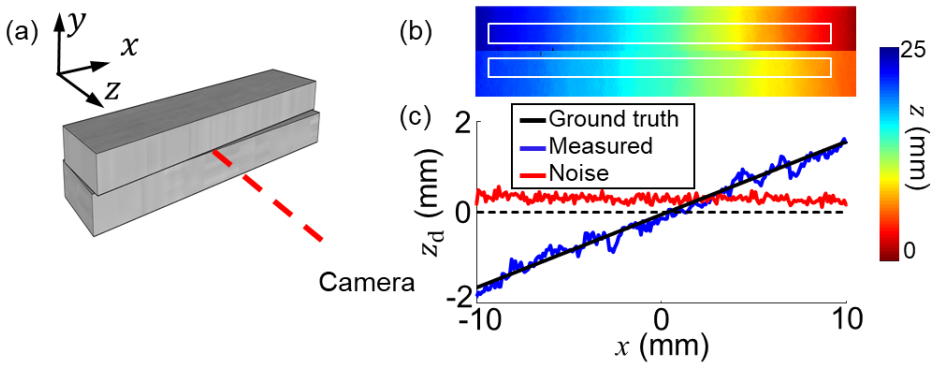

Figure 2.4 Quantification of depth resolution.

(a) Experimental setup. (b) 3D images of the planar surfaces (top image) and measured depth difference (bottom plot) of CI-BLIP. The white boxes represent the selected regions for analysis.

Lateral resolution:

The lateral resolution of BLIP is mainly determined by the numerical aperture of the camera lens and the SNR in the measurement. The lateral resolution is quantified by imaging the sharp edge of a planar surface. The reconstructed images are shown in Fig. 2.5(a). The edge spread function (ESF) is calculated by averaging along each line parallel to the x axis [Fig. 2.5(b)]. Taking the derivative of the ESF yielded the line spread function [LSF, Fig. 2.5(c)], which was further

interpolated with a cubic spline. The lateral resolution as defined by the full-width at half-maximum (FWHM) of the interpolated LSF.

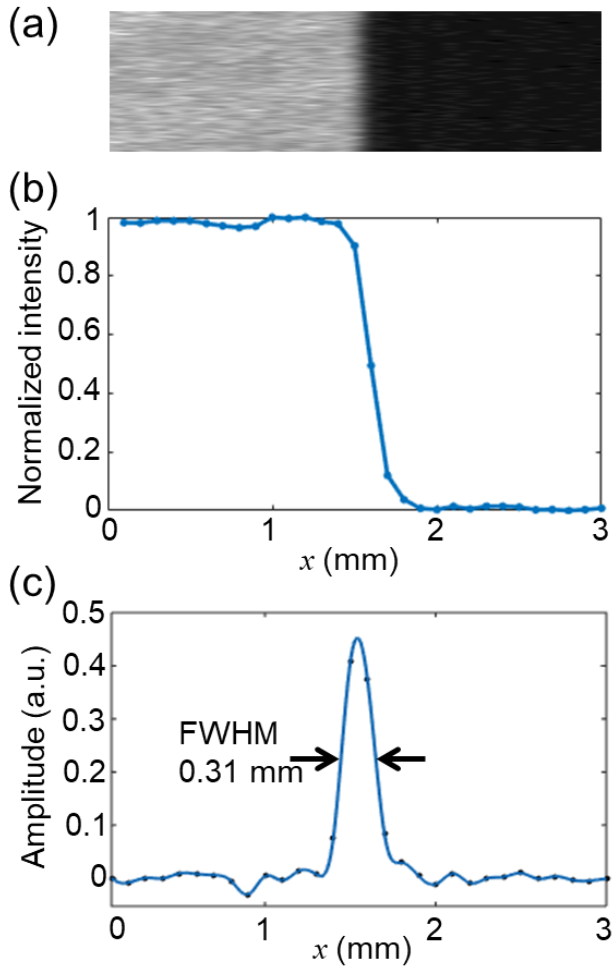

Figure 2.5 Quantification of lateral resolution.

(a) Reconstructed image of the edge of a planar surface. (b) ESF extracted from (a). (c) LSF extracted from ESF and interpolated with cubic spline.

Field of view and depth of view:

The field of view (FOV) in BLIP refers to the overlap region between the projection and detection, the portion of space that is both captured by the camera and within the area illuminated or covered by the projection. It is determined by the focal length of both the projection lens and the camera lens, as well as the size of the camera sensor.

The depth of field (DOF) in BLIP is used to describe the range of distance within the projected greyscale fringe patterns that appear acceptably sharp. The factors that determines

DOF include system's numerical aperture, focal length of the imaging lens, as well as the distance to the object. Specifically, the deterministic criteria of the acceptable shapness in BLIP is the fringe contrast, which essentially limits the accuracy of phase extraction as well as the 3D reconstruction. However, the accuracy of phase extraction, can be otherwise lifted up by increase the number of phase-shifting steps, which leases the restriction to the fringe contrast by sacrificing the 3D imaging speed. Thus, there is a trade-off between the DOF and the 3D imaging speed in BLIP. Specific requirements posed by different applications will determine the system optimization with the maxim DOF at acceptable 3D imaging speed and accuracy.

2.2 Single-pixel infrared imaging thermometry (SPIRIT)

2.2.1 Basic principle of single-pixel imaging

SPI encodes deterministic two-dimensional (2D) spatial patterns onto the imaging beam, with subsequent data acquisition performed by a single-pixel detector. During the subsequent process of image reconstruction, these encoding schemes are utilized as prior knowledge, aiding in the accurate retrieval of spatial information. Figure 2.6 shows the schematic diagram of a typical SPI system with an active illumination mode. Each spatial modulated pattern (masking pattern) s and the base image x are real-valued, discrete images of rectangular size $p \times q$. Consequently, they can be perceived as vectors in $\mathbb{R}^n$ with $n = p \times q$. In SPI, each detector measurement (also known as a 'bucket signal') $y \in \mathbb{R}$, corresponding to x, is interpreted as an optically computed inner product between the pixel elements of the base image vector x and a known measurement vector s. Hence, a series comprising m measurements therefore be specified by a measurement matrix $S \in \mathbb{R}^{m \times n}$, which describes the entire sequence of deterministic masking patterns deployed in the course of an imaging experiment. The rows of this matrix are constituted by the masking pattern vectors s_i for $1 \leq i \leq m$. Expressing this in matrix-vector notation yields

$$\mathbf{y} = Sx, \tag{2.10}$$

where $\mathbf{y} \in \mathbb{R}^m$ is the vector composed of the individual measurements y_i.

In the case that $m = n$, it is clear that the signal $\mathbf{y}$ can be obtained from Eq. (2.10) uniquely if the m rows of S span $\mathbb{R}^n$ (and thus S^{-1} exists). However, for the case that $m \leq n$, the linear system is under-determined and thus there exist infinitely many solution images consistent with the observations $\mathbf{y}$ via Eq. (2.10). Solving Eq. (2.10) for a specific measurement matrix design comprises a necessary and often non-trivial reconstruction step, which has lead to several approaches in both the choice of S and the computational method used to recover x.

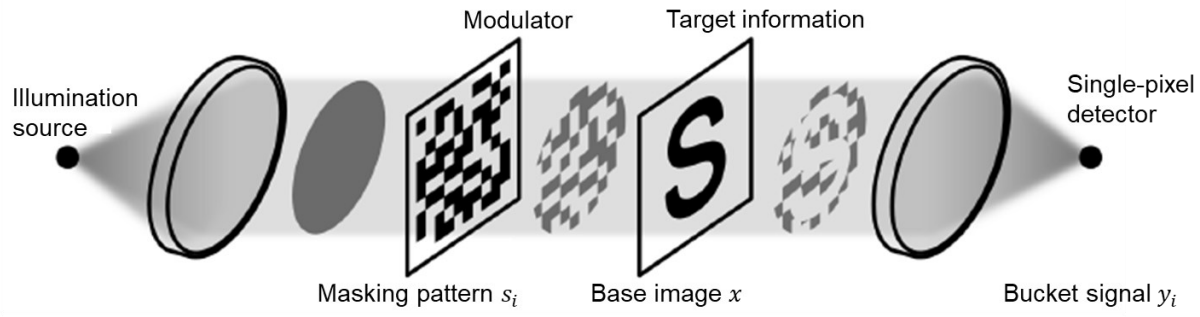

Figure 2.6 Illustration of a typical SPI system.

A lens first collimates a point source of light, directing it onto the modulator and subsequently illuminating the target. This light is then converged onto a single-pixel detector. The imaging beam's transverse profile at various segments is depicted in dark gray.

Noise presentation in SPI:

In the presence of noise, an important consideration for the design of a measurement matrix is its ability to suppress the mean squared error (MSE) of the recovered images (Sloane & Harwit, 1979). The measurements can be modified by adding a vector representing random errors

$$\mathbf{y} = Sx + N, \tag{2.11}$$

where N is an m-vector of uniform zero-mean Gaussian noise with variance σ^2. With x_{N} as a subsequently recovered image, the MSE can be defined as

$$\eta = Ave\{(x_{\mathrm{N}} - x)^T(x_{\mathrm{N}} - x)\}, \tag{2.12}$$

where $Ave\{\}$ represents an average taken over several experiments. In the simplest case where $m = n$ and S is invertible, the reconstructed image x_{N} can be recovered as

$$x_{\mathrm{N}} = S^{-1}\mathbf{y}. \tag{2.13}$$

It can be shown that, Eq. (2.13) provides an unbiased estimate of the underlying image in the sense that $Ave\{x_{\mathrm{N}}\} = x$. In this case, it can also be shown that the MSE becomes

$$\frac{\eta}{\sigma^2} = Tr((S^{-1})^T S^{-1}), \tag{2.14}$$

where $Tr(\,)$ represents the matrix trace (Sloane & Harwit, 1979). As a result, the noise performance of an SPI system depends strongly on the choice of measurement matrix S, with those matrices possessing smaller values of $Tr((S^{-1})^T S^{-1})$ being the most desired.

Hadamard matrices and S-matrices:

Due to the fact that masking patterns used in SPI systems must correspond to intensity modulations of an optical imaging beam, it is required that measurement matrices be comprised of entries with magnitudes not exceeding one. An important mathematical result is that those matrices S which possess entries satisfying $|S_{ij}| \leq 1$ and which minimize the MSE ratio of Eq. (2.14) and are uniquely the Hadamard matrices (Sloane & Harwit, 1979), which are defined as $\{+1, -1\}$-valued square matrices H of size $n \times n$ satisfying the property $HH^T = nI_n$. This result has been widely recognized, with the noise-favorable properties of Hadamard matrices motivating their use not only in many SPI systems, but also in the design of various spectrometers and imaging systems reaching back to the 1970s. For $n \times n$ Hadamard matrices, it may be proved that if $n > 2$, then n must be a multiple of 4. Interestingly the existence of Hadamard matrices of all orders $n = 4k$, known as the 'Hadamard conjecture', has remained a famous open problem despite over a century of study.

However, because they contain negative entries, experiments using Hadamard measurement matrices must acquire measurements in pairs ($2m$ total) which are subtracted in order to form the entries of $\mathbf{y}$. As a result, interest has also been sustained in using those matrices which minimize the MSE of Eq. (2.14) while only having entries in the set $\{0, +1\}$. These matrices are S-matrices, which exist in a one-to-one correspondence with Hadamard matrices. The nature of this correspondence is briefly explained below.

By definition, Hadamard matrices possess rows (and columns) which are mutually orthogonal. It follows that if H is an $n \times n$ Hadamard matrix, then another matrix H' is also a Hadamard matrix if it is obtained from H by exchanging rows or columns, or multiplying rows or columns by -1. Using these operations, it follows that all Hadamard matrices can be normalized, that is, brought into the form

$$H = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & & & \\ \vdots & & G & \\ 1 & & & \end{bmatrix}, \tag{2.15}$$

where G is a $\{+1, -1\}$ valued $(n-1) \times (n-1)$ sub-matrix. An S-matrix S may then be obtained by replacing -1 with $+1$, and $+1$ with 0 across the entries of G. Alternatively this can be written as $S = (G + J)/2$, where J is the $(n-1) \times (n-1)$ matrix filled with 1's.

It then follows (now considering $n = 4k - 1$ as the order of S) that S must obey the following properties:

$$SS^T = \frac{n+1}{4}(I+J), SJ = \frac{n+1}{2}J = JS^T, S^{-1} = \frac{2}{n+1}(2S^T - J). \quad (2.16)$$

2.2.2 Coding strategy with diagonally aggregated cyclic S-matrices

To apply SPI in LWIR imaging thermometry, we propose to implement cyclic S-matrices as the masking patterns, together with diagonal aggregation to facilitate a compact mask assembly.

Spatial reshaping of cyclic S-matrices:

During an SPI experiment related to $the\ m \times n$ measurement matrix S, there are typically m unique masking patterns derived from the reshaped rows of S. These are introduced to the imaging beam, yielding m associated photodiode measurements. When an independent modulator is employed, each pattern must be depicted as individual rectangular $(p \times q)$ areas of modulator pixels. Every pattern mandates a minimum transfer of control data to update the modulator's active area. Yet, by selecting S with a specific structure, several masking patterns' $p \times q$ fields can substantially overlap. This allows multiple patterns to be present concurrently on the modulator surface, reducing the control updates needed to achieve the measurement strategy dictated by the rows of S (Gopalsami *et al.*, 2012).

This overlap is attainable when S has a cyclic or left circulant structure. This means that given the first row of S, subsequent rows are formed by shifting the preceding row one unit to the left, using an overflow replacement method. When achieved, adjacent reshaped rows exhibit maximum overlap without being identical. There are known deterministic designs for S-matrices with this cyclic structure, though only three are recognized (Sloane & Harwit, 1979). The possible choices of n across these methods are limited to (i) $n = 2^k - 1$ for some positive integer k, (ii) n is prime, and (iii) $n = p \times q$ where p and q are twin primes. For the purposes of spatial imaging, rectangular 2D masks necessitate that n isn't a prime. As such, the third design (iii) offers the greatest adaptability and the sole assurance that n has factors of similar magnitude.

An example of the reshape of a cyclic S matrix is shown in Fig. 2.7(a) to (c), which utilize $p = 5$ and $q = 7$. The initial row of a cyclic S-matrix is shown in Fig. 2.7(a), while the complete cyclic S-matrces is displayed in Fig. 2.7(b) by shifting the previous row one element to the left with overflow replacement. By tiling copies of the reshaped initial row of a cyclic S-matrix, a compact mask, with a dimension of $(2p-1) \times (2q-1)$ is obtained [Fig. 2.7(c)], which has the property that the entire set of 2D masking patterns defined by the S-matrix can be found by forming all $p \times q$ contiguous sub-matrices from its elements.

Compact mask assembly with a diagonal aggregation:

For implementing SPIRIT in long-wavelength infrared (LWIR) SPI, a significant hardware constraint arises due to the reduced efficacy of spatial light modulators (SLMs) within the LWIR spectrum. This necessitates the use of physical masking patterns for spatial encoding. Compared to SLMs, physical masks are inherently less adaptable, especially when transitioning between different rows as shown in Fig. 2.7(b). Thus, to ensure ease of encoding with a physical mask and to keep the assembly of masking patterns simple, SPIRIT adopts a diagonal aggregation of the masking patterns. This guarantees a comprehensive coverage of the S-matrices while allowing for a unidirectional packing approach.

Here, we use cyclic S-matrices with dimensions of 5×7 pixels for a simplified explanation. As illustrated in Fig. 2.7(c) for a section of the entire diagonal aggregation, each colored box signifies a distinct masking pattern. This directly corresponds to a specific row in the full S-matrices, as highlighted by the box of the matching color in Fig. 2.7(b). Moreover, by replicating more copies of the initially reshaped S-matrix row and conducting a diagonal scan, which is highlighted by the red box in Fig. 2.7(d), the compact mask assembly approach becomes intuitive. Using this method, all mask patterns can be organized efficiently in an aggregated way. This design simplifies and condenses the physical mask assembly process.

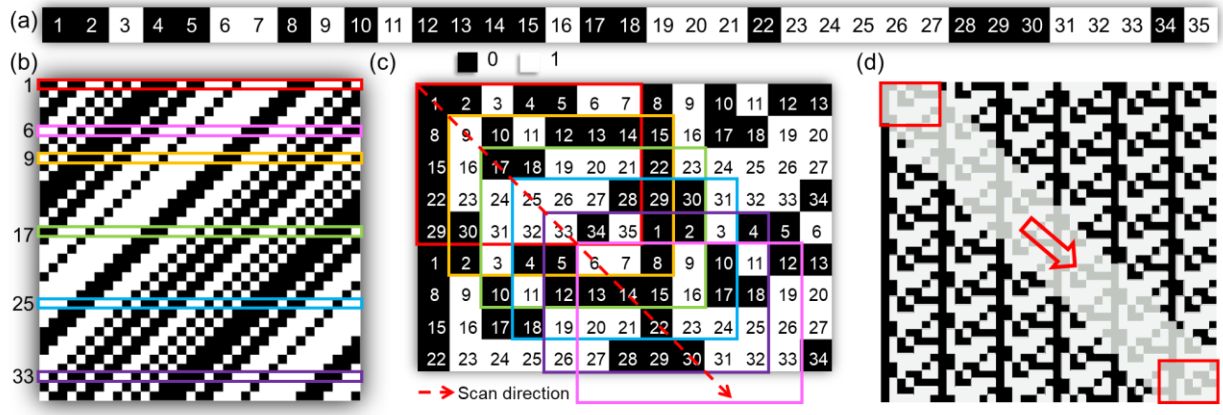

Figure 2.7 Illustration of the coding strategy of SPIRIT.

(a) The first row of S with $p = 5$ and $q = 7$. (b) The complete 35×35 cyclic S-matrices. (c) The compact mask uniquely covers the rows of the parent matrix. The colored boxes are representing masking patterns, matching with the same color boxes in (b). (d) A large tiling of copies of reshaped row cyclic S-matrix with diagonal aggregation. Red boxes indicate individual masking patterns and the arrow shows the scanning direction.

2.2.3 Reconstruction and temperature mapping

Image reconstruction with data interpolation:

The image reconstruction process in SPIRIT hinges on the understanding that multiplying a vector by an appropriate cyclic S-matrix can be seen as a two-dimensional discrete convolution operation. Thus, a specific relationship can be found between the matrix-vector product of an image x with S. Due to both horizontal and vertical shifts of ±1 by the diagonal aggregation, there's significant overlap between encoding patterns and the image. This overlap implies that bucket signals from SPI using cyclic S-matrices exhibit 2D smoothness when organized via row-major ordering.

This anticipated 2D smoothness allows for redundancy, which SPIRIT uses to reduce the size of the physical masking patterns. The design of these masks ensures that a linear scan results in diagonal bucket signals, offering a property that even an interrupted data collection provides measurements distributed uniformly, suitable for 2D interpolation. More specifics about matrix order n and scan size m are in Section 6.5.4.

For estimating unmeasured data, the discrete Laplace interpolation technique was utilized, apt for handling irregularly distributed data points (Gopalsami *et al.*, 2012). For a deeper dive, refer to Section 6.2.2. Once estimates of the unmeasured values are obtained, they are integrated with the measured data to form a comprehensive interpolated data vector $\tilde{y}$. Then, the final image $\tilde{x}$ is then generated by inverting the full-rank measurement matrix as

$$\tilde{x} = S^{-1}\tilde{\mathbf{y}}, \tag{2.17}$$

Temperature calculation:

According to Stefan-Boltzmann law, the temperature distribution will then be calculated by (Standardization., 2022)

$$T = \sqrt[4]{\frac{\tilde{x}}{C\varepsilon\sigma_{\mathrm{SB}}}}, \tag{2.18}$$

where C is a constant representing a conversion gain, ε=0.98±0.01 is the emissivity of human skin, and σ_{SB} is the Stefan-Boltzmann constant.

2.3 References

Bone DJ (1991) Fourier fringe analysis: the two-dimensional phase unwrapping problem. *Applied optics* 30(25):3627-3632.

Bouguet J-Y (2004) Camera calibration toolbox for matlab. *http://www. vision. caltech. edu/bouguetj/calib_doc/index. html*.

Chen K, Xi J & Yu Y (2013) Quality-guided spatial phase unwrapping algorithm for fast three-dimensional measurement. *Optics Communications* 294:139-147.

Chung K-L, Pei S-C, Pan Y-L, Hsu W-L, Huang Y-H, Yang W-N & Chen C-H (2011) A gradient-based adaptive error diffusion method with edge enhancement. *Expert Systems with Applications* 38(3):1591-1601.

Flynn TJ (1997) Two-dimensional phase unwrapping with minimum weighted discontinuity. *JOSA A* 14(10):2692-2701.

Ghiglia DC & Pritt MD (1998) Two-dimensional phase unwrapping: theory, algorithms, and software. *A Wiley Interscience Publication*.

Ghiglia DC & Romero LA (1994) Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods. *JOSA A* 11(1):107-117.

Gopalsami N, Liao S, Elmer TW, Koehl ER, Heifetz A, Raptis AC, Spinoulas L & Katsaggelos AK (2012) Passive millimeter-wave imaging with compressive sensing. *Optical Engineering* 51(9):091614-091614.

Huang PS & Zhang S (2006) Fast three-step phase-shifting algorithm. *Applied optics* 45(21):5086-5091.

Jiang C, Kilcullen P, Lai Y, Wang S, Ozaki T & Liang J (2022) Multi-scale band-limited illumination profilometry for robust three-dimensional surface imaging at video rate. *Optics Express* 30(11):19824-19838.

Jiang C, Kilcullen P, Liu X, Gribben J, Boate A, Ozaki T & Liang J (2020) Real-time high-speed three-dimensional surface imaging using band-limited illumination profilometry with a CoaXPress interface. *Optics Letters* 45(4):964-967.

Judge TR & Bryanston-Cross P (1994) A review of phase unwrapping techniques in fringe analysis. *Optics and Lasers in Engineering* 21(4):199-239.

Knox KT (1992) Error image in error diffusion. *Image Processing Algorithms and Techniques III.* SPIE, p 268-279.

Liang J, Kohn Jr RN, Becker MF & Heinzen DJ (2010) High-precision laser beam shaping using a binary-amplitude spatial light modulator. *Applied optics* 49(8):1323-1330.

Liang J, Wu S-Y, Kohn RN, Becker MF & Heinzen DJ (2012) Grayscale laser image formation using a programmable binary mask. *Optical Engineering* 51(10):108201.

Mao X, Chen W & Su X (2007) Improved Fourier-transform profilometry. *Applied optics* 46(5):664-668.

Sloane NJ & Harwit M (1979) Hadamard Transform Optics. (Academic Press New York).

Standardization. IOf (2022) Thermal insulation — Heat transfer by radiation — Vocabulary. *ISO* 9288:2022.

Su X & Chen W (2001) Fourier transform profilometry:: a review. *Optics and lasers in Engineering* 35(5):263-284.

Zhang H, Zhang Q, Li Y & Liu Y (2019) High speed 3D shape measurement with temporal Fourier transform profilometry. *Applied Sciences* 9(19):4123.

Zhang S & Huang PS (2006) Novel method for structured light system calibration. *Optical Engineering* 45(8):083601.

Zhang S & Yau S-T (2006) High-resolution, real-time 3D absolute coordinate measurement based on a phase-shifting method. *Optics Express* 14(7):2644-2649.

Zuo C, Feng S, Huang L, Tao T, Yin W & Chen Q (2018a) Phase shifting algorithms for fringe projection profilometry: A review. *Optics and Lasers in Engineering* 109:23-59.

Zuo C, Tao T, Feng S, Huang L, Asundi A & Chen Q (2018b) Micro Fourier Transform Profilometry (μFTP): 3D shape measurement at 10,000 frames per second. *Optics and Lasers in Engineering* 102:70-91.

3 BLIP OF REAL-TIME KHZ-LEVEL 3D IMAGING

Real-time high-speed three-dimensional surface imaging using band-limited illumination profilometry with a CoaXPress interface

Imagerie de surface tridimensionnelle à haute vitesse en temps réel utilisant la profilométrie d'éclairage à bande limitée par interface CoaXPress

Authors :

Cheng Jiang,[1,3] Patrick Kilcullen,[1,3] Xiangleil Liu,[1] Jeremy Gribben,[2] Alan Boate,[2] Tsuneyuki Ozaki,[1] and Jinyang Liang,[1,*]

[1]Centre Énergie Matériaux Télécommunications, Institut National de la Recherche Scientifique, 1650 boulevard Lionel-Boulet, Varennes, Québec J3X1P7, CANADA

[2]Ajile Light Industries, 5480 Canotek Road, Unit 6, Ottawa, Ontario K1J9H5, CANADA

[3]These authors contributed equally to this work.

Publication:

Optics Letters Vol. 45, Issue 4, pp 964-967

Submitted 25 September 2019, accepted 10 January 2020

https://doi.org/10.1364/OL.378939

Contribution of authors:

Cheng Jiang built the system, conducted the experiments, analyzed the data, and wrote the manuscript. Patrick Kilcullen developed the software and provided part of the manuscript. Jeremy Gribben and Alan Boate helped in the hardware operation. Tsuneyuki Ozaki provided editorial input in the manuscript preparation. Jinyang Liang initiated the project, proposed the concept, contributed to experimental design, and supervised the project.

Abstract

High-speed three-dimensional (3D) surface imaging by structured-light profilometry is currently driven by industrial needs, medical applications, and entertainment. However, the limited speeds in fringe pattern projection, image acquisition, and data transmission have strained the existing methods from reaching kilohertz (kHz)-level acquisition, processing, and display of 3D information during the occurrence of dynamic events (i.e., in real time). To overcome these limitations, we have developed CoaXPress-interfaced band-limited illumination profilometry (CI-BLIP), which enables real-time 3D surface information reconstruction at 1 kHz. We have demonstrated the system's performance by imaging various static and fast-moving 3D objects in real time. We have also applied this system in fluid mechanics by imaging dynamics of a flag, which allowed observation of the wave propagation, gravity-induced phase mismatch, and asymmetric flapping motion. We expect CI-BLIP to find diverse scientific and industrial applications.

3.1 Introduction

Three-dimensional (3D) surface imaging is used in widespread fields, including machine vision (Song, 2013), remote sensing (Beraldin *et al.*, 2000), biomedical engineering (Beraldin *et al.*, 2000), and entertainment (Seer *et al.*, 2014). The incessant demands for higher spatial resolution, a larger field of view (FOV), and a higher imaging speed have driven researchers to invent a myriad of 3D surface imaging methods (Gosta & Grgic, 2010; Hansard *et al.*, 2012; Zhang, 2010). Among them, structured-light profilometry (SLP) has become the mainstream technique (Zhang, 2018a). In a typical setup, a projector delivers structured patterns onto an object. The structured patterns, deformed by the object's 3D surface, are captured by a camera. The 3D information is then extracted by analyzing the patterns' deformation with respect to reference phase profiles. Among existing choices of structured patterns in SLP (Geng, 2011), the sinusoidal fringes are most widely used because of their provision of 3D data with both high spatial resolution and high depth accuracy (Zuo *et al.*, 2018a) as well as their suitability for high-speed 3D measurements (Gorthi & Rastogi, 2010). To generate these fringes, digital micromirror devices (DMDs) are commonly used. Each micromirror on the DMD can be independently tilted to either +12° or –12° from its surface normal to generate binary patterns at up to tens of kilohertz (kHz). By controlling the spatially dependent reflectance via temporal dithering of each micromirror, the DMD can produce grayscale sinusoidal patterns (Hornbeck, 1997). The DMD-based SLP, flexible in system development and accurate in 3D measurements (Su & Zhang, 2010), has been implemented in numerous scientific and industrial applications (Deetjen *et al.*, 2017; Li & Zhang, 2018).

Recently, the development of DMD-based SLP has increasingly emphasized high-speed 3D surface imaging in real time (defined as image acquisition, processing, and display during the occurrence of dynamic events (Van der Jeught & Dirckx, 2016)). Nonetheless, most existing systems fall short in accomplishing these objectives from the aspects of fringe pattern generation as well as image acquisition and processing. Micromirror-dithering clamps the speed of generating 8-bit grayscale images to around 100 Hz (Zhang, 2018b), which excludes 3D profilometry of many moving objects (Zuo *et al.*, 2018b). This limitation can be mitigated by using binary defocusing (Li *et al.*, 2014), which generates a pseudo-sinusoidal pattern at an unconjugate plane to the DMD by slightly defocusing the projector. However, this method compromises the depth-sensing range (Lei & Zhang, 2009) and is less flexible when used with binary patterns with different periods. Meanwhile, the DMD's uneven surface (Parthiban *et al.*, 2016) could induce image distortion to the defocused sinusoidal patterns at the unconjugate plane, which may decrease measurement accuracy, especially under coherent illumination.

Besides illumination schemes, image acquisition devices and image processing modules in existing SLP systems also have limitations. Most high-speed cameras deployed in SLP, despite having ultra-high imaging speeds, are not equipped with a high-speed interface to transfer data on time. This bottleneck also hampers the on-line processing software developed based on graphics processing units (GPUs) (Han & Abdelrahman, 2010). As a result, these systems cannot continuously stream data, further limiting their application scope to highly synchronous events.

3.2 Results

3.2.1 Operating principle of CI-BLIP

To overcome these limitations, we developed CoaXPress-interfaced band-limited illumination profilometry [CI-BLIP, Fig. 3.1(a)]. A continuous-wave laser (671-nm wavelength and 200-mW power) is used as the light source. After expansion and collimation (L1 with a focal length of -20 mm and L2 with a focal length of 200 mm), the laser beam is directed to a 0.45” DMD at an incident angle of ~24° with respect to its surface normal. Binary patterns, loaded onto the DMD, generate sinusoidal fringes at the intermediate image plane by a band-limited 4f imaging system that consists of two lenses (L3 with a focal length of 120 mm and L4 with a focal length of 175 mm) and one pinhole. Then, a camera lens projects these fringes onto a 3D object with an FOV of 100 mm × 100 mm. The deformed structured images, scattered from the object, are captured by a high-speed CMOS camera at 5 kHz with 512 × 512 pixels in each frame. The captured images are transferred to a host computer via a four-channel CoaXPress (CXP) cable connected to a frame grabber. A GPU-based image processing software developed in house is used for on-line 3D image reconstruction.

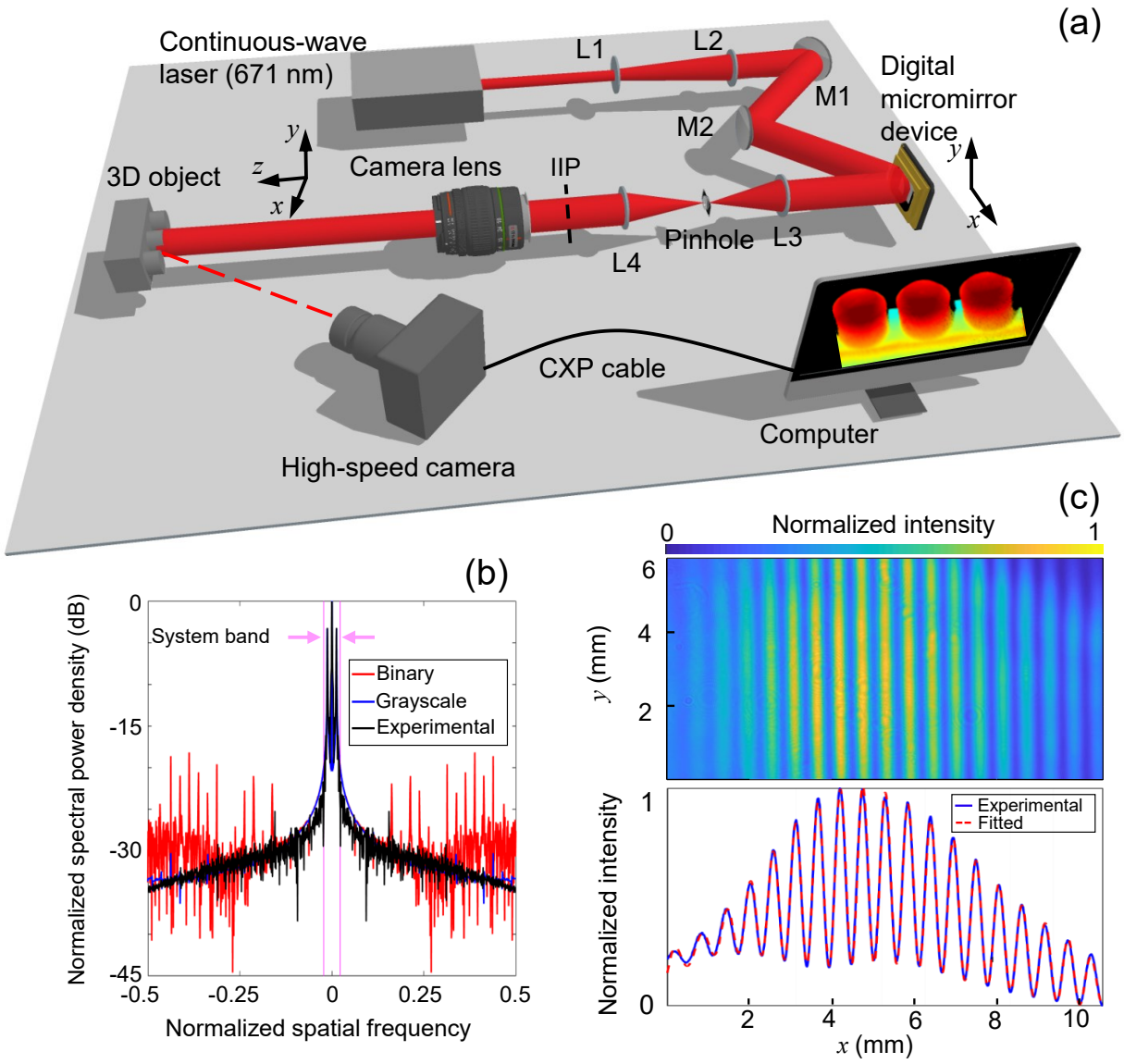

Figure 3.1 Operating principle of CI-BLIP.

(a) System schematic. Equipment details: continuous-wave laser (MRL-III-671, CNI Lasers), camera lens (AF-P DX NIKKOR, Nikon), digital micromirror device (AJD-4500, Ajile Light Industries), frame grabber (Cyton-CXP, Bitflow), graphics processing unit (GeForce GTX 1060, NVIDIA), high-speed camera (CP70-1HS-M-1900, Optronis), IIP: intermediate image plane, Lenses: L1 (32-990, Edmund Optics), L2 (45-169, Edmund Optics), L3 (63-559, Edmund Optics), L4 (45-895, Edmund Optics), Mirrors: M1 and M2 (BB2-E02, Thorlabs), Pinhole (P150D, Thorlabs). (b) Comparison of the central cross-sections of the spectral power density of the target sinusoidal pattern, the corresponding binary DMD pattern, and experimentally captured image. (c) Top panel: Grayscale sinusoidal pattern generated experimentally by band-limited illumination. Bottom panel: The averaged cross-section from the top panel with a fitted result.

The real-time, high-speed 3D surface imaging in CI-BLIP is enabled by the following two innovations. First, we implement band-limited illumination to generate and project high-quality sinusoidal fringe patterns at 5 kHz. The set of target images contains four grayscale sinusoidal fringe patterns (648-μm period on the DMD) with phase shifts from 0 to $3\pi/2$ with a step of $\pi/2$ as

well as one stripe pattern, which contains only one vertical stripe at the center of DMD. We first convert each individual grayscale fringe pattern to its corresponding binary DMD pattern using an adaptive error diffusion algorithm (Chung *et al.*, 2011; Liang *et al.*, 2010). The generated binary DMD patterns possess blue noise characteristics in the spatial frequency domain (Liang *et al.*, 2012b), which is manifested by precisely matched imaging content to the corresponding grayscale pattern within the system's bandwidth [Fig. 3.1(b)]. To filter high-spatial-frequency noise, we place a 150-μm pinhole at the Fourier plane in the 4f imaging system (Liang *et al.*, 2012a). The resultant beam profiles are high-quality grayscale sinusoidal fringes [Fig. 3.1(c)], with a root-mean-square error of 4.18% with respect to the target pattern. Compared with the binary defocusing techniques, the band-limited illumination scheme generates accurate sinusoidal patterns at the DMD's conjugate plane, which avoids the depth-dependent blurring effect and image distortion, thus assisting in improving the operating flexibility.

The second innovation in CI-BLIP lies in the synergy between the high-speed camera interface and the GPU-based processing, which allows real-time high-speed 3D image reconstruction. The CXP cable, with a bandwidth of 25 Gbps, enables continuous streaming of fringe images to the host computer during image acquisition at 5 kHz. These images are stored into RAM via direct memory access operations independent from the CPU. After determining the vertical stripe pattern, we process the calibration frame and four fringe patterns in GPU, beginning with the parallelized extraction of wrapped phase and quality map information. We then carry out phase unwrapping via our GPU accelerated implementation of a two-dimensional (2D) weighted phase unwrapping algorithm (Ghiglia & Pritt, 1998; Ghiglia & Romero, 1994), which uses the quality map information to guide the unwrapping procedure against errors induced by noise. The selected algorithm relies on the independent manipulation of points using 2D discrete cosine transformation (https://developer.nvidia.com/cufft.). The absolute phase is then computed in parallel by adding a constant determined from the vertical stripe pattern. Finally, we recover coordinate information from the absolute phase map from each pixel via matrix inversion. The resulted software enables real-time 3D position tracking at 1 kHz.

3.2.2 Depth resolution quantification of CI-BLIP

To quantify the depth resolution of CI-BLIP, we imaged two stacked planar surfaces offset by ~5° [Fig. 3.2(a)]. In this configuration, the depth difference between the two surfaces along the x axis increased monotonously, starting from a negative maximum at the right edge, to zero at the center, and reaching a positive maximum at the left edge. The reconstructed image of the 3D

object [Fig. 3.2(b)] allowed analyzing cross-sections of depth at different x positions. We calculated the depth difference (denoted by Δz) of the two surfaces, which closely agrees with the ground truth [Fig. 3.2(c)]. In addition, standard deviations of the measured depths were used as the system's noise level. The depth resolution was defined as when Δz equals to two times of the system's noise level. CI-BLIP's depth resolution was 0.15 mm.

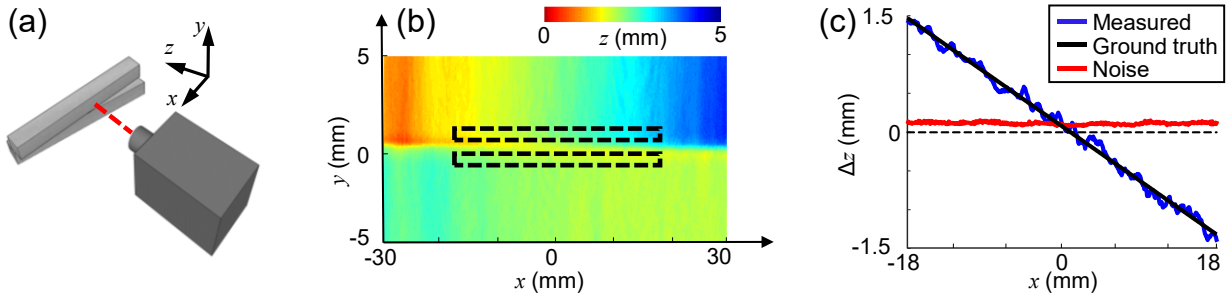

Figure 3.2 Quantification of CI-BLIP's depth resolution

(a) Experimental setup. (b) Reconstructed image of the planar surfaces. Black dashed boxes represent the selected regions for analysis. (c) Measured depth difference between the planar surfaces.

3.2.3 Static 3D imaging of CI-BLIP

To examine the feasibility of CI-BLIP, we imaged various static 3D objects. First, we measured three interlocking brick toys, whose studs had the heights of 23 mm, 15 mm, and 3 mm, respectively. Reconstructed results and selected centerlines are plotted in Fig. 3.3(a). The structure information of these brick toys is accurately recovered, manifested by the identification of shallow dents with depths of ~1 mm at the centers of the studs [see white arrows in Fig. 3.3(a)]. We also conducted proof-of-concept experiments on single and multiple animal toys [Figs. 3.3(b) and (c)]. In both cases, the depth information is shown in two perspective images and the detailed surface structures are illustrated by close-up views.

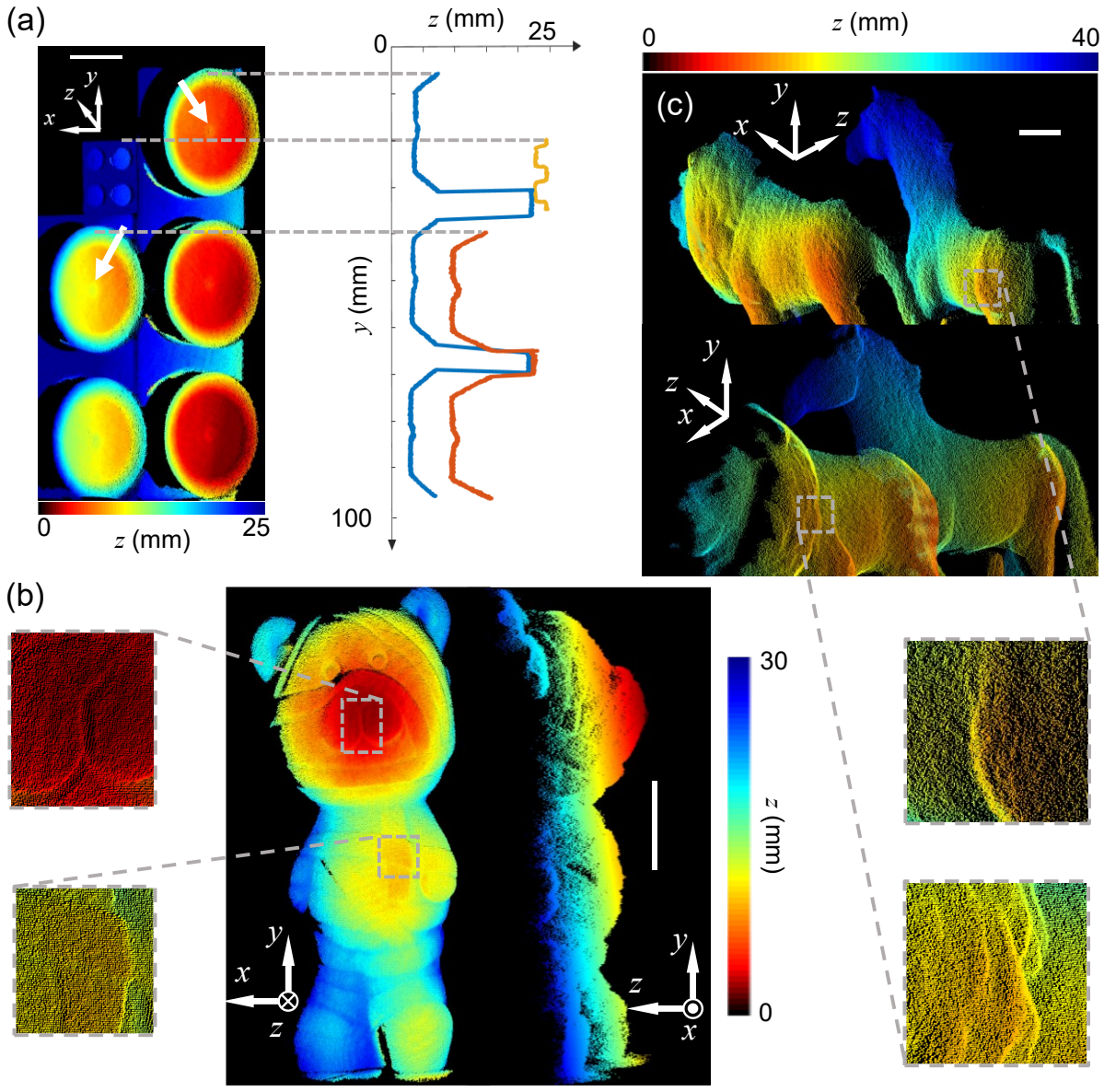

Figure 3.3 CI-BLIP of static 3D objects

(a) Reconstructed image of interlocking brick toys (left) and depth profiles of selected centerlines (right). (b) and (c) Reconstruction results of a bear toy (b) as well as a lion toy and a horse toy (c) with different perspective angles and close-up views. Scalar bar: 10 mm.

3.2.4 Dynamics 3D imaging of CI-BLIP

To verify real-time high-speed 3D surface profilometry, we used CI-BLIP to image an operating electric fan that rotated at 300 rounds per minute. The fan has seven blades, each having a tilt angle of ~30°. Figure 3.4(a) shows the front view of the reconstructed 3D images at four time points (full evolution is shown in *Visualization 3.1*). Along with a detailed 3D shapes of the base, center hub, side cover, blades, and the bars underneath the blades. CI-BLIP enabled tracking the

evolution of depth over the entire FOV. As an example, Fig. 3.4(c) gives line profiles along the radial direction [marked by the black solid line in the first panel of Fig. 3.4(a)] over time. The blade quickly swept through the radial profile in 10 ms, resulting in a maximum change in depth of 6 mm. To further analyze details in depth dynamics, in Fig. 3.4(d), we plot the depth profiles of three points on the line—one on the central hub (p_A), two along a blade (p_B and p_C). The depth profiles of p_A to p_C clearly illustrate the linear relationship between the displacements of the fan blades with different radii at a constant angular speed of ~30 rad/s. The depth of p_A does not shown apparent change, which is in accord with the flat central hub. The periodic displacements in depth of p_B and p_C show that the fan had a rotation period of approximately 210 ms. These measured results well agree with preset experimental conditions. Finally, *Visualization 3.2* shows real-time, high-speed imaging of a swinging pendulum and a bear statue using CI-BLIP. Selected results from the real-time reconstruction of this scene were displayed at 60 Hz. The demonstration also showed real-time 1-kHz 3D position tracking of a user-selectable single point.

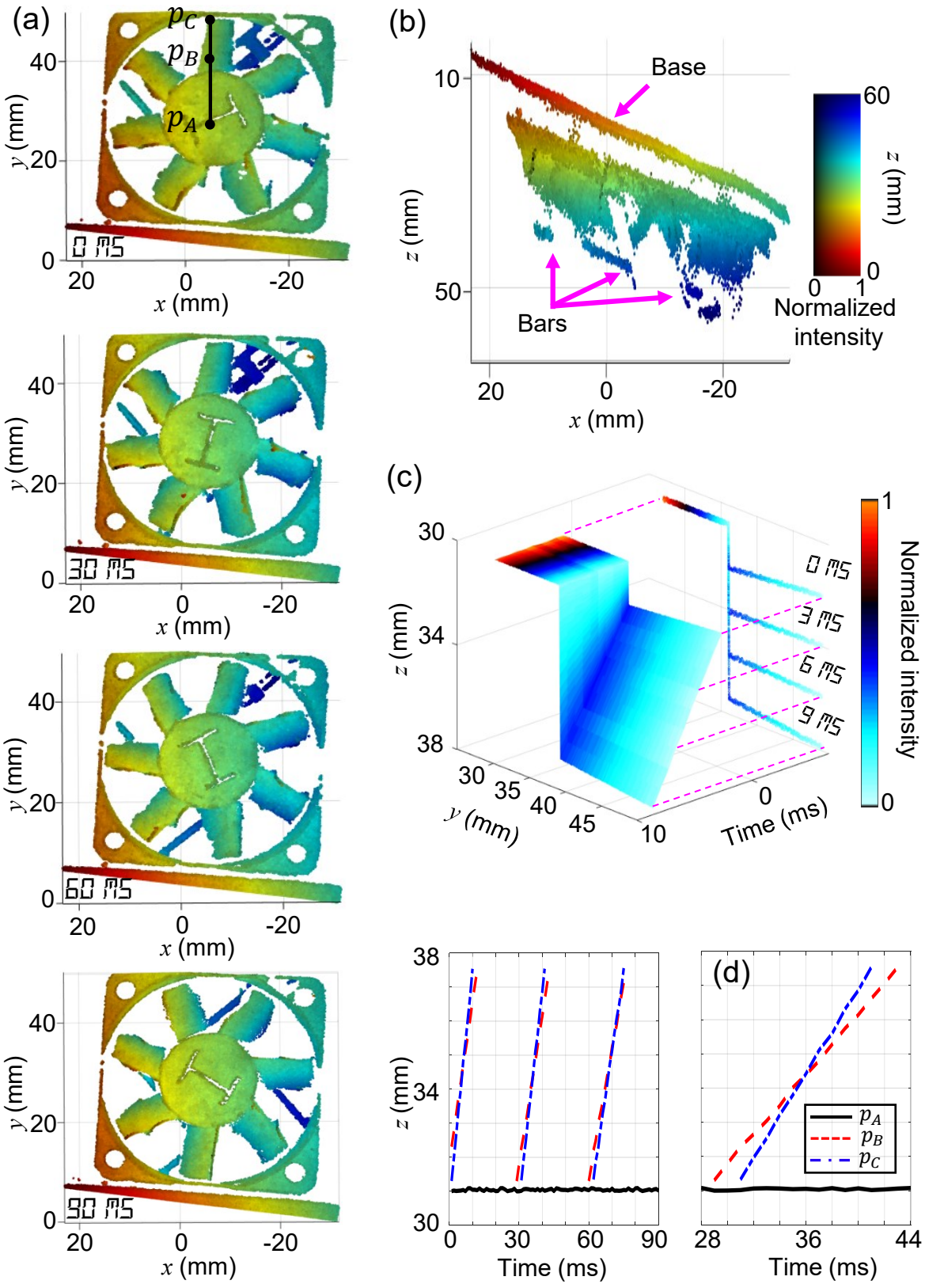

Figure 3.4 **CI-BLIP of a high-speed operating fan.**

(a) Front view of reconstructed 3D images at four different time points. (b) Side view of the reconstructed 3D image at 60 ms. (c) Depth line profiles along the radial direction over time. (d) Depth dynamics of the three selected points marked in the first panel in (a).

To highlight the broad utility of CI-BLIP, we used it to image flapping dynamics of a flag. In our experiment, a flag with a maple leaf pattern (80 mm × 50 mm in size) was mounted on a pole. An air blower generated strong wind interacting with the flag. Figure 3.5(a) depicts four representative 3D images of the instantaneous poses of the flapping flag from two perspective views (see the full sequence in *Visualization 3.3*), showing a wave traveling toward the edge at approximately 2 m/s. Displayed in Fig. 3.5(b), we analyzed time histories of the streamwise (x axis), spanwise (y axis), and transverse (z axis) displacements of three selected points [marked as p_A, p_B, and p_C in Fig. 3.5(a)]: p_A is at the mid-point in the y axis; p_B and p_C have the same x coordinate. The displacements of p_A have the smallest amplitudes in all three directions, showing a less intense flapping motion in the part closer to the pole. Moreover, the streamwise and transverse displacements of p_B and p_C show an apparent phase difference, which is probably attributed to the gravity-induced sagging effect. Furthermore, the phase relation between the spanwise and transverse displacements of p_C [Fig. 3.5(c)] show an elliptical shape in both experimental and fitted results, demonstrating that the flapping motion is dominated by single-frequency sinusoidal waves. Finally, the depth curves of the flag's centerline in all reconstructed images [Fig. 3.5(d)] show asymmetric flapping motion toward the -z direction, which indicates the uneven forces to the flag surface and a relatively high degree of turbulent flow.

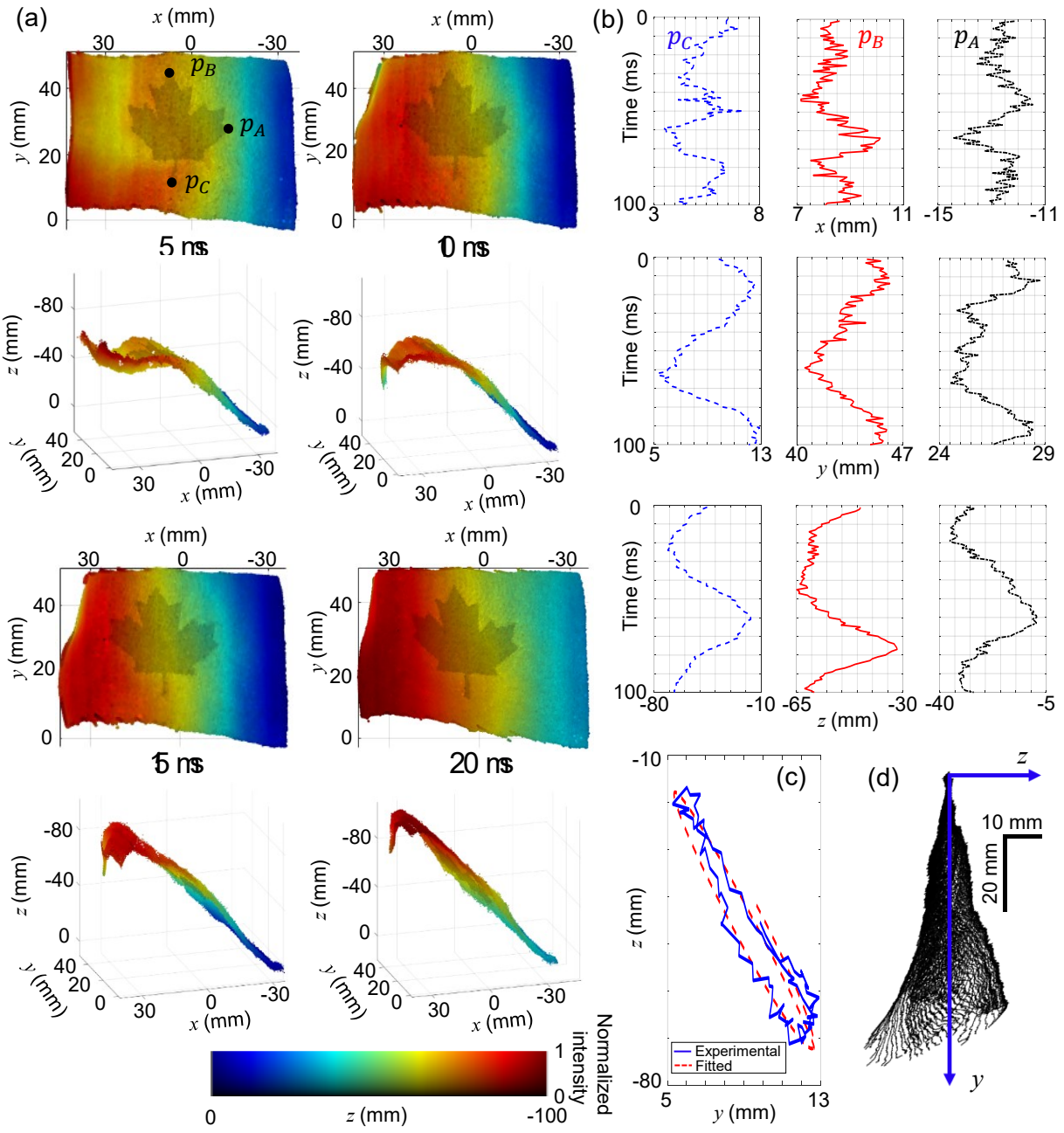

Figure 3.5 CI-BLIP of a flapping flag.

(a) Two perspective views of reconstructed 3D images of a flapping flag at different times. (b) Evolution of 3D positions of the three selected points marked in the first panel in (a). (c) Phase relation of the y-z positions of p_C and the fitted result. (d) Superimposed centerlines of the flag over time. Note: The depth of the pole is defined to zero in the z axis.

3.3 Discussion and Conclusions

We have developed CI-BLIP that integrates band-limited illumination, CoaXPress-interfaced high-speed image acquisition, and GPU-based image reconstruction. This high-speed SLP system has

enabled, for the first time, real-time 3D position tracking at 1 kHz. The major specifications could be further improved by using a high-power laser, a faster camera interface (Liu *et al.*, 2019; Waide, 2017), and advanced phase unwrapping algorithms (Hyun & Zhang, 2016).

CI-BLIP will likely open new avenues for diverse applications, including in-situ industrial inspection (Zhong *et al.*, 2015), dynamic biometric analysis (Chatterjee *et al.*, 2019), and acoustic vibrometry (Akutsu *et al.*, 2013). In particular, we expect it to immediately benefit studies of flag-fluid interactions. The mechanisms that govern these phenomena has generated scientific interests in diverse studies (Banerjee *et al.*, 2015), including animal ethology (Shelley & Zhang, 2011), paper engineering (Watanabe *et al.*, 2002), and hydroelectric power generation (Taylor *et al.*, 2001). Due to the flag's complex geometries and freely moving boundaries (Huang & Sung, 2010), study the gravity-induced impact in its interaction with fluid remains a challenge. Thus far, although many simulation works have been carried out to investigate the relation between gravity and the flag's stability with different flow air conditions (Alben & Shelley, 2008; Connell & Yue, 2007), experimental visualization of real-time flag movement has not kept up. A limited number of previous experiments were conducted using low flow speeds to avoid the complication induced by flag dragging its support (Shelley *et al.*, 2005; Zhang *et al.*, 2000). In addition, rigid flags were used in these experiments, which, however, restricted the analysis of fluid dynamics in 2D. In contrast, CI-BLIP has empowered the 3D analysis of the wave propagation, gravity-induced phase mismatch, and asymmetric flapping motion of a non-rigid flag at 1 kHz. These experiments may pave the way for potential future studies of high-speed nonlinear flapping dynamics (Argentina & Mahadevan, 2005). CI-BLIP's ability to continuously stream acquired data could also contribute to imaging unpredicted or non-repeatable flow dynamics.

3.4 Future perspectives

The real-time high-speed surface measurement of BLIP has drawn interest from 3D vision-assisted surgery (Wagner *et al.*, 2012). In response, we have initiated a collaboration to incorporate BLIP into a robot-assisted surgical platform. As depicted in Fig. 3.6(a), a near-infrared (NIR) continuous-wave laser at 980 nm serves as the light source. To ensure the adaptability of the NIR-BLIP in various surgical settings, a compact, mountable design has been adopted with dimensions of 12”×12”×5” [see Fig. 3.6(b)]. After installation and calibration, NIR-BLIP will undergo an initial proof-of-concept test, focusing on real-time 3D measurements of target tissues. Subsequently, integrating NIR-BLIP into the robotic surgery system will facilitate the transmission of 3D vision data to the robot's control software, offering real-time feedback.

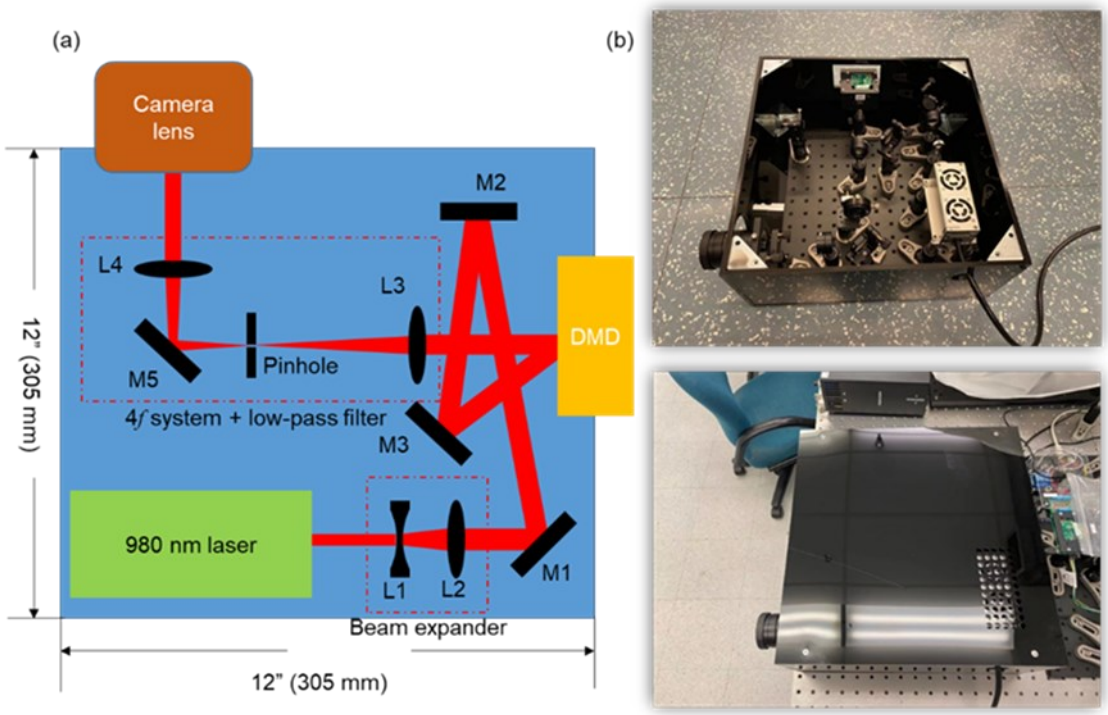

Figure 3.6 System schematic and construction of NIR-BLIP.

(a) System schematic. L1-L4, lenses; M1-M5, mirrors. (b) Images of the constructed system.

3.5 References

Akutsu M, Oikawa Y & Yamasaki Y (2013) Extract voice information using high-speed camera. *Proceedings of Meetings on Acoustics ICA2013.* ASA, p 055019.

Alben S & Shelley MJ (2008) Flapping states of a flag in an inviscid fluid: bistability and the transition to chaos. *Physical Review Letters* 100(7):074301.

Argentina M & Mahadevan L (2005) Fluid-flow-induced flutter of a flag. *Proceedings of the National Academy of Sciences* 102(6):1829-1834.

Banerjee S, Connell BS & Yue DK (2015) Three-dimensional effects on flag flapping dynamics. *Journal of Fluid Mechanics* 783:103-136.

Beraldin J-A, Blais F, Boulanger P, Cournoyer L, Domey J, El-Hakim S, Godin G, Rioux M & Taylor J (2000) Real world modelling through high resolution digital 3D imaging of objects and structures. *ISPRS J. Photogramm. Remote Sensing* 55(4):230-250.

Chatterjee A, Singh P, Bhatia V & Prakash S (2019) Ear biometrics recognition using laser biospeckled fringe projection profilometry. *Optics & Laser Technology* 112:368-378.

Chung K-L, Pei S-C, Pan Y-L, Hsu W-L, Huang Y-H, Yang W-N & Chen C-H (2011) A gradient-based adaptive error diffusion method with edge enhancement. *Expert Systems with Applications* 38(3):1591-1601.

Connell BS & Yue DK (2007) Flapping dynamics of a flag in a uniform stream. *Journal of Fluid Mechanics* 581:33-67.

Deetjen ME, Biewener AA & Lentink D (2017) High-speed surface reconstruction of a flying bird using structured light. *Journal of Experimental Biology* 220(11):1956-1961.

Geng J (2011) Structured-light 3D surface imaging: a tutorial. *Advances in optics and photonics* 3(2):128-160.

Ghiglia DC & Pritt MD (1998) Two-dimensional phase unwrapping: theory, algorithms, and software. *Wiely-Interscience, first ed.(April 1998)*.

Ghiglia DC & Romero LA (1994) Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods. *JOSA A* 11(1):107-117.

Gorthi SS & Rastogi P (2010) Fringe projection techniques: whither we are? *Optics and Lasers in Engineering* 48(IMAC-REVIEW-2009-001):133-140.

Gosta M & Grgic M (2010) Accomplishments and challenges of computer stereo vision. *Proceedings ELMAR-2010.* IEEE, p 57-64.

Han TD & Abdelrahman TS (2010) hiCUDA: High-level GPGPU programming. *IEEE Transactions on Parallel and Distributed systems* 22(1):78-90.

Hansard M, Lee S, Choi O & Horaud RP (2012) *Time-of-flight cameras: principles, methods and applications.* Springer Science & Business Media,

Hornbeck LJ (1997) Digital light processing for high-brightness high-resolution applications. *Projection Displays III.* International Society for Optics and Photonics, p 27-41.

https://developer.nvidia.com/cufft. (https://developer.nvidia.com/cufft.

Huang W-X & Sung HJ (2010) Three-dimensional simulation of a flapping flag in a uniform flow. *Journal of Fluid Mechanics* 653:301-336.

Hyun J-S & Zhang S (2016) Enhanced two-frequency phase-shifting method. *Applied Optics* 55(16):4395-4401.

Lei S & Zhang S (2009) Flexible 3-D shape measurement using projector defocusing. *Optics Letters* 34(20):3080-3082.

Li B, Wang Y, Dai J, Lohry W & Zhang S (2014) Some recent advances on superfast 3D shape measurement with digital binary defocusing techniques. *Optics and Lasers in Engineering* 54:236-246.

Li B & Zhang S (2018) Novel method for measuring a dense 3d strain map of robotic flapping wings. *Measurement Science and Technology* 29(4):045402.

Liang J, Kohn Jr RN, Becker MF & Heinzen DJ (2010) High-precision laser beam shaping using a binary-amplitude spatial light modulator. *Applied optics* 49(8):1323-1330.

Liang J, Kohn Jr RN, Becker MF & Heinzen DJ (2012a) Homogeneous one-dimensional optical lattice generation using a digital micromirror device-based high-precision beam shaper. *Journal of Micro/Nanolithography, MEMS, and MOEMS* 11(2):023002-023002.

Liang J, Wu S-Y, Kohn Jr RN, Becker MF & Heinzen DJ (2012b) Grayscale laser image formation using a programmable binary mask. *Optical Engineering* 51(10):108201-108201.

Liu X, Liu J, Jiang C, Vetrone F & Liang J (2019) Single-shot compressed optical-streaking ultra-high-speed photography. *Optics Letters* 44(6):1387-1390.

Parthiban V, Kohn Jr RN, Liang J & Becker MF (2016) Experimental demonstration of precise holograms using complex light modulation. *Emerging Digital Micromirror Device Based Systems and Applications VIII.* SPIE, p 88-95.

Seer S, Brändle N & Ratti C (2014) Kinects and human kinetics: A new approach for studying pedestrian behavior. *Transportation research part C: emerging technologies* 48:212-228.

Shelley M, Vandenberghe N & Zhang J (2005) Heavy flags undergo spontaneous oscillations in flowing water. *Physical Review Letters* 94(9):094302.

Shelley MJ & Zhang J (2011) Flapping and bending bodies interacting with fluid flows. *Annual Review of Fluid Mechanics* 43:449-465.

Song Z (2013) *Handbook of 3D machine vision : optical metrology and imaging.*

Su X & Zhang Q (2010) Dynamic 3-D shape measurement method: a review. *Optics and Lasers in Engineering* 48(2):191-204.

Taylor GW, Burns JR, Kammann S, Powers WB & Welsh TR (2001) The energy harvesting eel: a small subsurface ocean/river power generator. *IEEE Journal of Oceanic Engineering* 26(4):539-547.

Van der Jeught S & Dirckx JJ (2016) Real-time structured light profilometry: a review. *Optics and Lasers in Engineering* 87:18-31.

Wagner O, Hagen M, Kurmann A, Horgan S, Candinas D & Vorburger S (2012) Three-dimensional vision enhances task performance independently of the surgical method. *Surgical endoscopy* 26(10):2961-2968.

Waide D (2017) The Future of Machine Vision Standards. *Quality* :19VS.

Watanabe Y, Isogai K, Suzuki S & Sugihara M (2002) A theoretical study of paper flutter. *Journal of Fluids and Structures* 16(4):543-560.

Zhang J, Childress S, Libchaber A & Shelley M (2000) Flexible filaments in a flowing soap film as a model for one-dimensional flags in a two-dimensional wind. *Nature* 408(6814):835.

Zhang S (2010) Recent progresses on real-time 3D shape measurement using digital fringe projection techniques. *Optics and lasers in engineering* 48(2):149-158.

Zhang S (2018a) Absolute phase retrieval methods for digital fringe projection profilometry: A review. *Optics and Lasers in Engineering* 107:28-37.

Zhang S (2018b) High-speed 3D shape measurement with structured light methods: A review. *Optics and Lasers in Engineering* 106:119-131.

Zhong K, Li Z, Zhou X, Li Y, Shi Y & Wang C (2015) Enhanced phase measurement profilometry for industrial 3D inspection automation. *The International Journal of Advanced Manufacturing Technology* 76(9-12):1563-1574.

Zuo C, Feng S, Huang L, Tao T, Yin W & Chen Q (2018a) Phase shifting algorithms for fringe projection profilometry: A review. *Optics and lasers in engineering* 109:23-59.

Zuo C, Tao T, Feng S, Huang L, Asundi A & Chen Q (2018b) Micro Fourier Transform Profilometry (μFTP): 3D shape measurement at 10,000 frames per second. *Optics and Lasers in Engineering* 102:70-91.

4 BLIP OF ENHANCED KHZ-LEVEL 3D IMAGING

High-speed dual-view band-limited illumination profilometry using temporally interlaced acquisition

Profilometrie d'eclairage limite en bande a double vue a haute vitesse utilisant une acquisition entrelacee temporellement

Authors:

Cheng Jiang,[1] Patrick Kilcullen,[1] Yingming Lai, Tsuneyuki Ozaki, and Jinyang Liang[*]

Centre Énergie Matériaux Télécommunications, Institut National de la Recherche Scientifique, 1650 boulevard Lionel-Boulet, Varennes, Québec J3X1P7, CANADA

[1]These authors contributed equally to this work.

Publication:

Photonics Research Vol. 8, Issue 11, pp 1808-1817 (2020)

Submitted 4 June 2020, accepted 20 September 2020

https://doi.org/10.1364/PRJ.399492

Contribution of authors:

Cheng Jiang built the system, developed the software, conducted the experiments, analyzed the data, and wrote the manuscript. Patrick Kilcullen contributed to the algorithm development and provided part of the manuscript. Yingming Lai conducted some of the experiments. Tsuneyuki Ozaki provided editorial input in the manuscript preparation. Jinyang Liang initiated the project, proposed the concept, contributed to experimental design, and supervised the project.

Abstract

We report dual-view band-limited illumination profilometry (BLIP) with temporally interlaced acquisition (TIA) for high-speed three-dimensional (3D) imaging. Band-limited illumination based on a digital micromirror device enables sinusoidal fringe projection at up to 4.8 kHz. The fringe patterns are captured alternately by two high-speed cameras. A new algorithm, which robustly matches pixels in acquired images, recovers the object's 3D shape. The resultant TIA-BLIP system enables 3D imaging over 1000 frames per second on a field of view (FOV) of up to 180×130 mm^2 (corresponding to 1180 ×860 pixels) in captured images. We demonstrated TIA-BLIP's performance by imaging various static and fast-moving 3D objects. TIA-BLIP was applied to imaging glass vibration induced by sound and glass breakage by a hammer. Compared with existing methods in multi-view phase-shifting fringe projection profilometry, TIA-BLIP eliminates information redundancy in data acquisition, which improves the 3D imaging speed and the FOV. We envision TIA-BLIP to be broadly implemented in diverse scientific studies and industrial applications.

4.1 Introduction

Three-dimensional (3D) surface imaging has been extensively applied in numerous fields in industry, entertainment, and biomedicine (Kilcullen *et al.*, 2020; Su & Zhang, 2010). Among existing methods, structured-light profilometry has gained increasing popularity in measuring dynamic 3D objects because of its high measurement accuracy and high imaging speeds (Gorthi & Rastogi, 2010; Jinyang, 2020; Van der Jeught & Dirckx, 2016; Zhang, 2018). As the most widely used method in structured-light profilometry, phase-shifting fringe projection profilometry (PSFPP) uses a set of sinusoidal fringe patterns as the basis for coordinate encoding. In contrast to other approaches of structured light, such as binary pattern projection (Ishii *et al.*, 2007), the pixel-level information carried by the phase of the fringe patterns is insensitive to variations in reflectivity across an object's surface, which enables high accuracy in 3D measurements (Geng, 2011). The sinusoidal fringes employed in PSFPP are commonly generated by using digital micromirror devices (DMDs). Each micromirror on the DMD can be independently tilted to either +12° or –12° from its surface normal to generate binary patterns at up to tens of kilohertz. Despite being a binary amplitude spatial light modulator (Liang *et al.*, 2012a), the DMD can be used to generate grayscale fringe patterns at high speeds (Hornbeck, 1997; Jiang *et al.*, 2020; Lei & Zhang, 2009; Li *et al.*, 2014b; Liang *et al.*, 2012b). The conventional dithering method controls the average reflectance rate of each micromirror to form a grayscale image. However, the projection rate of fringe patterns is clamped at hundreds of hertz. To improve the projection speed, binary defocusing techniques (Li *et al.*, 2014b) have been developed to produce a quasi-sinusoidal pattern by slightly defocusing a single binary DMD pattern. Nonetheless, the image is generated at a plane unconjugate to the DMD, which compromises the depth-sensing range and is less convenient to operate with fringe patterns of different frequencies. Recently, these limitations are lifted by the development of band-limited illumination (Jiang *et al.*, 2020), which controls the system bandwidth by placing a pinhole low-pass filter at the Fourier plane of a 4*f* imaging system. Both the binary defocusing method and the band-limited illumination scheme allow generating one grayscale sinusoidal fringe pattern from a single binary DMD pattern. Thus, the fringe projection speed matches the DMD's refreshing rate.

High-speed image acquisition is also indispensable to DMD-based PSFPP. In the standard phase-shifting methods, extra calibration patterns must be used to avoid phase ambiguity (Wang & Zhang, 2011), which reduces the overall 3D imaging speed (Zuo *et al.*, 2012). A solution to this problem is to place multiple cameras at both sides of the projector to simultaneously capture the full sequence of fringe patterns (Bräuer-Burchardt *et al.*, 2011; Li *et*

al., 2010; Li *et al.*, 2013; Yin *et al.*, 2019). These multi-view approaches bring in enriched observation of 3D objects in data acquisition. Pixel matching between different views is achieved with various assistance, including epipolar-line rectification (Li *et al.*, 2010), measurement-volume-dependent geometry (Bräuer-Burchardt *et al.*, 2011), and wrapped phase monotonicity (Li *et al.*, 2013). Using these methods, the object's 3D surface information can be directly retrieved from the wrapped phase maps (Zhang, 2018). Consequently, the necessity of calibration patterns is eliminated in data acquisition and phase unwrapping. This advancement, along with the incessantly increasing imaging speeds of cameras (Hyun & Zhang, 2017; Jiang *et al.*, 2016; Karpinsky *et al.*, 2014; Li *et al.*, 2014a), has endowed multi-view PSFPP systems with image acquisition rates that keep up with the DMD's refreshing rates.

Despite these advantages, existing multi-view PSFPP systems are limited mainly in two aspects. First, each camera captures the full sequence of fringe patterns. This requirement imposes redundancy in data acquisition, which ultimately clamps the systems' imaging speeds. Given the finite readout rates of camera sensors, a sacrifice of the field of view (FOV) is inevitable for higher imaging speeds. Although advanced signal processing approaches, such as image interpolation (Unser *et al.*, 1991) and compressed sensing (Liu *et al.*, 2019), have been applied to mitigate this trade-off, they usually accompany with high computational complexity and reduced image quality (Lei *et al.*, 2017). Second, the cameras are mostly placed on different sides of the projector. This arrangement could induce a large intensity difference from the directional scattering light and the shadow effect from the occlusion by local surface features, both of which reduce the reconstruction accuracy and pose challenges in imaging non-Lambertian surfaces (Yin *et al.*, 2019).

To overcome these limitations, we have developed dual-view band-limited illumination profilometry (BLIP) with temporally interlaced acquisition (TIA). A new algorithm is developed for coordinate-based 3D point matching from different views. Implemented with two cameras, TIA allows each to capture half of the sequence of the phase-shifted patterns, reducing the data transfer load of each camera by 50%. This freed capacity is used either to transfer data from more pixels on each camera's sensor or to support using higher frame rates of both cameras. In addition, the two cameras are placed as close as possible on the same side of the projector, which largely mitigates the intensity difference and shadow effects. Leveraging these advantages, TIA-BLIP has enabled high-speed 3D imaging of glass vibration induced by sound and glass breakage by a hammer.

4.2 Method

4.2.1 System setup of TIA-BLIP

The schematic of the TIA-BLIP system is shown in Fig. 4.1(a). A 200-mW continuous-wave laser (wavelength λ=671 nm, MRL-III-671, CNI Lasers) is used as the light source. After expansion and collimation, the laser beam is directed to a 0.45" DMD (AJD-4500, Ajile Light Industries) at an incident angle of ~24° to its surface normal. Four phase-shifting binary vertical fringe patterns, generated by an error diffusion algorithm (Liang *et al.*, 2009) from their corresponding grayscale sinusoidal patterns, are loaded onto the DMD. The phase of each fringe pattern are set to be 0, π/2, π, and 3π/2. A band-limited 4f imaging system that consists of two lenses [Lens 1 and Lens 2 in Fig. 4.1(a)] and one pinhole converts these binary patterns to grayscale fringes at the intermediate image plane. The two lenses have focal lengths of f_1=120 mm and f_2=175 mm. The pinhole works as a low-pass filter. Its diameter, determined by the system bandwidth, is calculated as

$$D = \frac{\lambda f_1}{p_{\mathrm{f}}} \quad , \tag{4.1}$$

where p_{f}=324 μm denotes the fringe period composed of 30 DMD pixels. Thus, the required pinhole diameter is *D* =248.52 μm. In the experiment, a 300-μm-diameter pinhole is used to ensure all spatial frequency content of the sinusoidal fringe pattern passing through the system. Then, a camera lens (AF-P DX NIKKOR, Nikon, 18-55 mm focal length) projects these fringe patterns onto a 3D object. The deformed structure images are captured alternately by two high-speed CMOS cameras (CP70-1HS-M-1900, Optronis) with camera lenses (AZURE-3520MX5M, AZURE Photonics, 35 mm focal length) placed side by side. The distance between these two cameras is ~12 cm. The difference in their viewing angles to the 3D object is ~10°. Depending on their roles in image reconstruction, they are denoted as the main camera and the auxiliary camera. Synchronized by the DMD's trigger signal, each camera captures half of the sequence [Fig. 4.1(b)]. The acquired images from each camera are transferred to a computer via a CoaXPress cable connected to a frame grabber (Cyton-CXP, Bitflow).

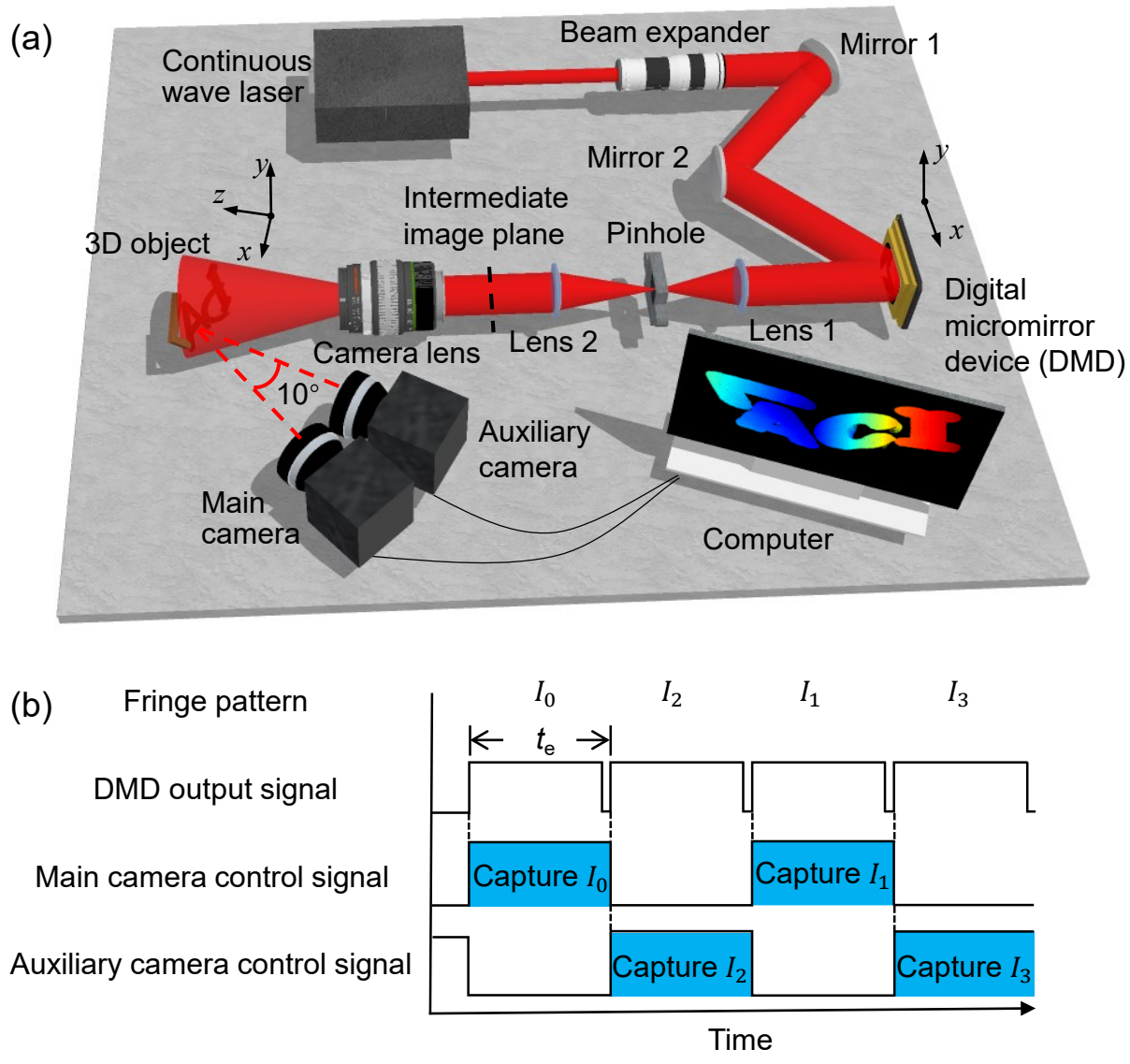

Figure 4.1 Operating principle of TIA-BLIP.

(a) System schematic. (b) Timing diagram and acquisition sequence. t_e: camera's exposure time.

4.2.2 System calibration of TIA-BLIP

To recover 3D information from the mutually incomplete images provided by the interlaced acquisition, TIA-BLIP relies on a coordinate-based understanding of the spatial relationship of the projector and both cameras in image formation. In particular, a "pinhole" model (Zhang & Huang, 2006),

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & \alpha & u_{\text{pp}} \\ 0 & f_v & v_{\text{pp}} \\ 0 & 0 & 1 \end{bmatrix} [R \quad T] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \tag{4.2}$$

describes the projection of 3D coordinates (x, y, z) onto camera coordinates (u, v). The extrinsic parameters, *R* and *T*, describe the rotation and translation of coordinates, respectively. The intrinsic parameters, which link to the camera's properties in image formation, consist of f_u and f_v (representing effective focal lengths along each axis of the camera's sensor), u_{pp} and v_{pp} (describing the coordinates of the camera's principal point), and α (accounting for pixel skewness). The column vectors $[u, v, 1]^T$ and $[x, y, z, 1]^T$ represent (u, v) and (x, y, z) in homogeneous coordinates, which through the scalar factor *s* allow for the numerical extraction of (u, v) from Eq. (4.2).

Based on the pinhole model, both cameras and the projector can be calibrated to determine the values of these parameters. Using a checkerboard as the calibration object, we adopted the established calibration procedure and software toolbox (Bouguet). Since the direct image acquisition is not possible for a projector, the phase-based mapping method (Zhang & Huang, 2006) was used to synthesize projector-centered images of the calibration object. These images were subsequently sent to the toolbox with calibration proceeding in the same manner as for the cameras.

4.2.3 Coordinate-based 3D point determination of TIA-BLIP

In the context of the pinhole model of Eq. (4.2), a coordinate-based method is used to recover 3D information from a calibrated imaging system. To a point on a 3D object with the coordinates (x, y, z), there correspond two independent coordinates: (u, v) for the camera and (u'', v'') for the projector. In a calibrated PSFPP system, any three of these coordinates (i.e.,(u, v, u'', v'')) can be determined. Then a linear system of the form $E = M[x, y, z]^T$ is derived. The elements of *E* and *M* are found by using each device's calibration parameters as well as by using the scalar factors and the three determined coordinates (Zhang *et al.*, 2006). In this way, 3D information of an object point can be extracted via matrix inversion.

This analysis can be adapted to dual-view TIA-BLIP. First, images from a selected calibrated camera are used to provide the coordinates (u, v) of a point on a 3D object. Along with the system's calibration parameters, an epipolar line is determined on the other camera. The horizontal coordinate in the images of this camera is recovered using search-based algorithms along this epipolar line—a procedure commonly referred to as stereo vision. Second, by substituting a calibrated projector in place of the secondary camera, structured light methods use the intensity values of the pixel (u, v) across a sequence of images to recover information about

a coordinate of the projector. By incorporating both aspects, 3D information can be extracted pixel by pixel based on interlaced image acquisition.

Data acquisition:

In data acquisition, four vertical fringe patterns, whose phases are equally shifted by $\pi/2$, illuminate a 3D object. The intensity value for the pixel (u, v) in the *k*th acquired image, $I_k\,(u, v)$, is expressed as

$$I_k(u,v) = I_{\mathrm{b}}(u,v) + I_{\mathrm{va}}(u,v)\cos\left[\varphi(u,v) - \frac{\pi k}{2}\right], \tag{4.3}$$

where $k \in [0,3]$. $I_{\mathrm{b}}(u,v)$, $I_{\mathrm{va}}(u,v)$, and $\varphi(u,v)$ represent background intensity, the variation of intensity, and depth-dependent phase, respectively.

Equation (4.3) allows analyzing two types of intensity matching conditions for the order of pattern projection shown in Fig. 4.1(b). The coordinates of a selected pixel in the images of the main camera are denoted by $(u_{\mathrm{m}}, v_{\mathrm{m}})$. For a pixel $(u'_{\mathrm{a}}, v'_{\mathrm{a}})$ in the images of the auxiliary camera that perfectly corresponds with $(u_{\mathrm{m}}, v_{\mathrm{m}})$, Eq. (4.3) allows us to write

$$I_0(u_{\mathrm{m}}, v_{\mathrm{m}}) + I_2(u'_{\mathrm{a}}, v'_{\mathrm{a}}) = I_1(u_{\mathrm{m}}, v_{\mathrm{m}}) + I_3(u'_{\mathrm{a}}, v'_{\mathrm{a}}). \tag{4.4}$$

Alternatively, rearrangement of Eq. (4.4) leads to the equivalent expression

$$I_0(u_{\mathrm{m}}, v_{\mathrm{m}}) - I_1(u_{\mathrm{m}}, v_{\mathrm{m}}) = I_3(u'_{\mathrm{a}}, v'_{\mathrm{a}}) - I_2(u'_{\mathrm{a}}, v'_{\mathrm{a}}). \tag{4.5}$$

In contrast to Eq. (4.4), each side of Eq. (4.5), containing images captured by the same camera, computes images with a residual fringe component. Retaining sinusoidal characteristics, this residual has the effect of improving the efficiency of line-constrained searches by regularizing encountered patterns of local maxima and minima and by including additional phase information. Moreover, by interpreting its right-hand side as a continuously varying function along the epipolar line, Eq. (4.5) together with bi-linear interpolation allows for the selection of discrete candidates with sub-pixel accuracy. Thus, Eq. (4.5) is selected as the intensity matching condition.

Image reconstruction:

We developed a four-step algorithm to recover the 3D image of the object pixel by pixel. In brief, for a selected pixel $(u_{\mathrm{m}}, v_{\mathrm{m}})$ of the main camera, the algorithm locates a matching point $(u'_{\mathrm{a}}, v'_{\mathrm{a}})$ in the images of the auxiliary camera. From knowledge of the camera calibration, this point then enables determining estimated 3D coordinates as well as recovering a wrapped phase. Using knowledge of the projector calibration, this phase value is used to calculate a horizontal

coordinate on the projector's plane. A final 3D point is then recovered using the coordinate-based method. A flowchart of this algorithm is provided in Fig. 4.2(a).

In the first step, $(I_0 + I_1)/2$ and $(I_2 + I_3)/2$ are calculated. Then, a threshold intensity, calculated from a selected background region, is used to eliminate pixels with low intensities. The thresholding results in a binary quality map [see Step I in Fig. 4.2(a)]. Subsequently, only pixels that fall within the quality map of the main camera are considered for 3D information recovery.

In the second step, the selected pixel $(u_\mathrm{m}, v_\mathrm{m})$ determines an epipolar line containing the matching point within the auxiliary camera's images. Then, the algorithm extracts the candidates $(u'_{\mathrm{a}i}, v'_{\mathrm{a}i})$ that satisfy the intensity matching condition [i.e., Eq. (4.5). Illustrative data are shown in Fig. 4.2(b)] in addition to three constraints [see Step II in Fig. 4.2(a)]. The subscript "i" denotes the i^{th} candidate. As displayed in the illustrative data in Fig. 4.2(c), the first constraint requires candidates to fall within the quality map of the auxiliary camera. The second constraint requires that candidates occur within a segment of the epipolar line determined by a fixed transformation that approximates the location of the matching point. This approximation is provided by a two-dimensional projective transformation (or homography) that determines the estimated corresponding point $(u'_\mathrm{e}, v'_\mathrm{e})$ by (Hartley & Zisserman, 2003)

$$s'[u'_\mathrm{e}, v'_\mathrm{e}, 1]^T = H[u_\mathrm{m}, v_\mathrm{m}, 1]^T, \tag{4.6}$$

where s' is a scalar factor representing extraction of the pair $(u'_\mathrm{e}, v'_\mathrm{e})$ from its homogeneous coordinates. H is determined by applying Eq. (4.6) to four points chosen as the corners of a flat rectangular plane when imaged by both cameras at the approximate center of the measurement volume. Once $(u'_\mathrm{e}, v'_\mathrm{e})$ are determined, the search along the epipolar line is confined to the segment occurring over the horizontal interval $[u'_\mathrm{e} - r_0, u'_\mathrm{e} + r_0]$, where r_0 is an experiment-dependent constant. In general, r_0 should be as small as possible while still covering the targeted depth range. For our experiments, the value of r_0 was set to 40 pixels.

The third constraint requires that the selected point and candidates must have the same sign of their wrapped phases [Fig. 4.2(d)]. Estimates of the wrapped phases are obtained using the technique of Fourier transform profilometry (Zhang, 2018). In particular, by band-pass filtering the left-hand side of Eq. (4.5), i.e., $I_0 - I_1$, the intensity of pixel $(u_\mathrm{m}, v_\mathrm{m})$ in the filtered image is

$$I_\mathrm{f}(u_\mathrm{m}, v_\mathrm{m}) = \frac{\sqrt{2}}{2} I_\mathrm{va}(u_\mathrm{m}, v_\mathrm{m}) \exp\left[j\left(\varphi(u_\mathrm{m}, v_\mathrm{m}) + \frac{\pi}{4}\right)\right]. \tag{4.7}$$

The wrapped phase estimation ω_m of $(u_\mathrm{m}, v_\mathrm{m})$is recovered by

$$\omega_{\mathrm{m}} = \tan^{-1}\left\{\frac{\Im[I_{\mathrm{f}}(u_{\mathrm{m}}, v_{\mathrm{m}})]}{\Re[I_{\mathrm{f}}(u_{\mathrm{m}}, v_{\mathrm{m}})]}\right\}, \tag{4.8}$$

where $\Im[\cdot]$ and $\Re[\cdot]$ respectively denote the imaginary part and the real part of a complex variable. The same treatment is applied to the right-hand side of Eq. (4.5), i.e., $I_3 - I_2$. For the candidate $(u'_{\mathrm{a}i}, v'_{\mathrm{a}i})$, the estimate of its wrapped phase is calculated as

$$\omega'_{\mathrm{a}i} = \tan^{-1}\left\{\frac{\Im[I'_{\mathrm{f}}(u'_{\mathrm{a}i}, v'_{\mathrm{a}i})]}{\Re[I'_{\mathrm{f}}(u'_{\mathrm{a}i}, v'_{\mathrm{a}i})]}\right\}. \tag{4.9}$$

The constraint requires ω_{m} and $\omega'_{\mathrm{a}i}$ to have the same sign in the interval $(-\pi, \pi]$. The output of the second step is a pool of candidates for further evaluation. If no candidate is found, the algorithm abandons the following process, and this step is re-initiated for the next pixel in the main camera.

In the third step, three criteria are used to calculate penalty scores for each candidate [see Step III in Fig. 4.2(a). The scheme is shown in Fig. 4.2(e)]. The first and primary criterion compares the phase values of the candidates using two methods. First, the phase inferred from the intensities of candidates and the pixel $(u_{\mathrm{m}}, v_{\mathrm{m}})$ is calculated by

$$\varphi'_{\mathrm{a}i} = \tan^{-1}\left[\frac{I_1(u_{\mathrm{m}}, v_{\mathrm{m}}) - I_3(u'_{\mathrm{a}i}, v'_{\mathrm{a}i})}{I_0(u_{\mathrm{m}}, v_{\mathrm{m}}) - I_2(u'_{\mathrm{a}i}, v'_{\mathrm{a}i})}\right]. \tag{4.10}$$

Meanwhile, for each candidate $(u'_{\mathrm{a}i}, v'_{\mathrm{a}i})$, the coordinate triple $(u_{\mathrm{m}}, v_{\mathrm{m}}, u'_{\mathrm{a}i})$ and knowledge of camera calibration allows determining an estimated 3D point $\boldsymbol{P_i}$ by using the stereo vision method. In addition, with the knowledge of the projector calibration, a point with coordinates $(u''_{\mathrm{p}i}, v''_{\mathrm{p}i})$ on the projector's plane is determined for each candidate. Then, an unwrapped phase value $\varphi''_{\mathrm{p}i}$ is calculated by

$$\varphi''_{\mathrm{p}i} = \frac{2\pi}{p}\left(u''_{\mathrm{p}i} - u''_{\mathrm{d}}\right), \tag{4.11}$$

where u''_{d} is a horizontal datum coordinate on the projector's plane associated with the zero phase, and p is the fringe period in units of projector pixels. Since these independently inferred phase values must agree if the candidate correctly matches $(u_{\mathrm{m}}, v_{\mathrm{m}})$, a penalty score A_i, as a normalized difference of these two phase values, is calculated by

$$A_i = \frac{\left|R\left(\varphi'_{\mathrm{a}i} - \varphi''_{\mathrm{p}i}\right)\right|}{\pi}, \tag{4.12}$$

where the rewrapping function $R(\cdot)$computes the subtracted difference between wrapped and unwrapped phase values.

To improve the robustness of the algorithm, two additional criteria are implemented using data available from the second step. B_i is a normalized distance score favoring candidates located closer to the estimated matching point $(u'_{\mathrm{e}}, v'_{\mathrm{e}})$, which is calculated by

$$B_i = \frac{|u'_{\mathrm{e}} - u'_{\mathrm{a}i}|}{r_0}. \tag{4.13}$$

Moreover, C_i is a normalized difference of wrapped phase values using the wrapped phases ω_{m} and $\omega'_{\mathrm{a}i}$, which is calculated by

$$C_i = \frac{|R(\omega_{\mathrm{m}} - \omega'_{\mathrm{a}i})|}{\pi}. \tag{4.14}$$

A total penalty score $S_i$for each candidate is then computed as a weighted linear combination of three individual scores,

$$S_i = \eta_1 A_i + \eta_2 B_i + \eta_3 C_i, \tag{4.15}$$

where the normalized weights [η_1, η_2, η_3] = [0.73, 0.09, 0.18] are empirically chosen to lead to the results that are most consistent with physical reality. Finally, the candidate with the minimum S_i is chosen as the matching point $(u'_{\mathrm{a}}, v'_{\mathrm{a}})$. Its phase values, calculated by using Eqs. (4.10) and (4.11), are denoted as φ'_{a} and φ''_{p}, respectively.

In the final step, the algorithm determines the final 3D coordinates [see Step IV in Fig. 4.2(a) and the scheme in Fig. 4.2(e)]. First, φ'_{a} is unwrapped as $\varphi'_{\mathrm{a}} + 2\pi q$, where q is an integer making $\varphi''_{\mathrm{p}} - (\varphi'_{\mathrm{a}} + 2\pi q) \in (-\pi, \pi]$. Then, the coordinate on the projector's plane, u''_{p}, is recovered with sub-pixel resolution as

$$u''_{\mathrm{p}} = u''_{\mathrm{d}} + P\left(\frac{\varphi'_{\mathrm{a}}}{2\pi} + q\right), \tag{4.16}$$

from which the final 3D coordinates (x, y, z) are computed using calibration information associated with the coordinate triple $(u_{\mathrm{m}}, v_{\mathrm{m}}, u''_{\mathrm{p}})$.

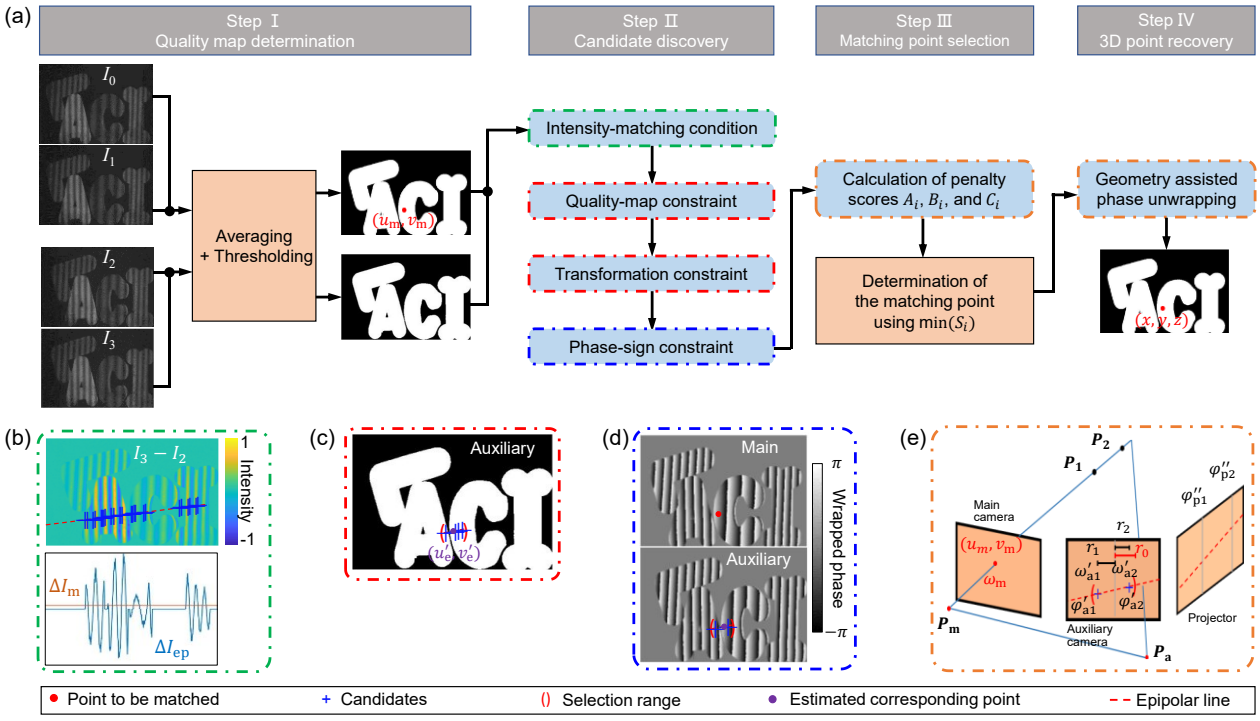

Figure 4.2 Coordinate-based 3D point determination algorithm.of TIA-BLIP.

(a) Flowchart of the algorithm. $(u_{\mathrm{m}}, v_{\mathrm{m}})$, Coordinates of the point to be matched for the main camera; (x, y, z), recovered 3D coordinates. (b) Illustrative data of the intensity matching condition. $\Delta I_{\mathrm{m}} = I_0(u_{\mathrm{m}}, v_{\mathrm{m}}) - I_1(u_{\mathrm{m}}, v_{\mathrm{m}})$; ΔI_{ep}, Intensity profile of $I_3 - I_2$ along the epipolar line. (c) Illustrative data of the quality map constraint and the transformation constraint. $(u'_{\mathrm{e}}, v'_{\mathrm{e}})$, Coordinates of the estimated corresponding point for the auxiliary camera. (d) Illustrative data of the phase sign constraint. (e) Scheme for penalty score calculation and phase unwrapping. r_i, Horizontal distance between the candidates and the estimated corresponding point; ω_{m}, Phase value of the selected point in the main camera calculated by the Fourier transform profilometry method; $\omega'_{\mathrm{a}i}$, Phase value of the candidate points in the auxiliary camera calculated by the Fourier transform profilometry method; $\varphi'_{\mathrm{a}i}$, Phase value calculated by the phase-shifting method; $\varphi''_{\mathrm{p}i}$, Phase value determined on the projector's plane; P_i, 3D points determined by candidates; P_{m}, Principal point of the main camera; P_{a}, Principal point of the auxiliary camera.

4.3 Results

4.3.1 Quantification of depth resolution

To quantify the depth resolution of TIA-BLIP with different exposure times, we imaged two stacked planar surfaces offset by ~9° (Fig. 4.3). Reconstructed results at four representative exposure times (denoted as t_{e}) are shown in Fig. 4.3(a). One area on each surface [marked as white solid boxes in Fig. 4.3(a)] was selected in the reconstructed image. The depth information on the *x*-axis was calculated by averaging the depth values along the y axis. The difference in depths between these two surfaces is denoted by z_{d}. In addition, the noise is defined as the averaged value of the standard deviations in depth from both surfaces. The depth resolution is defined as

when z_d equals twice the system's noise level. As shown in the four plots in Fig. 4.3(a), the reconstruction results deteriorate with shorter exposure times, manifested by increased noise levels and more points incapable of retrieving 3D information. As a result, the depth resolution degrades from 0.06 mm at t_e=950 μs to 0.45 mm at t_e=150 μs [Fig. 4.3(b)]. At t_e=100 μs, TIA-BLIP fails in 3D measurements. The region of unsuccessful reconstruction prevails across most of the planar surfaces. The noise dominates the calculated depth difference, which is attributed to the low signal-to-noise ratio in the captured images.

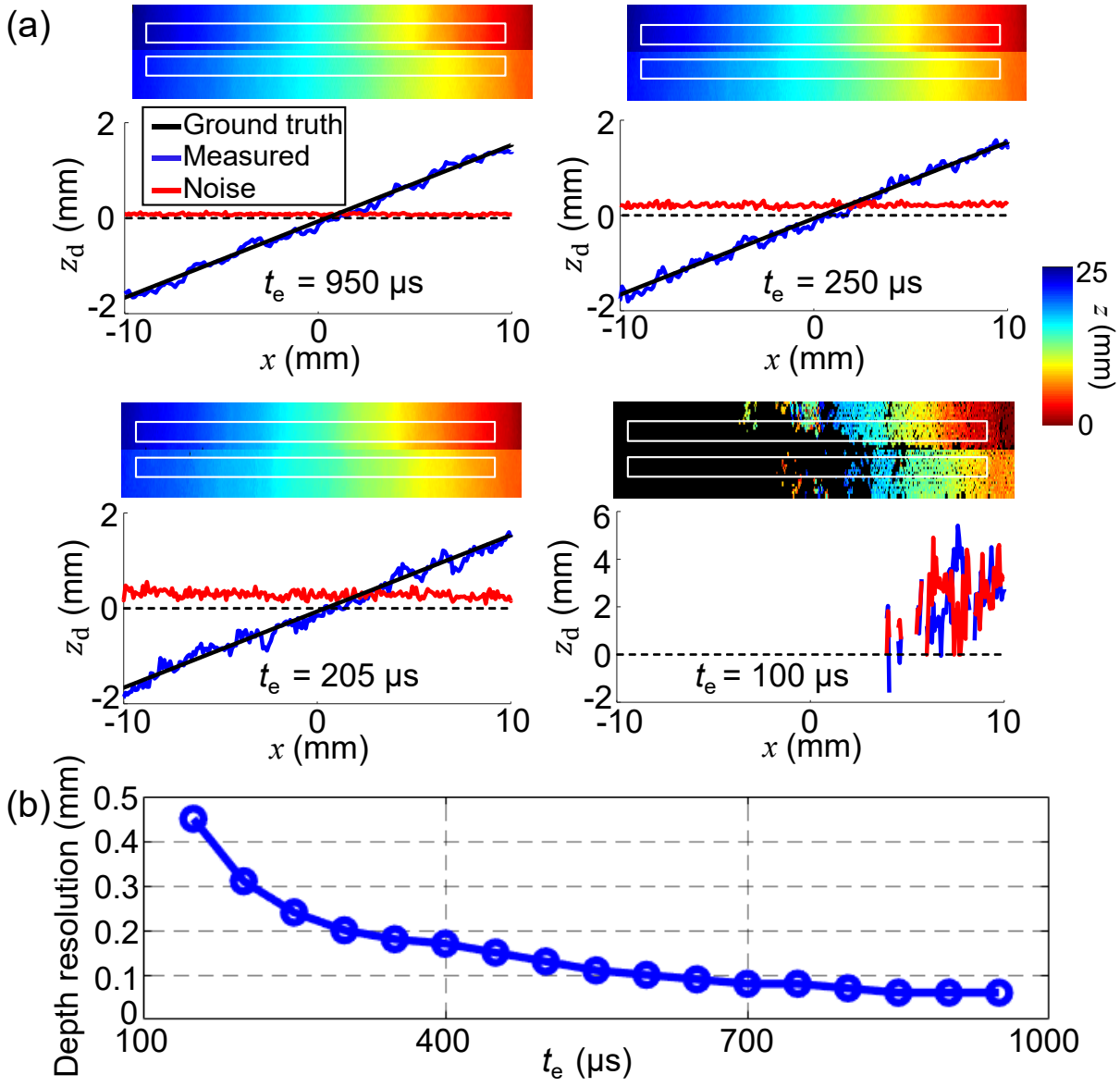

Figure 4.3 Quantification of TIA-BLIP's depth resolution.

(a) 3D images of the planar surfaces (top image) and measured depth difference (bottom plot) at four exposure times. The white boxes represent the selected regions for analysis. (b) Relation between measured depth differences and their corresponding exposure times.

4.3.2 Imaging of static 3D objects

To examine the feasibility of TIA-BLIP, we imaged various static 3D objects. First, two sets of 3D distributed letter toys that composed the words of "LACI" and "INRS" were imaged. Shown in Fig. 4.4(a), the two perspective views of the reconstructed results reveal the 3D position of each letter toy. The detailed surface structures are illustrated by the selected depth profiles [see the white dashed lines and the magenta dashed boxes in Fig. 4.4(a)]. We also conducted a proof-of-concept experiment on three cube toys with fine structures (with a depth of ~4 mm) on the surfaces. Depicted in Fig. 4.4(b), the detailed structural information of these cube toys is recovered by TIA-BLIP.

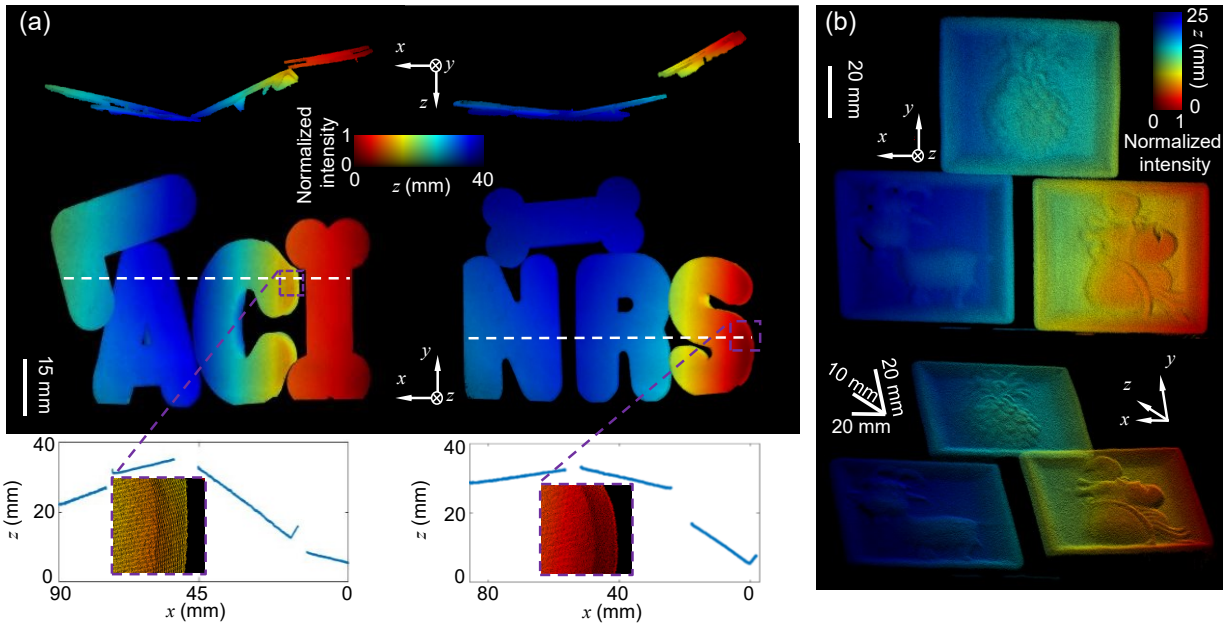

Figure 4.4 TIA-BLIP of static 3D objects.

(a) Reconstructed results of letter toys. Two perspective views are shown in the top row. Selected depth profiles (marked by the white dashed lines in the top images) and close-up views are shown in the bottom row. (b) Two perspective views of the reconstruction results of three toy cubes.

4.3.3 Imaging of dynamic 3D objects

To verify high-speed 3D surface profilometry, we used TIA-BLIP to image two dynamic scenes: a moving hand and three bouncing balls. The fringe patterns were projected at 4 kHz. The exposure times of both cameras were t_e=250 μs. Under these experimental conditions, TIA-BLIP had a 3D

imaging speed of 1 thousand frames per second (kfps), an FOV of 180×130 mm^2 (corresponding to 1180×860 pixels) in captured images, and a depth resolution of 0.24 mm. Figure 4.5(a) shows the reconstructed 3D images of the moving hand at five time-points from 0 ms to 60 ms with a time interval of 15 ms (see the full evolution in Visualization 4.1). TIA-BLIP's high-speed 3D imaging allowed tracking the movements of four fingertips. Shown in Fig. 4.5(b), all the four fingers have apparent movement in both the x axis and the z axis but stay relatively stationary in the y axis, which agrees with the experimental condition.

In the second experiment, three white balls, each of which was marked by a different letter on its surface, bounced in an inclined transparent container. Figure 4.5(c) shows five representative reconstructed images from 8 ms to 28 ms with a time interval of 5 ms. The changes of the letter "C" on B_1 and the letter "L" on B_2 [marked in the third panel of Fig. 4.5(c)] clearly show the rotation of the two balls (see the full evolution in Visualization 4.2). TIA-BLIP enabled tracking the 3D centroids of each ball over time. Shown in Fig. 4.5(d), B_1 collides with B_2 at 16 ms, resulting in a sudden change in the moving directions. This collision temporarily interrupted the free fall of B_1, represented by the two turning points in the curve of evolution along the y-axis [see the second panel of Fig. 4.5(d)]. The collision also changed the moving direction of B_2, making it touch the base at 27 ms and then bounce up. In this scene, B_3 maintained its movement in a single direction in both the x axis and the z axis. It fell onto the base and bounced back at $16\ \mathrm{ms}$, resulting in a turning point in its y-t curve. Because of the inclined bottom plane, the y-value of B_3 at 16 ms was smaller than that of B_2 at 27 ms.

Under the same experimental settings and pattern sequence choice, TIA-BLIP surpasses the existing PSFPP techniques in pixel counts and hence the imaging FOV. At the 1-kfps 3D imaging speed, the systems of standard single-camera PSFPP (Jiang *et al.*, 2020) and multi-view PSFPP (Li *et al.*, 2013) would restrict their imaging FOV to 512×512 pixels and 768×640 pixels, respectively (GmbH, 2018). In contrast, TIA, with a frame size of 1180×860, increases the FOV by 3.87 and 2.07 times, respectively.

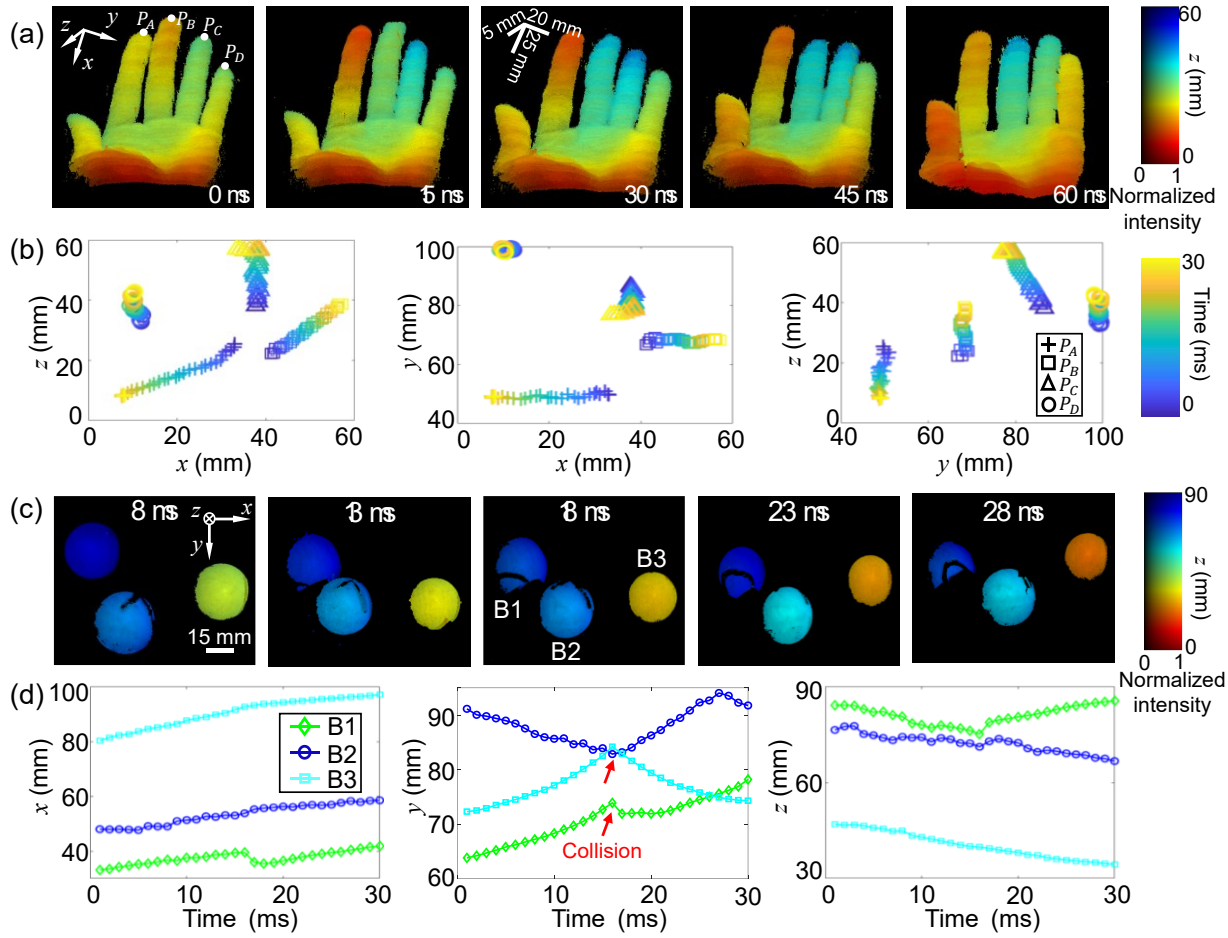

Figure 4.5 TIA-BLIP of dynamic objects.

(a) Reconstructed 3D images of a moving hand at five time points. (b) Movement traces of four fingertips [marked in the first panel in (a)]. (c) Front view of the reconstructed 3D image of the bouncing balls at five different time points. (d) Evolution of 3D positions of the three balls [marked in the third panel in (c)].

4.3.4 Imaging of sound-induced vibration on glass

To highlight the broad utility of TIA-BLIP, we imaged sound-induced vibration on glass. In this experiment [Fig. 4.6(a)], a glass cup was fixed on a table. The glass's surface was painted white. A function generator drove a speaker to produce single-frequency sound signals (from 450 Hz to 550 Hz with a step of 10 Hz) through a sound channel placed close to the cup's wall. To image the vibration dynamics, fringe patterns were projected at 4.8 kHz. The cameras had an exposure time of t_e=205 μs. This configuration enabled a 3D imaging speed of 1.2 kfps, an FOV of 146×130 mm^2 (correspond 960×860 pixels) in captured images, and a depth resolution of 0.31 mm. Figure 4.6(b) shows four representative 3D images of the instantaneous shapes of the glass cup driven by the 500-Hz sound signal (the full sequence is shown in Visualization 4.3), showing the dynamic

of structural deformation of the glass cup. The evolution of depth changes was analyzed using five selected points [marked by P_{A} to P_{E} in the first panel of Fig. 4.6(b)]. Shown in Fig. 4.6(c), the depth changes of five points are in accordance, which is attributed to the rigidness of the glass.

We further analyzed time histories of averaged depth displacements under different sound frequencies. Figure 4.6(d) shows the results at the driving frequencies of 490 Hz, 500 Hz, and 510 Hz. Each result was fitted by a sinusoidal function with a frequency of 490.0 Hz, 499.4 Hz, and 508.6 Hz, respectively. These results show that the rigid glass cup vibrated in compliance with the driving frequency. Moreover, the amplitudes of fitted results, Δz_{fit}, were used to determine the relationship between the depth displacement and the sound frequency [Fig. 4.6(e)]. We fitted this result by the Lorentz function, which determined the resonant frequency of this glass cup to 499.0 Hz.

It is worth noting that this phenomenon would be difficult to be captured by using previous methods given the same experimental settings and pattern sequence choice. With an image size of 960×860 pixels, the maximum frame rate for the used cameras is 2.4 kfps, which transfers to a 3D imaging speed of 480 fps for single-camera-based PSFPP and 600 fps for multi-view PSFPP. Neither provides sufficient imaging speed to visualize the glass vibration at the resonance frequency. In contrast, TIA improves the 3D imaging speed to 1.2 kfps, well capable of sampling the glass vibration dynamics in the tested frequency range of 450–550 Hz.

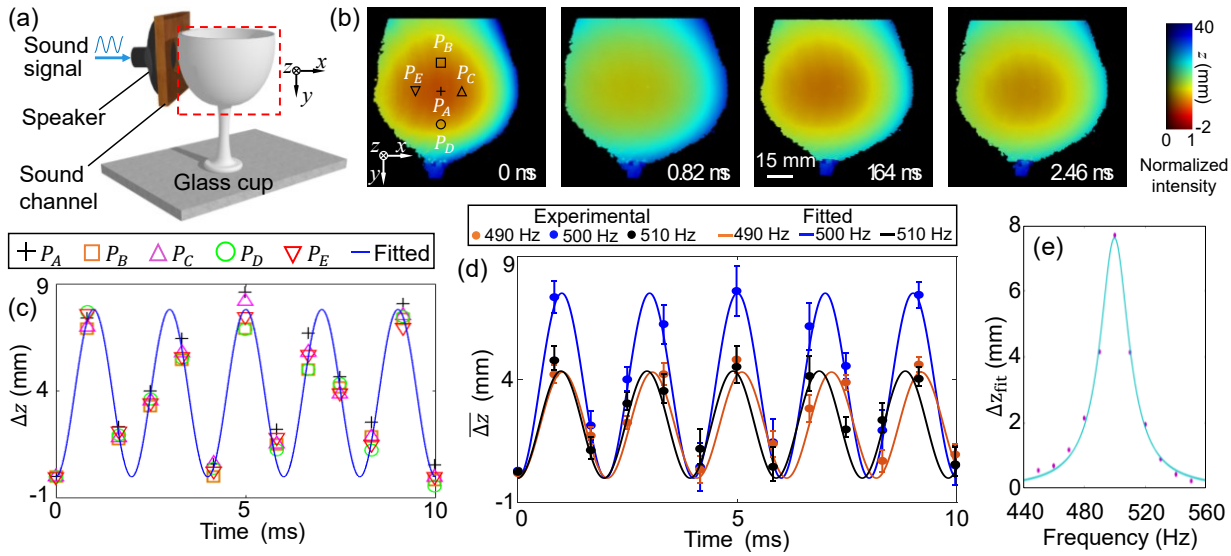

Figure 4.6 TIA-BLIP of sound-induced vibration on glass.

(a) Schematic of the experimental setup. The field of view is marked by the red dashed box. (b) Four reconstructed 3D images of the cup driven by a 500-Hz sound signal. (c) Evolution of the depth change of five points marked in the first panel of (b) with the fitted result. (d) Evolution of the averaged depth change with the

fitted results under the driving frequencies of 490 Hz, 500 Hz, and 510Hz. Error bar: standard deviation of Δz calculated from the five selected pixels. (e) Response of the depth displacements to sound frequencies. The cyan curve is a fitted result of a Lorentz function.

4.3.5 Imaging of glass breakage

To further apply TIA-BLIP to recording non-repeatable 3D dynamics, we imaged the process of glass breaking by a hammer (the full sequence is shown in Visualization 4.4). As displayed in Fig. 4.7(a), the growth of cracks and the burst of fragments with different shapes and sizes were clearly shown in the reconstructed 3D images. The time courses of velocities of four fragments [marked by F_{A} to F_{D} in Fig. 4.7(a)] are plotted in Fig. 4.7(b). The velocities in the y axis are considerably small compared to the other two directions, which indicates the impact of the hammer force was exerted on the x-z plane. v_y of fragments F_{A} and F_{C} shows that they moved upward until $13\ \mathrm{ms}$ and fell afterward. v_y of fragments F_{B} and F_{D} reveals that they fell onto the remaining base of the cup at $15\ \mathrm{ms}$ and kept sliding down on the surface. The data of v_z illustrates that F_{A} and F_{C} moved closer to the cameras, which were directly driven by the hammer's force. However, F_{B} and F_{D}, which collided with other pieces, maintaining their positive directions in v_z to move away from the cameras. The corresponding accelerations are displayed as in Fig. 4.7(c), which indicates the influence of both the main strike and the ensuing collision among different fragments. At $14\ \mathrm{ms}$, the collision with other fragments, which applied an impact along the $+x$ direction, dominated the acceleration direction for all four tracked fragments. In contrast, at $15\ \mathrm{ms}$, another collision produced an impact in the $-x$ direction, causing a sharp decrease in the acceleration for F_{A} and F_{C}. In addition, the direction of acceleration for F_{D} along the y-axis changed several times, which is attributed in several collisions of F_{D} with the base of the glass cup while sliding down.

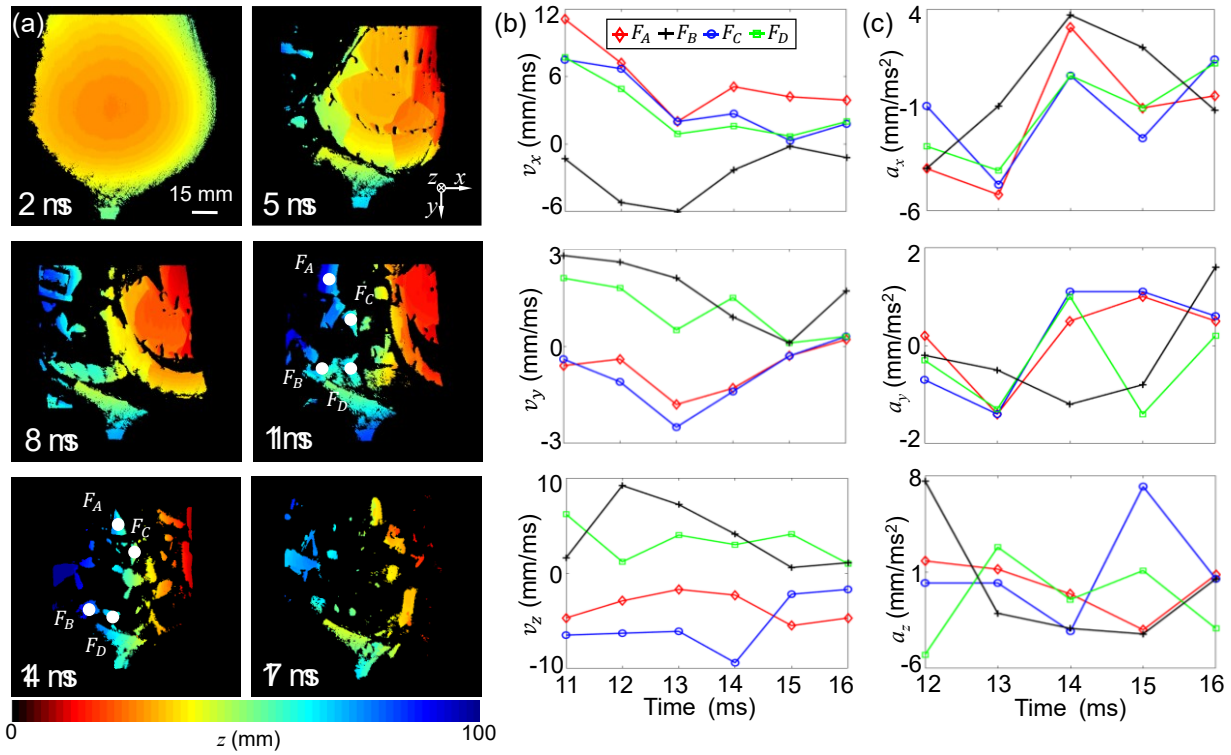

Figure 4.7 TIA-BLIP of glass breakage.

(a) Six reconstructed 3D images showing a glass cup broken by a hammer. (b) Evolution of 3D velocities of four selected fragments marked in the fourth and fifth panels in (a). (c) Evolution of the corresponding 3D accelerations of the four selected fragments.

4.4 Discussion and conclusions

We have developed TIA-BLIP with a kfps-level 3D imaging speed over an FOV of up to 180×130 mm^2 (corresponding to 1180×860 pixels) in captured images. This technique implements TIA in multi-view 3D PSFPP systems, which allows each camera capturing half of the sequence of phase-shifting fringes. Leveraging the characteristics indicated in the intensity matching condition [i.e., Eq. (4.5)], the newly developed algorithm applies constraints in geometry and phase to find the matching pair of points in the main and auxiliary cameras and guides phase unwrapping to extract the depth information. TIA-BLIP has empowered the 3D visualization of glass vibration induced by sound and the glass breakage by a hammer.

TIA-BLIP possesses many advantages. First, TIA eliminates the redundant capture of fringe patterns in data acquisition. The roles of the main camera and the auxiliary camera are interchangeable. Despite demonstrated only with high cameras, TIA-BLIP is a universal imaging paradigm easily adaptable to other multi-view PSFPP systems. Second, TIA reduces the

workload for each camera employed in the multi-view systems. The freed capacity is used to enhance the technical specifications in PSFPP. In particular, at a certain frame rate, more pixels on the sensors of the deployed cameras can be used, which increases the imaging FOV. Alternatively, if the FOV is fixed, TIA supports these cameras to have higher frame rates, which thus increases the 3D imaging speed. Both advantages shed light on implementing TIA-BLIP with an array of cameras to simultaneously accomplishing high accuracy and high speed 3D imaging over a larger FOV. Third, the two cameras deployed in the current TIA-BLIP system are placed side by side. Compared with the existing dual-view PSFPP systems that mostly place the cameras at different sides of the projector, the arrangement in TIA-BLIP circumvents the intensity difference induced by the directional scattering light from the 3D object and reduces the shadow effect by occlusion. Both merits support robust pixel matching in the image reconstruction algorithm to recover 3D information on non-Lambertian surfaces.

The future work will be carried out in the following aspects. First, we plan to further improve TIA-BLIP's imaging speed and FOV by separating the workload to an array of cameras, by implementing a faster DMD, and by using a more powerful laser. Moreover, we will implement depth-range estimation and on-line feedback to reduce the time in candidate discovery. Furthermore, parallel computing will be used to increase the speed of image reconstruction toward real-time operation (Gao & Kemao, 2012). Finally, to robustly image 3D objects with different sizes, we will generate fringe patterns with adaptive periods by using a slit or a pinhole array as the spatial filter (Chang *et al.*, 2013). Automated size calculation (Hoppe *et al.*, 2012) will also be integrated into the imaging processing software to facilitate in determining the proper fringe period.

Besides technical improvement, we will continue exploring new applications of TIA-BLIP. For example, it could be integrated into structure illumination microscopy (Qian *et al.*, 2015) and frequency-resolved multi-dimensional imaging (Dorozynska *et al.*, 2020). TIA-BLIP could also be implemented in the study of the dynamic characterization of glass in its interaction with the external forces in non-repeatable safety test analysis (Bedon *et al.*, 2019; Haldimann *et al.*, 2008; Ramos *et al.*, 2013). As another example, TIA-BLIP could trace and recognize the hand gesture in 3D space to provide information for human-computer interaction (Li, 2012). Furthermore, in robotics, TIA-BLIP could provide a dual-view 3D vision for object tracking and reaction guidance (Huang *et al.*, 2018). Finally, TIA-BLIP can function as an imaging accelerometer for vibration monitoring in rotating machinery (Randall, 2004) and for behavior quantification in biological science (Van Walsum *et al.*, 2020).

4.5 References

Bedon C, Fasan M & Amadio C (2019) Vibration analysis and dynamic characterization of structural glass elements with different restraints based on Operational Modal Analysis. *Buildings* 9(1):13.

Bouguet J-Y (*Camera Calibration Toolbox for Matlab.* http://www.vision.caltech.edu/bouguetj/calib_doc/

Bräuer-Burchardt C, Munkelt C, Heinze M, Kühmstedt P & Notni G (2011) Using geometric constraints to solve the point correspondence problem in fringe projection based 3D measuring systems. *International Conference on Image Analysis and Processing.* Springer, p 265-274.

Chang C, Liang J, Hei D, Becker MF, Tang K, Feng Y, Yakimenko V, Pellegrini C & Wu J (2013) High-brightness X-ray free-electron laser with an optical undulator by pulse shaping. *Optics express* 21(26):32013-32018.

Dorozynska K, Kornienko V, Aldén M & Kristensson E (2020) A versatile, low-cost, snapshot multidimensional imaging approach based on structured light. *Optics Express* 28(7):9572-9586.

Gao W & Kemao Q (2012) Parallel computing in experimental mechanics and optical measurement: A review. *Optics and Lasers in Engineering* 50(4):608-617.

Geng J (2011) Structured-light 3D surface imaging: a tutorial. *Adv. Opt. Photon.* 3(2):128-160.

GmbH O (2018) User Manual CP70-1HS-M/C-1900 4.

Gorthi SS & Rastogi P (2010) Fringe projection techniques: whither we are? *Optics and Lasers in Engineering* 48(IMAC-REVIEW-2009-001):133-140.

Haldimann M, Luible A & Overend M (2008) *Structural use of glass.* labse,

Hartley R & Zisserman A (2003) *Multiple view geometry in computer vision.* Cambridge university press,

Hoppe C, Klopschitz M, Rumpler M, Wendel A, Kluckner S, Bischof H & Reitmayr G (2012) Online Feedback for Structure-from-Motion Image Acquisition. *BMVC.* p 6.

Hornbeck LJ (1997) Digital light processing for high-brightness high-resolution applications. *Projection Displays III.* International Society for Optics and Photonics, p 27-41.

Huang S, Shinya K, Bergström N, Yamakawa Y, Yamazaki T & Ishikawa M (2018) Dynamic compensation robot with a new high-speed vision system for flexible manufacturing. *The International Journal of Advanced Manufacturing Technology* 95(9-12):4523-4533.

Hyun J-S & Zhang S (2017) Superfast 3D absolute shape measurement using five binary patterns. *Optics and Lasers in Engineering* 90:217-224.

Ishii I, Yamamoto K, Doi K & Tsuji T (2007) High-speed 3D image acquisition using coded structured light projection. *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems.* IEEE, p 925-930.

Jiang C, Bell T & Zhang S (2016) High dynamic range real-time 3D shape measurement. *Optics express* 24(7):7337-7346.

Jiang C, Kilcullen P, Liu X, Gribben J, Boate A, Ozaki T & Liang J (2020) Real-time high-speed three-dimensional surface imaging using band-limited illumination profilometry with a CoaXPress interface. *Optics Letters* 45(4):964-967.

Jinyang L (2020) Punching holes in light: Recent progress in single-shot coded-aperture optical imaging. *Reports on Progress in Physics*.

Karpinsky NL, Hoke M, Chen V & Zhang S (2014) High-resolution, real-time three-dimensional shape measurement on graphics processing unit. *Optical Engineering* 53(2):024105.

Kilcullen P, Jiang C, Ozaki T & Liang J (2020) Camera-free three-dimensional dual photography. *Optics Express* 28(20):29377-29389.

Lei C, Wu Y, Sankaranarayanan AC, Chang S-M, Guo B, Sasaki N, Kobayashi H, Sun C-W, Ozeki Y & Goda K (2017) GHz optical time-stretch microscopy by compressive sensing. *IEEE Photonics Journal* 9(2):1-8.

Lei S & Zhang S (2009) Flexible 3-D shape measurement using projector defocusing. *Optics Letters* 34(20):3080-3082.

Li B, Ou P & Zhang S (2014a) High-speed 3D shape measurement with fiber interference. *Interferometry XVII: Techniques and Analysis.* International Society for Optics and Photonics, p 920310.

Li B, Wang Y, Dai J, Lohry W & Zhang S (2014b) Some recent advances on superfast 3D shape measurement with digital binary defocusing techniques. *Optics and Lasers in Engineering* 54:236-246.

Li D, Zhao H & Jiang H (2010) Fast phase-based stereo matching method for 3D shape measurement. *2010 International Symposium on Optomechatronic Technologies.* IEEE, p 1-5.

Li Y (2012) Hand gesture recognition using Kinect. *2012 IEEE International Conference on Computer Science and Automation Engineering.* IEEE, p 196-199.

Li Z, Zhong K, Li YF, Zhou X & Shi Y (2013) Multiview phase shifting: a full-resolution and high-speed 3D measurement framework for arbitrary shape dynamic objects. *Optics letters* 38(9):1389-1391.

Liang J, Becker MF, Kohn RN & Heinzen DJ (2012a) Homogeneous one-dimensional optical lattice generation using a digital micromirror device-based high-precision beam shaper. *Journal of Micro/Nanolithography, MEMS, and MOEMS* 11(2):023002.

Liang J, Kohn Jr RN, Becker MF & Heinzen DJ (2009) 1.5% root-mean-square flat-intensity laser beam formed using a binary-amplitude spatial light modulator. *Applied optics* 48(10):1955-1962.

Liang J, Wu S-Y, Kohn RN, Becker MF & Heinzen DJ (2012b) Grayscale laser image formation using a programmable binary mask. *Optical Engineering* 51(10):108201.

Liu X, Liu J, Jiang C, Vetrone F & Liang J (2019) Single-shot compressed optical-streaking ultra-high-speed photography. *Optics letters* 44(6):1387-1390.

Qian J, Lei M, Dan D, Yao B, Zhou X, Yang Y, Yan S, Min J & Yu X (2015) Full-color structured illumination optical sectioning microscopy. *Scientific reports* 5:14513.

Ramos A, Pelayo F, Lamela M, Canteli AF, Huerta C & Acios A (2013) Evaluation of damping properties of structural glass panes under impact loading. *COST Action TU0905 Mid-Term Conference on Structural Glass; Belis, J., Louter, C., Mocibob, D., Eds.*

Randall RB (2004) State of the art in monitoring rotating machinery-part 1. *Sound and vibration* 38(3):14-21.

Su X & Zhang Q (2010) Dynamic 3-D shape measurement method: a review. *Optics and Lasers in Engineering* 48(2):191-204.

Unser M, Aldroubi A & Eden M (1991) Fast B-spline transforms for continuous image representation and interpolation. *IEEE Transactions on Pattern Analysis & Machine Intelligence* (3):277-285.

Van der Jeught S & Dirckx JJ (2016) Real-time structured light profilometry: a review. *Optics and Lasers in Engineering* 87:18-31.

Van Walsum TA, Perna A, Bishop CM, Murn CP, Collins PM, Wilson RP & Halsey LG (2020) Exploring the relationship between flapping behaviour and accelerometer signal during ascending flight, and a new approach to calibration. *Ibis* 162(1):13-26.

Wang Y & Zhang S (2011) Superfast multifrequency phase-shifting technique with optimal pulse width modulation. *Optics Express* 19(6):5149-5155.

Yin W, Feng S, Tao T, Huang L, Trusiak M, Chen Q & Zuo C (2019) High-speed 3D shape measurement using the optimized composite fringe patterns and stereo-assisted structured light system. *Optics express* 27(3):2411-2431.

Zhang S (2018) Absolute phase retrieval methods for digital fringe projection profilometry: A review. *Opt. Lasers Eng.* 107:28-37.

Zhang S & Huang PS (2006) Novel method for structured light system calibration. *Optical Engineering* 45(8):083601.

Zhang S, Royer D & Yau S-T (2006) GPU-assisted high-resolution, real-time 3-D shape measurement. *Optics Express* 14(20):9120-9129.

Zuo C, Chen Q, Gu G, Feng S & Feng F (2012) High-speed three-dimensional profilometry for multiple objects with complex shapes. *Optics express* 20(17):19493-19510.

5 BLIP OF ROBUST VIDEO-RATE MULTI-SCALE 3D IMAGING

Multi-scale band-limited illumination profilometry for robust three-dimensional surface imaging at video rate

Profilometrie d'illumination a bande limitee multi-echelle pour une imagerie de surface tridimensionnelle robuste a la cadence video

Authors:

Cheng Jiang, Patrick Kilcullen, Yingming Lai, Siqi Wang, Tsuneyuki Ozaki, and Jinyang Liang

Centre Énergie Matériaux Télécommunications, Institut National de la Recherche Scientifique, 1650 boulevard Lionel-Boulet, Varennes, Québec J3X1P7, CANADA

Publication:

Optics Express Vol. 30, Issue 11, pp 19824-19838 (2022)

Submitted 1 Mar 2022, accepted 9 May 2022

https://doi.org/10.1364/OE.457502

Contribution of authors:

Cheng Jiang built the system, developed the software, conducted the experiments, analyzed the data, and wrote the manuscript. Patrick Kilcullen contributed to the algorithm development. Yingming Lai and Siqi Wang conducted some of the experiments. Tsuneyuki Ozaki provided editorial input in the manuscript preparation. Jinyang Liang initiated the project, proposed the concept, contributed to experimental design, and supervised the project.

Abstract

Dynamic three-dimensional (3D) surface imaging by phase-shifting fringe projection profilometry has been widely implemented in diverse applications. However, existing techniques fall short in simultaneously providing the robustness in solving spatially isolated 3D objects, the tolerance of large variation in surface reflectance, and the flexibility of tunable working distances with meter-square-level fields of view (FOVs) at video rate. In this work, we overcome these limitations by developing multi-scale band-limited illumination profilometry (MS-BLIP). Supported by the synergy of dual-level intensity projection, multi-frequency fringe projection, and an iterative method for distortion compensation, MS-BLIP can accurately discern spatially separated 3D objects with highly varying reflectance. MS-BLIP is demonstrated by dynamic 3D imaging of a translating engineered box and a rotating vase. With an FOV of up to 1.7 m × 1.1 m and a working distance of up to 2.8 m, MS-BLIP is applied to capturing full human-body movements at video rate.

5.1 Introduction

Three-dimensional (3D) surface imaging techniques (Gilles *et al.*, 2016; Hyun & Zhang, 2017; Yu & Xu, 2010; Zuo *et al.*, 2018) have found diverse applications, including industrial manufacturing, archaeological inspection, entertainment, and biomedicine (Jiang *et al.*, 2015; Li & Zhang, 2017; Qian *et al.*, 2019; Wu *et al.*, 2021; Zhong *et al.*, 2015). The incessant demands of higher accuracy in extracting surface spatial information in complex scenes pose strict requirements to existing 3D imaging methods. For example, a millimeter-level spatial resolution is desired in reliable 3D facial recognition for attendance checks (Jha, 2007). Meanwhile, video-rate 3D imaging is necessary for fluid-flag interaction analysis in applied bionics (Alben & Shelley, 2008). Moreover, spatially isolated 3D objects with large depth discontinuities must be identified for *in-situ* monitoring and robotic classification (Mazzetto *et al.*, 2019). Furthermore, a high dynamic range is indispensable to measure highly reflective objects in defect inspection of steel castings for the automobile industry (Kosmopoulos & Varvarigou, 2001). Finally, a large field of view (FOV), ideally covering the full human body, is needed to detect and collect 3D movements for virtual reality gaming (Cui *et al.*, 2012).

Among existing techniques, phase-shifting fringe projection profilometry (PSFPP) has proven to be a potent technique for 3D surface imaging (Zhang, 2010). In the most widely used configuration, PSFPP works by first projecting sets of phase-shifting sinusoidal fringe patterns onto 3D objects and then analyzing deformed structure images reflected from the objects to retrieve 3D surface information (Geng, 2011). In contrast to other structured-light 3D imaging techniques (Hall-Holt & Rusinkiewicz, 2001; Huang *et al.*, 2005; Ishii *et al.*, 2007; Jia *et al.*, 2007), PSFPP uses sinusoidal fringe patterns to encode grayscale phase values for each pixel, followed by phase unwrapping to extract 3D information with a high depth resolution (Zhang, 2018).

The speed of sinusoidal fringe projection is one of the most important specifications in PSFPP. These patterns are commonly generated by digital micromirror devices (DMDs) (Liang *et al.*, 2012). Though being a binary amplitude spatial light modulator (Liang), the DMD can generate grayscale sinusoids in various ways. The conventional dithering method forms a grayscale image by controlling the average reflectance of each micromirror over time, which clamps the projection rate at hundreds of hertz (Song, 2013). To improve the fringe projection speed, binary defocusing techniques (Hyun *et al.*, 2017; Li *et al.*, 2014) and band-limited illumination profilometry (BLIP) (Jiang *et al.*, 2020a; Jiang *et al.*, 2020b) have been developed, both of which can produce a grayscale sinusoidal pattern from a single binary DMD mask. Their

fringe pattern projection speeds can keep up with the DMD's refreshing rate [up to tens of kilohertz (kHz)], showing promise for high-speed 3D visualizations.

Although overcoming the limitation in projection speed, binary defocusing could bring instability in pattern quality. In this method, quasi-sinusoidal fringe patterns are generated behind the system's image plane. Consequently, DMD's uneven surface could induce image distortion to the defocused binary DMD masks at any unconjugate plane, which may decrease the 3D measurement accuracy, especially under coherent illumination (Parthiban *et al.*, 2016). In addition, because the proper defocusing degree to eliminate the high-order harmonics from the binary mask is frequency-dependent (Wang *et al.*, 2018), this technique is less compatible with sinusoidal patterns of different periods. Furthermore, the defocusing operation reduces the contrast of quasi-sinusoidal fringes (Li *et al.*, 2017), especially at the far end of the measurement volume. Difficult to maintain the defocusing degree for fringe patterns with different frequencies at a fixed image plane, binary defocusing techniques pose challenges in imaging spatially isolated 3D objects (Zuo *et al.*, 2016). Finally, the binary defocusing technique compromises the depth-sensing range (Hyun *et al.*, 2017). Hence, it has not reached a cubic meter (m^3)-level measurement volume, falling short in multi-scale 3D imaging with tunable working distances and FOVs.

BLIP can overcome the limitations in binary defocusing techniques. It generates sinusoids by controlling the optical bandwidth of a 4*f* imaging system. As a result, regardless of their frequencies, sinusoidal fringe patterns are always generated on the image plane of the DMD. This configuration eliminates the distortion brought by uneven DMD surface and maintains the contrast of projected fringe patterns. It is also inherently compatible with multi-frequency fringe projection, varying working distances, and different FOVs. Despite these attractive advantages, to date, BLIP has not yet been demonstrated for imaging 3D objects with highly reflective surfaces. Meanwhile, BLIP has only been applied with the single-frequency fringe projection method, which is less robust for reconstructing the 3D information of spatially separated objects. Finally, limited by the power of the light sources and the focal lengths of the lenses on the projector and the camera in the existing apparatus (Jiang *et al.*, 2021), BLIP's FOV has not yet reached the meter-square (m^2) level.

To surmount these limitations, in this paper, we report multi-scale (MS) BLIP for robust 3D surface imaging at video rate. MS-BLIP implements multi-frequency fringe projection with the associated phase unwrapping, which enables imaging spatially isolated 3D objects. Meanwhile, MS-BLIP adopts dual-level intensity projection to enhance its dynamic range, which allows

recovering 3D information from surfaces with high reflectance. Moreover, MS-BLIP applies an iterative method for distortion compensation, which improves the 3D reconstruction quality over a m^2-level FOV. Finally, MS-BLIP is demonstrated for video-rate 3D surface measurements at varying working distances of up to 2.8 m with tunable FOVs of up to 1.7 m × 1.1 m.

5.2 Method

5.2.1 System setup of MS-BLIP

The schematic of the MS-BLIP system is shown in Fig. 5.1. A pulsed laser with an average power of 400 mW (AONano 559-0.2-10, 559 nm, Advanced Optowave) is used as the light source. After expansion and collimation, the laser beam is directed to a 0.45" DMD (AJD-4500, Ajile Light Industries) at an incident angle of ~24° to its surface normal. Binary fringe masks, generated by an error diffusion algorithm from their corresponding grayscale patterns (Chang *et al.*, 2013), are loaded onto the DMD and displayed at up to 1 kHz. A band-limited 4f imaging system that consists of two lenses with a same focal length as of 75 mm (Lens 1 and Lens 2 in Fig. 5.1; AC254-075-A, Thorlabs) and one pinhole converts these binary patterns to grayscale fringes at the intermediate image plane. The smallest period in the used sinusoidal fringe patterns is 388.8 μm (i.e., 36 DMD pixels), which demands a 150-μm-diameter pinhole (P150K, Thorlabs) to pass the spatial frequency components of these patterns while filtering all noise induced by the digital halftoning (Lau & Arce, 2018). A dove prism (PS992M, Thorlabs), placed between Lens 2 and the intermediate image plane, rotates the generated fringe patterns to match the aspect ratio of the targeted scene. Then, a camera lens (AF-P DX NIKKOR 10-20mm f/4.5-5.6G VR, Nikon) projects these fringe patterns onto 3D objects. The deformed structure images are captured by a high-speed CMOS camera (CP70-1HS-M-1900, Optronis) with a camera lens (AZURE-3520MX5M, AZURE Photonics). Synchronized by the DMD's trigger signal, the acquired images are transferred to a computer via a CoaXPress cable connected to a frame grabber (Cyton-CXP, Bitflow).

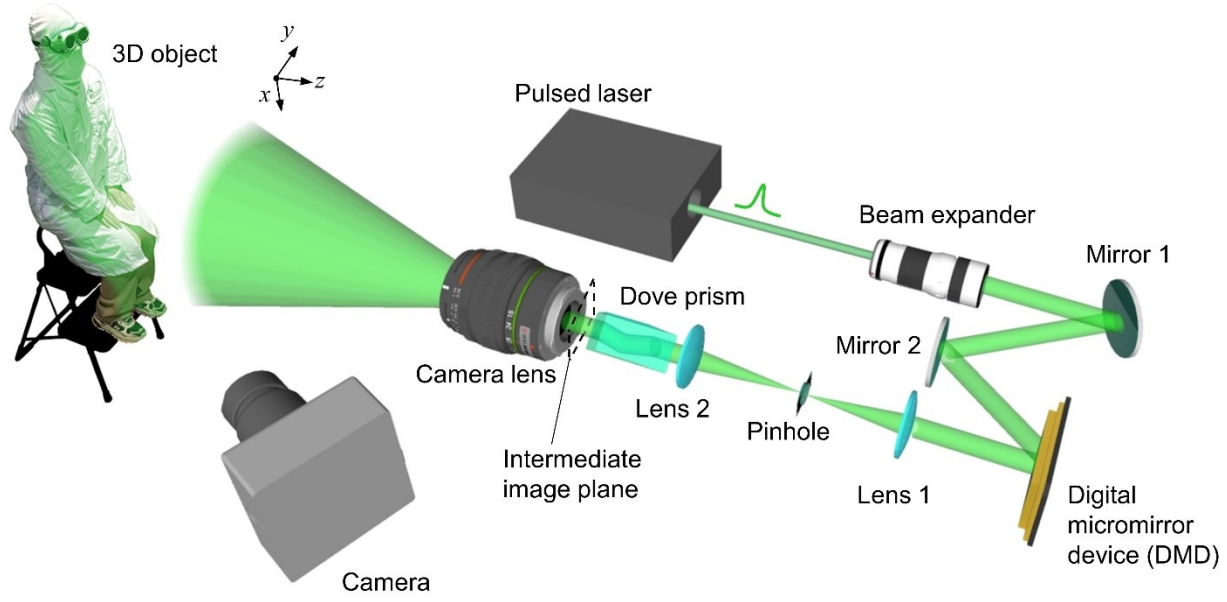

Figure 5.1 Schematic of the MS-BLIP system.

5.2.2 Operating principle of MS-BLIP

System modeling:

The projective behaviors of the camera and the projector in the MS-BLIP system are modeled based on the "pinhole model" with the consideration of image distortion. For a 3D point (x, y, z) in the world coordinate system, in a perfect imaging condition, the pinhole model describes its corresponding pixel on the camera sensor, $(\hat{u}, \hat{v})$, to be

$$s\begin{bmatrix}\hat{u}\\ \hat{v}\\ 1\end{bmatrix} = A[R\ T]\begin{bmatrix}x\\ y\\ z\\ 1\end{bmatrix} = \begin{bmatrix}f_u & 0 & u_{\mathrm{pp}}\\ 0 & f_v & v_{\mathrm{pp}}\\ 0 & 0 & 1\end{bmatrix}[R\ T]\begin{bmatrix}x\\ y\\ z\\ 1\end{bmatrix}. \tag{5.1}$$

Here, the matrix A contains the camera's intrinsic parameters: f_u and f_v describe the effective focal lengths along the axes of the camera, and $(u_{\mathrm{pp}}, v_{\mathrm{pp}})$ is the coordinate of the camera's principal point. Defined as the camera's extrinsic parameters, R is a 3×3 matrix accounting for rotation, and T is a 3×1 matrix for translation. Finally, s is a scalar factor of numerical extraction of $(\hat{u}, \hat{v})$. Overall, this operation is expressed by $(\hat{u}, \hat{v}) = \mathrm{Proj_c}(x, y, z)$. Similarly, the projection of the 3D point to the projector is modeled by $(\hat{u}'', \hat{v}'') = \mathrm{Proj_p}(x, y, z)$. The superscript of "double prime" is used hereafter to refer to the coordinates for the projector.

To take into consideration of distortion in the acquisition of deformed structure images, we first define the normalized camera coordinate without distortion as

$$\begin{bmatrix}\hat{u}_{\mathrm{n}}\\ \hat{v}_{\mathrm{n}}\end{bmatrix} = \begin{bmatrix}(u-u_{\mathrm{pp}})/f_u\\ (v-v_{\mathrm{pp}})/f_v\end{bmatrix}. \tag{5.2}$$

Then, the distorted normalized camera coordinate $(u_{\mathrm{n}}, v_{\mathrm{n}})$ is determined by (Heikkila & Silvén, 1997)

$$\begin{bmatrix}u_{\mathrm{n}}\\ v_{\mathrm{n}}\end{bmatrix} = \left(1 + d_1\hat{r}_{\mathrm{n}}{}^2 + d_2\hat{r}_{\mathrm{n}}{}^4 + d_3\hat{r}_{\mathrm{n}}{}^6\right)\begin{bmatrix}\hat{u}_{\mathrm{n}}\\ \hat{v}_{\mathrm{n}}\end{bmatrix} + \begin{bmatrix}2d_4\hat{u}_{\mathrm{n}}\hat{v}_{\mathrm{n}} + d_5\left(\hat{r}_{\mathrm{n}}{}^2 + 2\hat{u}_{\mathrm{n}}{}^2\right)\\ 2d_5\hat{u}_{\mathrm{n}}\hat{v}_{\mathrm{n}} + d_4\left(\hat{r}_{\mathrm{n}}{}^2 + 2\hat{v}_{\mathrm{n}}{}^2\right)\end{bmatrix}, \tag{5.3}$$

where d_1,…, d_5 are the camera system distortion coefficients, and $\hat{r}_{\mathrm{n}}^2 = \hat{u}_{\mathrm{n}}^2 + \hat{v}_{\mathrm{n}}^2$. Once $(u_{\mathrm{n}}, v_{\mathrm{n}})$ is computed, the distorted camera pixel coordinate (u, v) is found by using Eq. (5.2). Overall, this operation is expressed as $(u, v) = \mathrm{Dist}_{\mathrm{c}}(\hat{u}, \hat{v})$. Similarly, the distortion for the projector is modeled by $(u'', v'') = \mathrm{Dist}_{\mathrm{p}}(\hat{u}'', \hat{v}'')$.

System calibration:

MS-BLIP's system calibration aims to compute the intrinsic parameters, extrinsic parameters, and image distortion coefficients for both the camera and the projector. In the calibration of the camera, a checkerboard pattern was imaged with 20–40 poses. A MATLAB toolbox (Bouguet) was used to extract the grid corners, which allowed calculating all the calibration parameters. In the calibration of the projector, the projector was treated as another camera to adopt the similar process used in camera calibration (Zhang & Huang, 2006). The method involved capturing additional images of the checkerboard pattern under the illumination of both horizontally and vertically shifted fringe patterns to determine the pixel-wise mapping between the camera view and the projector view. Then, the camera-captured images of the checkerboard were transformed to generate the corresponding images from the perspective of the projector. Finally, the same MATLAB toolbox used in camera calibration was implemented to compute the corresponding calibration parameters of the projector.

Dual-level intensity projection:

To image 3D objects with a large variety of reflectance, projection intensity modulation (Zhang & Yau, 2009) is implemented to improve MS-BLIP's dynamic range. In particular, two normalized intensity levels, empirically chosen to be 1.0 and 0.3, are set for each fringe pattern. For any camera pixel in the acquired deformed structure images, if any values in the sequence of high-intensity projection are saturated, then it is replaced with the corresponding sequence with low intensity.

Coordinate recovery using multi-frequency fringe projection:

MS-BLIP implements the multi-frequency fringe projection method (Tian *et al.*, 2008). The intensity of the camera pixel (u, v) in the captured deformed structure images can be expressed as

$$I_{km}(u,v) = I_{\mathrm{b}}(u,v) + I_{\mathrm{va}}(u,v)\cos[\varphi_m(u,v) - 2\pi k/K]\,. \tag{5.4}$$

Here, $I_{km}(u,v)$ represents the intensity in the k^{th} deformed structure image in the m^{th} sequence. K is the number of phase-shifting steps. In this work, we use four-step phase shifting in six sets of fringes, so that $K = 4$, $k \in [0,3]$, and $m \in [0,5]$. $I_{\mathrm{b}}(u,v)$, $I_{\mathrm{va}}(u,v)$, and $\varphi_m(u,v)$ represent the background, intensity variation, and depth-dependent phase for the m^{th} sequence, respectively. From Eq. (5.4), the calculated $\varphi_m(u,v)$ is wrapped in the interval $(-\pi, \pi]$. Then, phase unwrapping is performed to determine the horizontal coordinate of the projector, which is then used to solve the 3D surface information by triangulation.

In single-frequency fringe projection, a single unwrapped phase map is generated by a two-dimensional weighted phase unwrapping algorithm (Ghiglia & Romero, 1994), which must make assumptions about the surface smoothness. The unwrapped phase value of each pixel is derived according to the phase values within a local neighborhood of this pixel. Consequently, this method falls short in solving the phase ambiguities induced by the depth-discontinued regions or isolated objects (Saldner & Huntley, 1997).

This limitation is lifted by multi-frequency fringe projection that unwraps the phase value of each pixel individually. The values of P_{p}^m in our work were determined by synthetically considering the system's signal-to-noise ratios and the measurement accuracy. The coarsest fringe pattern is chosen to have no more than one period so that its absolute phase, defined as $\tilde{\varphi}_0(u,v)$, equals the wrapped phase map $\varphi_0(u,v)$. The ensuing sets of fringe patterns, i.e., $m > 0$, have the periods of $P_{\mathrm{p}}^m = Round(N_u/2^m)$. Here, N_u is the width of the DMD expressed in the unit of DMD pixel. Because unwrapped phase maps of two adjacent sets of fringe projection have the relationship of $\tilde{\varphi}_m(u,v) = (P_{\mathrm{p}}^{m-1}/P_{\mathrm{p}}^m)\tilde{\varphi}_{m-1}(u,v)$, the absolute phase map for $m > 0$ can be calculated by

$$\tilde{\varphi}_m(u,v) = \varphi_m(u,v) + 2\pi \cdot Round\left[\frac{\left(P_{\mathrm{p}}^{m-1}/P_{\mathrm{p}}^m\right)\tilde{\varphi}_{m-1}(u,v) - \varphi_m(u,v)}{2\pi}\right]. \tag{5.5}$$

Equation (5) shows that, unlike single-frequency phase unwrapping, the phase of each pixel is unwrapped independently, which avoids incorrect depth quantification in analyzing spatially isolated objects (Saldner & Huntley, 1997). Using this method, the unwrapped phase

map of each set of fringe patterns is computed successively, and $\tilde{\varphi}_5(u,v)$ is used to compute the horizontal coordinate of the projector, denoted by u'', by

$$u'' = \tilde{\varphi}_5(u,v)\frac{P_{\mathrm{p}}^5}{2\pi} + \frac{1}{2}(N_u - 1). \tag{5.6}$$

Distortion compensation and 3D point recovery:

Distortion compensation contains two major steps. First, the undistorted camera pixel coordinate $(\hat{u},\hat{v})$ is extracted by performing the inverse operation of the distortion model (i.e., $\mathrm{Dist}_{\mathrm{c}}$). However, the direct inversion of Eq. (5.3) takes the form of a 7th degree polynomial in both $\hat{u}_{\mathrm{n}}$ and $\hat{v}_{\mathrm{n}}$, rendering direct recovery difficult. Thus, we apply a fixed point iteration method (Burden *et al.*, 2010) for computing the undistorted normalized camera coordinate $(\hat{u}_{\mathrm{n}},\hat{v}_{\mathrm{n}})$. With the initial condition of $\begin{bmatrix}\hat{u}_{\mathrm{n},0}\\ \hat{v}_{\mathrm{n},0}\end{bmatrix} = \begin{bmatrix}(u-u_{\mathrm{pp}})/f_u\\ (v-v_{\mathrm{pp}})/f_v\end{bmatrix}$, the i^{th} iteration is described by

$$\begin{bmatrix}\hat{u}_{\mathrm{n},i+1}\\ \hat{v}_{\mathrm{n},i+1}\end{bmatrix} = \frac{\begin{bmatrix}\hat{u}_{\mathrm{n},i}\\ \hat{v}_{\mathrm{n},i}\end{bmatrix} - \begin{bmatrix}2d_4\hat{u}_{\mathrm{n},i}\hat{v}_{\mathrm{n},i} + d_5\left(\hat{r}_{\mathrm{n},i}{}^2 + 2\hat{u}_{\mathrm{n},i}{}^2\right)\\ 2d_5\hat{u}_{\mathrm{n},i}\hat{v}_{\mathrm{n},i} + d_4\left(\hat{r}_{\mathrm{n},i}{}^2 + 2\hat{v}_{\mathrm{n},i}{}^2\right)\end{bmatrix}}{1 + d_1\hat{r}_{\mathrm{n},i}{}^2 + d_2\hat{r}_{\mathrm{n},i}{}^4 + d_3\hat{r}_{\mathrm{n},i}{}^6}, \tag{5.7}$$

where $\hat{r}_{\mathrm{n},i}^2 = \hat{u}_{\mathrm{n},i}^2 + \hat{v}_{\mathrm{n},i}^2$. The corresponding undistorted pixel coordinate $(\hat{u},\hat{v})$ on the camera plane is then calculated by using Eq. (5.2). Overall, this operation is expressed as $(\hat{u},\hat{v}) = \mathrm{DistComp}_{\mathrm{c}}(u,v)$.

The next step is to obtain the corresponding undistorted projector horizontal pixel coordinate $\hat{u}''$. In Eq. (5.1), we define $A[R\ T] = \begin{bmatrix}a_{11} & \cdots & a_{14}\\ \vdots & \ddots & \vdots\\ a_{31} & \cdots & a_{34}\end{bmatrix}$ and $A''[R''\ T''] = \begin{bmatrix}b_{11} & \cdots & b_{14}\\ \vdots & \ddots & \vdots\\ b_{31} & \cdots & b_{34}\end{bmatrix}$. By using $\mathrm{Proj}_{\mathrm{c}}$ and $\mathrm{Proj}_{\mathrm{p}}$, the coordinate of the distortion-compensated 3D point is calculated by triangulation as

$$\begin{bmatrix}x\\ y\\ z\end{bmatrix} = \mathrm{Tri}(\hat{u},\hat{v},\hat{u}'') = \begin{bmatrix}a_{11}-\hat{u}a_{31} & a_{12}-\hat{u}a_{32} & a_{13}-\hat{u}a_{33}\\ a_{21}-\hat{v}a_{31} & a_{22}-\hat{v}a_{32} & a_{23}-\hat{v}a_{33}\\ b_{11}-\hat{u}''b_{31} & b_{12}-\hat{u}''b_{32} & b_{13}-\hat{u}''b_{33}\end{bmatrix}^{-1}\begin{bmatrix}\hat{u}a_{34}-a_{14}\\ \hat{v}a_{34}-a_{24}\\ \hat{u}''b_{34}-b_{14}\end{bmatrix}. \tag{5.8}$$

Ideally, $\hat{u}''$ could be computed by $(\hat{u}'',\hat{v}'') = \mathrm{DistComp}_{\mathrm{p}}(u'',v'')$. However, v'' cannot be calculated from the projection of vertical fringe patterns. To overcome this limitation, we have developed an iterative method that recovers distortion-compensated 3D information without prior knowledge of v'' from fringe measurements. We first estimate the coordinate of the 3D point by the Tri function [i.e., Eq. (5.8)] by using the coordinate $(\hat{u},\hat{v},u'')$. This 3D point is then projected to the projector plane to extract the initial estimate of v'' with the function of $\mathrm{Dist}_{\mathrm{p}}$, which is input

to an iterative algorithm. Each iteration starts with performing the function $\mathrm{DistComp_p}$ by using u'' calculated from Eq. (6) and the variable v'' to compute the distortion compensated projector coordinate $\hat{u}''$, which is then used to extract the 3D coordinate by the function Tri. The method is presented as the following pseudo-code:

Algorithm to recover 3D information with distortion compensation

Input: $(\hat{u}, \hat{v}, u'')$, *maximum iterations N, vertical projector pixel error tolerance TOL, calibration parameters of camera: A, R, T, s,* d_1, d_2, d_3, d_4, d_5;
calibration parameters of projector: A'', R'', T'', s'', d_1'', d_2'', d_3'', d_4'', d_5''

Output: (x, y, z)

Variables: q, h_1, h_2, h_3, h_4, v_q'', $\hat{u}''$

1 *Set* $q = 1$
2 *compute* $(x, y, z) = \mathrm{Tri}(\hat{u}, \hat{v}, u'')$
3 *compute* $(h_1, h_2) = \mathrm{Proj_p}(x, y, z)$
4 *compute* $(\sim, v_q'') = \mathrm{Dist_p}(h_1, h_2)$
5 ***while*** $(q \leq N)$ *do steps 6 – 11*
6 *compute* $(\hat{u}'', \sim) = \mathrm{DistComp_p}(u'', v_q'')$
7 *compute* $(x, y, z) = \mathrm{Tri}(\hat{u}, \hat{v}, \hat{u}'')$
8 *compute* $(h_3, h_4) = \mathrm{Proj_p}(x, y, z)$
9 *compute* $(\sim, v_{q+1}'') = \mathrm{Dist_p}(h_3, h_4)$
10 ***if*** $|v_{q+1}'' - v_q''| \leq TOL$, *go to step 13*
11 *Set* $q = q + 1$
12 ***End***
13 ***Output*** (x, y, z)
14 ***End***

5.3 Results

5.3.1 Demonstration of MS-BLIP

To verify the feasibility of MS-BLIP, we imaged static 3D objects at a working distance of 0.5 m, with an FOV of 0.3 m × 0.2 m. The depth resolution, quantified by the standard deviation of a reconstructed planar surface (Jiang *et al.*), was measured to be 0.3 mm. The lateral resolution, quantified by imaging the sharp edge of the planar surface (Liang *et al.*, 2013), was 0.4 mm.

To demonstrate MS-BLIP's ability to tolerate a large variety in reflectance, we imaged an open box that had a highly reflective matte surface with a weakly reflective foam surface [Fig. 5.2(a)]. The representative fringe masks with two intensity levels are displayed in Fig. 5.2(b). The effect of different intensity levels in deformed structure images is compared in Fig. 5.2(c). The high-intensity projection (i.e., $I_{km}^{1.0}$) results in a large number of saturated pixels in the left region of the matte surface. The low-intensity projection (i.e., $I_{km}^{0.3}$), despite eliminating saturation, fails to resolve the white foam surface and the right bottom part of the matte surface. The dynamic ranges of $I_{km}^{1.0}$ and $I_{km}^{0.3}$ were calculated to be 48.1 dB and 41.8 dB, respectively. In comparison, dual-level intensity projection preserves low-intensity pixels while avoiding any saturation [Fig. 5.2(d)]. With an enhanced dynamic range of 58.6 dB, this method fully recovers this 3D object.

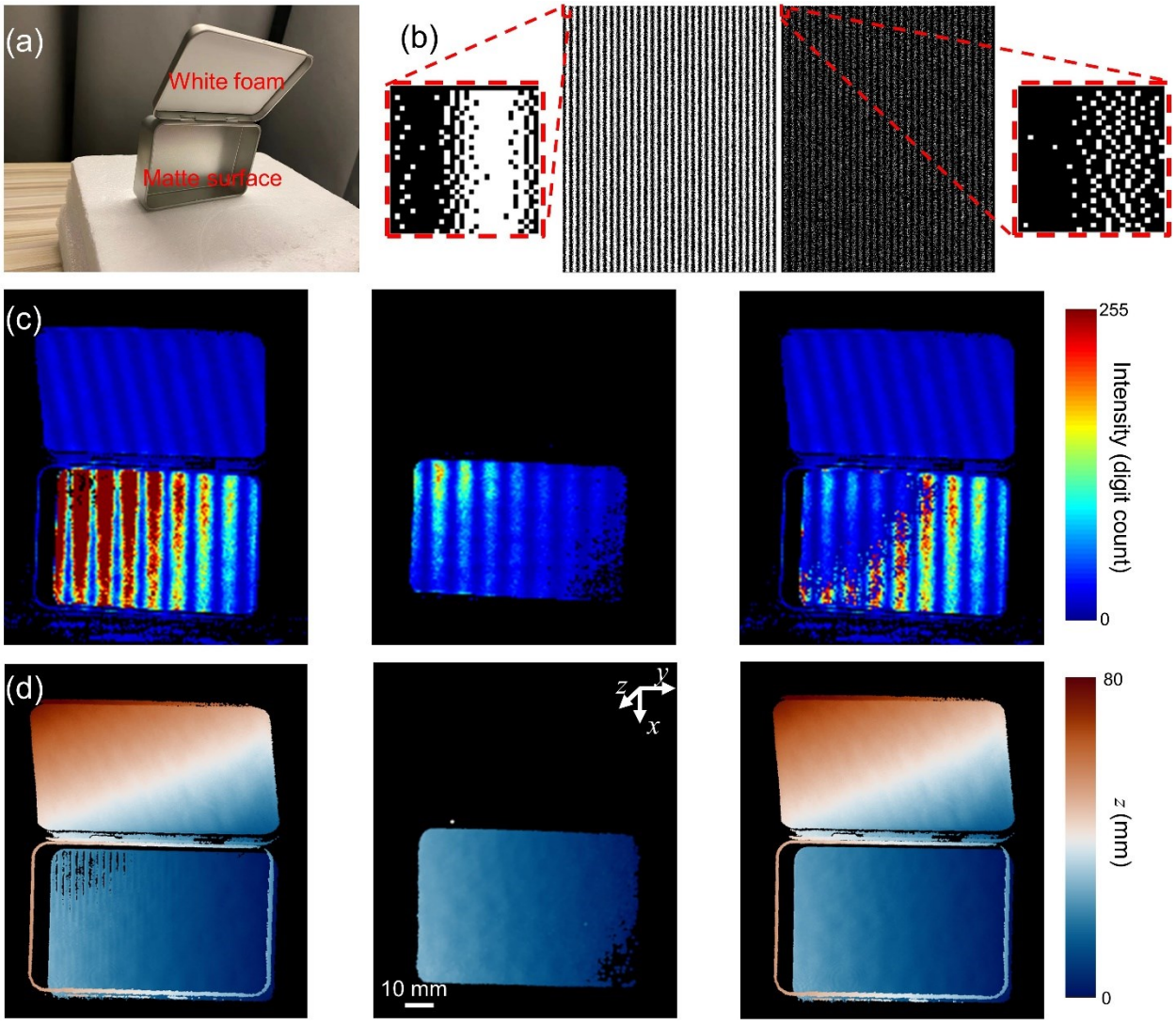

Figure 5.2 MS-BLIP of a 3D object with varying reflectance.

(a) Photo of the 3D object. (b) Representative binary fringe masks with the normalized intensity levels of 1.0 (left) and 0.3 (right). Red dashed boxes: close-up views. (c) Representatives fringe images by using high-intensity projection (left), low-intensity projection (middle), and their combination (right). (d) As (c), but showing the 3D reconstruction results in the three cases.

MS-BLIP's high accuracy in imaging 3D objects with depth discontinuities is proven by comparing the multi-frequency phase unwrapping method applied in MS-BLIP with the standard single-frequency counterpart. As shown in Fig. 5.3(a), a 50-mm-long cylinder with a 15-cm-diameter flat front surface stands on a base. A square foam plate is placed ~230 mm behind the cylinder. The single-frequency method used four phase-shifting fringe patterns with a period of 36 pixels (i.e., P_{p}^{5}) and an additional pattern of the DMD's vertical central line. Figures 5.3(b)–(c) show the captured deformed structure images by using the two methods. Because the distance between the cylinder and the foam plate resulted in a shift of multiple periods of fringes in deformed structure images, single-frequency fringe projection fails to correctly resolve the depth in the left half of the cylinder [Fig. 5.3(d)]. In contrast, displayed in Fig. 5.3(e), the multi-frequency method, adopted in MS-BLIP, accurately recovers the detailed 3D information. Specifically, the calculated depth between the cylinder's front surface to the foam plate is 281.7 mm, which well matches the ground truth (with a deviation of 0.61%).

To demonstrate the necessity of distortion compensation, we analyzed the 3D reconstruction results of a selected region on the foam plate [red dashed box in Fig. 5.3(a)]. As shown in Fig. 5.3(f), the system distortion mistakenly guides the 3D reconstruction, which displays a curved structure of the reconstructed surface. With the iterative method for distortion compensation applied in MS-BLIP, the surface with correct depth information is presented in Fig. 5.3(g).

MS-BLIP can accurately resolve 3D information of fine structures. Figure 5.3(h) shows the 3D reconstruction of the cylinder's flat round surface with two dents [blue dashed box in Fig. 5.3(a)]. The quantitative analysis of selected line profiles shows that the two dents on the flat surface have a depth of 6.0 mm and a diameter of 10.0 mm [Fig. 5.3(i)]. The deviations from the ground truth are 0.3 mm and 0.5 mm, respectively, which show an excellent agreement with the object's configuration.

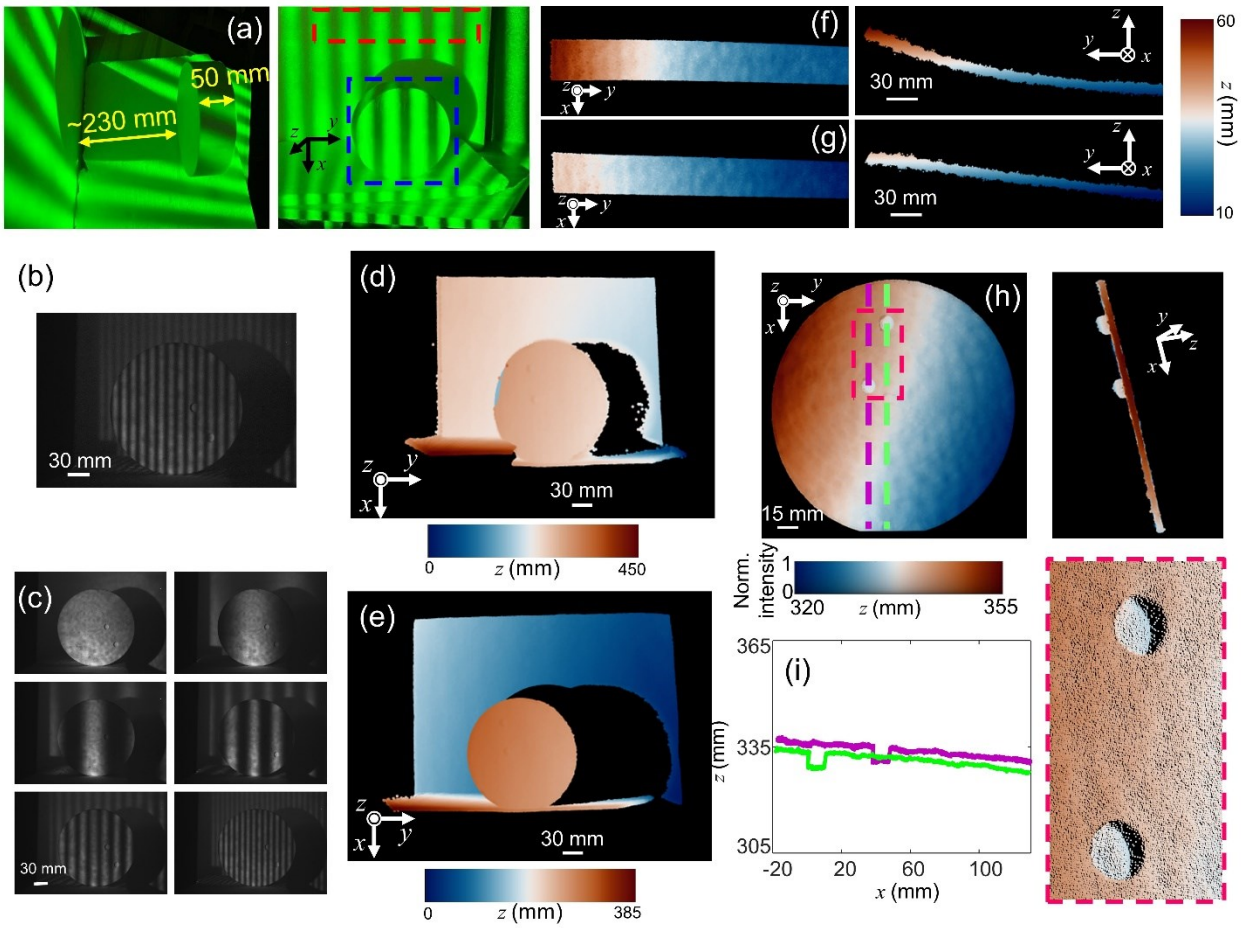

Figure 5.3 MS-BLIP of 3D objects with depth discontinuity.

(a) Side view (left) and front view (right) of the 3D objects. (b) Representative deformed structure image by using single-frequency fringe projection. (c) As (b), but using multi-frequency fringe projection. (d)–(e) 3D reconstruction from single-frequency (d) and multi-frequency (e) fringe projection. (f)–(g) 3D reconstruction without (f) and with (g) distortion compensation of a selected region on the foam plate, indicated by the red dashed box in (a). Front view (left) and side view (right) are displayed. (h) Front view (left) and side view (right) of the reconstructed flat round surface, indicated as the blue dashed box in (a). Red dashed box: Close-up view of the two dents. (i) Selected depth profiles [marked by the purple and green dashed lines in (h)].

5.3.2 MS-BLIP of translational movement

To highlight the potential application of MS-BLIP in automated detection and classification, we imaged an engineered box moving translationally at a speed of ~15 cm/s [Fig. 5.4(a)]. The box used a piece of curved cardboard as the base, on which pipes made of different materials were fixed. On the top row, the left pipe was made of black foam with low reflectance, while the other three metal pipes, as well as the white swelling glue, had high reflectance. MS-BLIP had a 3D imaging speed of 10.4 fps, an FOV of 0.3 m × 0.2 m at a working distance of 0.5 m, a depth resolution of 1.3 mm, and a lateral resolution of 1.1 mm.

Figures 5.4(b)–(d) show the advantage of dual-level intensity projection. For the $I_{km}^{1.0}$ projection, because saturated intensity distorted the pixel-wised sinusoidal intensity profile, fringe residues show up as artifacts on the regions with high reflectance [Fig. 5.4(b)]. Meanwhile, due to the lack of scattered light intensity from the $I_{km}^{0.3}$ projection, part of the black foam pipe was failed to be reconstructed [Fig. 5.4(c)]. In contrast, the dual-level intensity projection enhanced the dynamic range by over 12 dB. Presented in Fig. 5.4(d), the dual-level intensity projection successfully reconstructs the full shape of this 3D object.

Figure 5.4(e) shows the reconstructed 3D images of the moving engineered box at four time-lapse frames with an interval of 960 ms (see the full evolution in Visualization 5.1). MS-BLIP allowed tracking the 3D movements of all spatial points. As an example, the time-resolved positions of two points, selected from the pipes with both low and high reflectance [marked by P_{A} and P_{B} in Fig. 5.4(e)], are shown in Fig. 5.4(f). The results reveal that the two points have linear movement in the *y* axis at 15.4 cm/s but stay relatively stationary in the *z* axis, well matching the pre-set experimental condition.

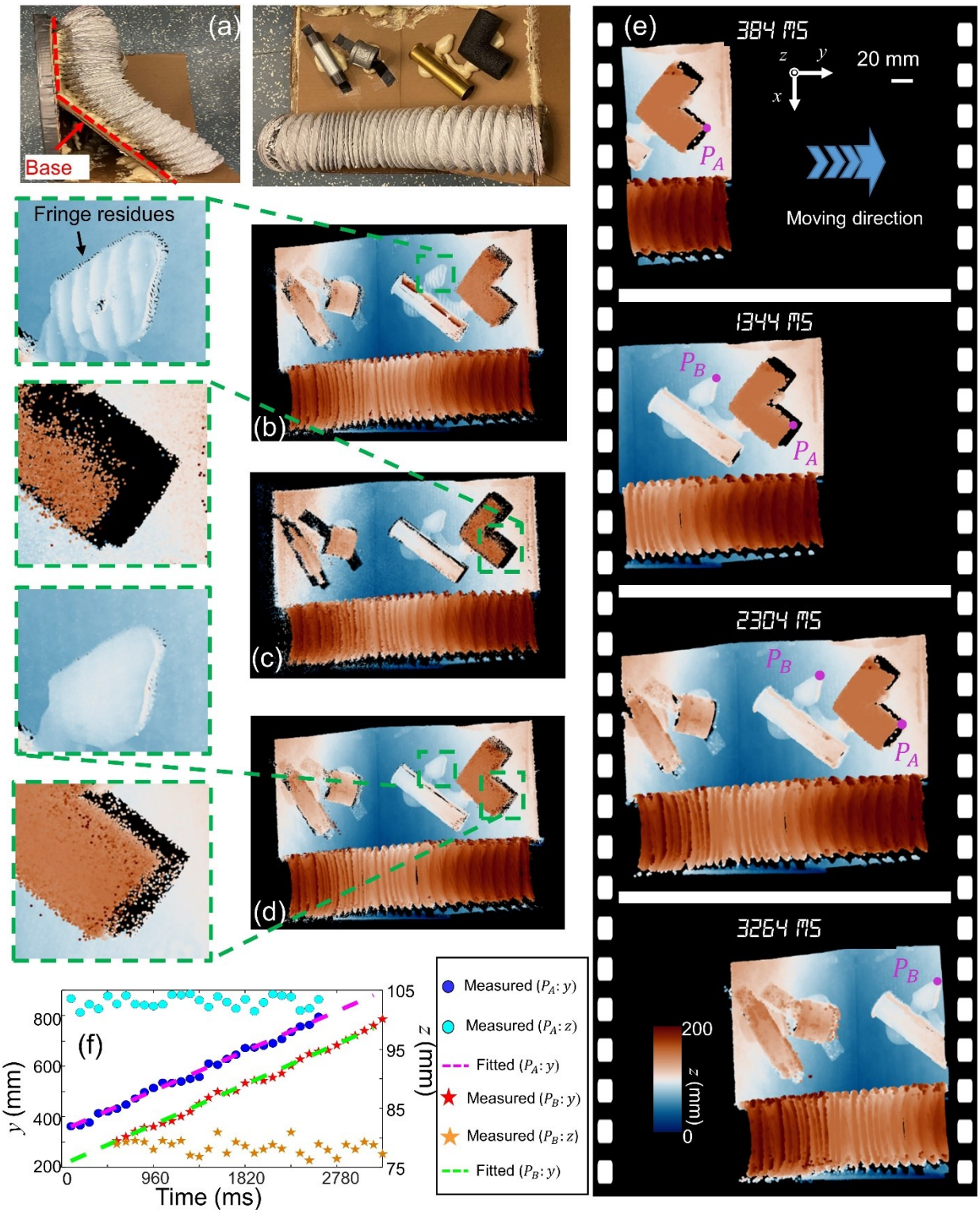

Figure 5.4 MS-BLIP of translational movement of an engineered box.

(a) Photo of the 3D object. (b)–(d) Reconstructed 3D images with high-intensity projection (b), low-intensity projection (c), and their combination (d). Close-up views are shown on the left column. (e) Time-lapse 3D images of the translationally moving engineered box. (f) Traces of two points [marked in (e)]c with fitting.

5.3.3 MS-BLIP of rotational movement

To demonstrate MS-BLIP's potential in industrial inspection, we imaged the rotational movement of a bamboo vase with extending branches [Fig. 5.5(a)]. With a height of 1.3 m, this 3D object was glued on a stand rotating at 0.6 rad/s. To fit the projection region with the desired imaging area, the dove prism was placed at 45°, which results in a 90° rotation of the projected patterns with respect to the fringe masks loaded onto the DMD. MS-BLIP was operated at a working distance of 2 m, with an FOV of 1.5 m × 1.0 m, and at a 3D imaging speed of 20.8 fps. Under these working conditions, the depth resolution was quantified to be 3.7 mm, and the lateral resolution was measured to be 1.7 mm.

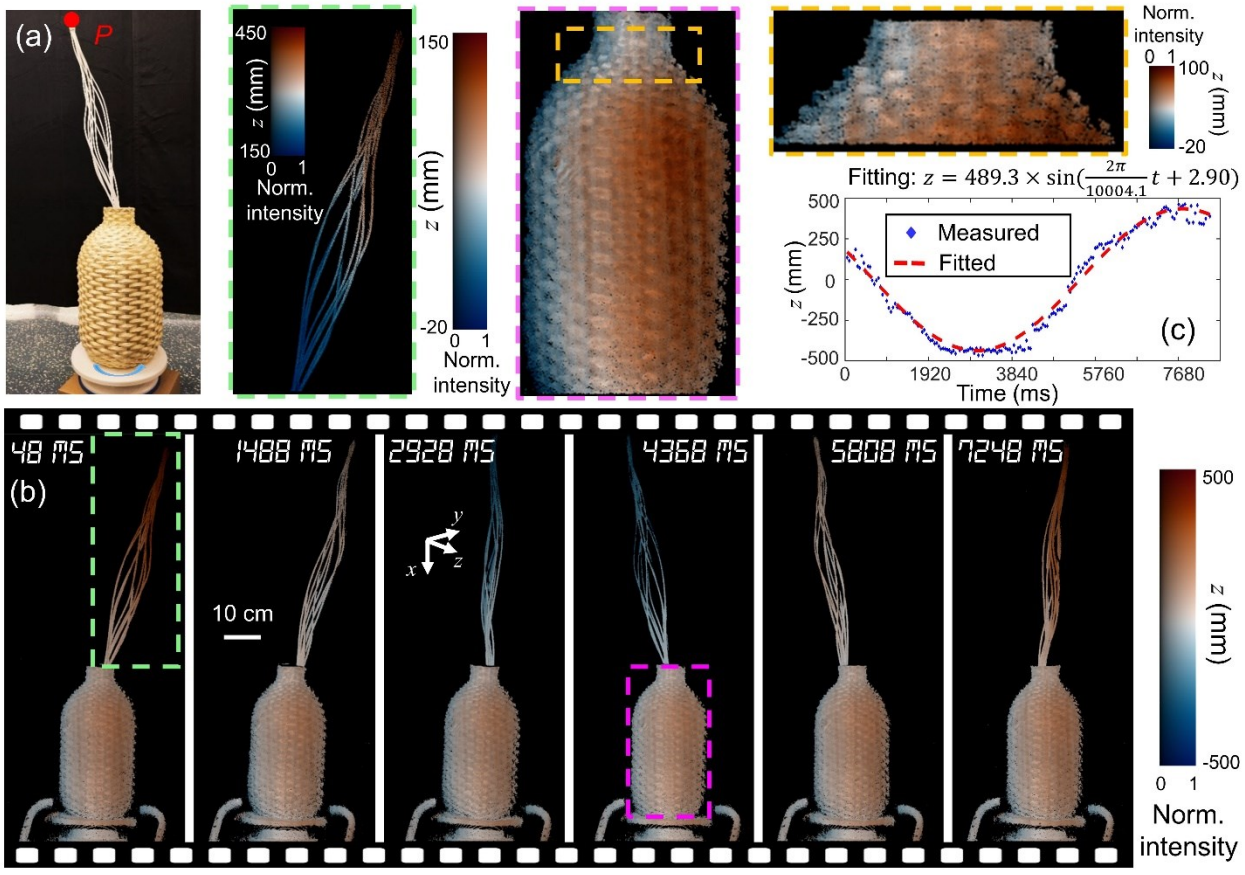

Figure 5.5 MS-BLIP of rotational movement of a vase with branches.

(a) Photo of the 3D object. (b) Time-lapse 3D images of the object. Close-up views are shown in the dashed boxes. (c) Evolution of the depth change of a selected point marked in (a) with sinusoidal fitting.

Figure 5.5(b) shows six time-lapse 3D images of the object (the full sequence is shown in Visualization 5.2). The close-up view of the vase mouth presents detailed structural information on its surface. The depth-encoded color change of the branches reflects the rotation movement of the object. By tracking the 3D position of the tip of the branches [marked by *P* in Fig. 5.5(a)],

MS-BLIP reveals its rotational radius of 489.3 mm and the period of 10004.1 ms [Fig. 5.5(c)], showing a uniform rotation with the pre-set speed.

5.3.4 MS-BLIP of human-body movement

To explore the potential of MS-BLIP in human-computer interaction, we imaged the full-body movements of a volunteer (the full sequence is shown in Visualization 5.3). All the experiments were conducted in accordance with the protocol (Project CÉR-22-649) approved by the human ethics research committee of Institut National de la Recherche Scientifique, Université du Québec. A photo of the scene is shown in Fig. 5.6(a). The volunteer had a height of ~1.6 m, wearing proper protective clothing and a laser goggle. Both hands were exposed to the laser illumination with intensity (and fluence) of 7.8×10^{-7} W/cm^2 (and 3.9×10^{-11} J/cm^2), which is much lower than the maximum permissible exposure (i.e., 0.2 W/cm^2 and 2×10^{-2} J/cm^2) regulated by the ANSI safety limit (Laser Institute of America, 2014). A part of an optical table on the left of this volunteer was also included in the FOV. Akin to the 3D imaging of rotational movement, the dove prism was set to 45°. The MS-BLIP system was operated with a working distance of 2.8 m and an FOV of 1.7 m × 1.1 m. It had a 3D imaging speed of 10.4 fps. The depth and the lateral resolutions were measured to be 4.7 mm and 2.4 mm, respectively.

Figure 5.6(b) shows five time-lapse 3D images of the instantaneous poses of the volunteer. The detailed surface structures of the volunteer's lab coat and hands are illustrated by the close-up views. The edge of an optical table and optomechanical components can also be seen as a static background. As displayed in Fig. 5.6(c), we tracked the 3D positions of two selected points over time: the ring finger’s tip of the volunteer’s left hand and the edge point of a marker on the volunteer's coat. In the experiment, the 3D image acquisition began when the volunteer sat on a stepladder with both hands on the knees. Afterward, the volunteer moved up both hands at 2016 ms, resulting in a change of the fingertip position along all three axes [marked by the arrows for “Hands up” in Fig. 5.6(c)]. The volunteer’s body stayed at the same position until 4128 ms when the volunteer stood up, which results in movements along the -x axis for both points [marked by the arrow for “Stand up” in Fig. 5.6(c)]. In addition, the forward incline movement of the volunteer’s body while standing up was also reviewed according to the position changes along the +z axes. At 6048 ms, the volunteer stretched out both hands, which is responsible for the sudden increase of the y value [marked by the arrow for “Hands out” in Fig. 5.6(c)].

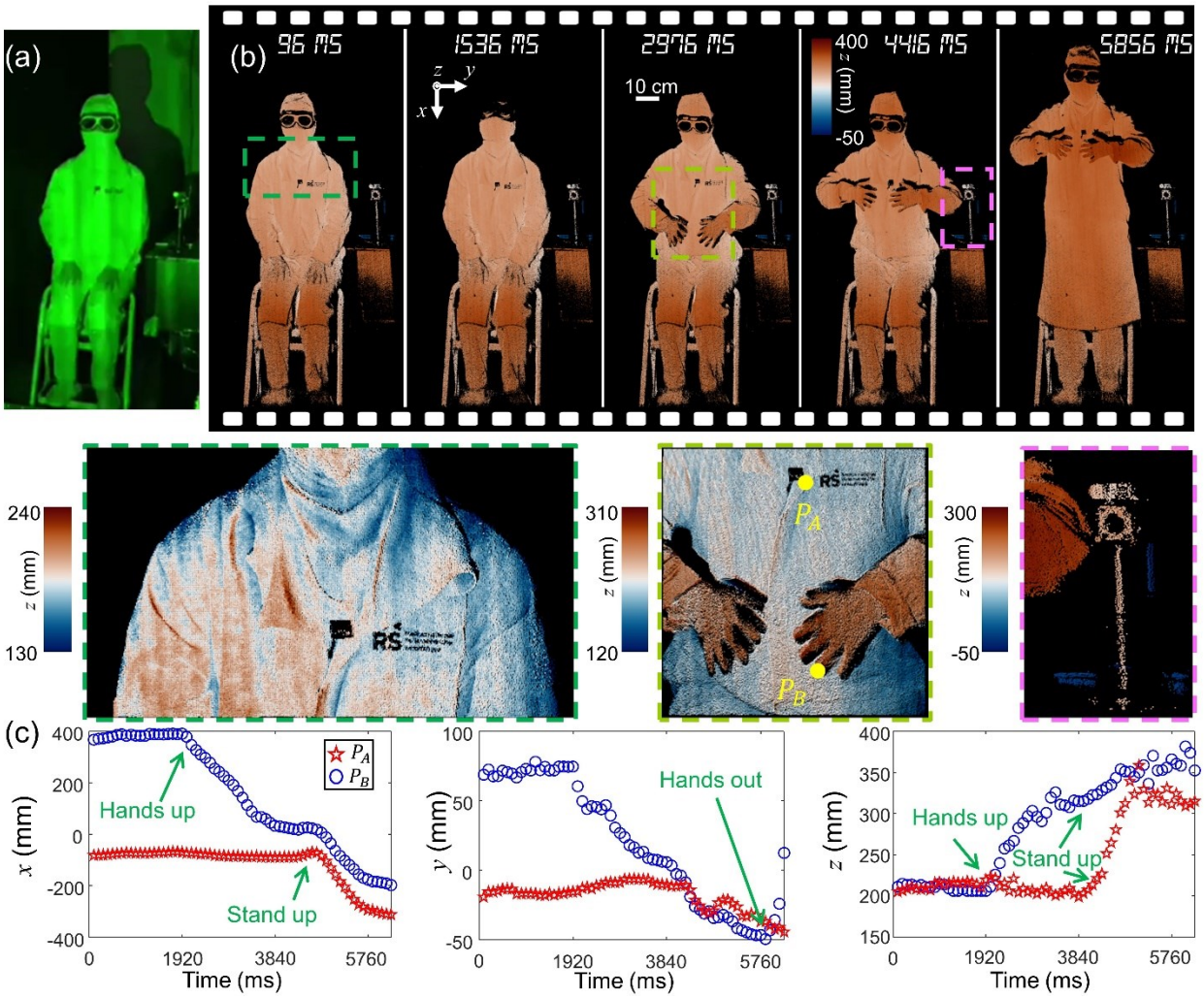

Figure 5.6 MS-BLIP of full human-body movements.

(a) Photo of the 3D object. (b) Five time-lapse reconstructed 3D images. Close-up views are shown in the dashed boxes. (c) Evolution of the 3D positions of the selected points [marked in the middle close-up view of (b)].

5.4 Discussion and conclusions

We have developed MS-BLIP for robust multi-scale 3D surface imaging at video rate. Compared to existing BLIP systems, MS-BLIP employs dual-level intensity projection to tolerate the variation in 3D objects' reflectance. It implements the multi-frequency fringe projection method to robustly resolve spatially isolated 3D objects with depth discontinuity. Finally, it uses an iterative method for distortion compensation in 3D point recovery to improve the quality of reconstruction. These supports have empowered MS-BLIP for dynamic 3D visualization of translational movements of an engineered box, rotational movements of a craft vase, and full human body movements at an imaging speed of up to 20.8 fps and a measurement volume of up to 1.5 m^3—a three order of

magnitude increase compared to existing BLIP systems. Specifications used for main demonstrations are summarized in Table 5.1.

Table 5.1 Major specifications of MS-BLIP in dynamic 3D imaging experiments

Application	Working distance (m)	(x, y) FOV (m^2)	Projection/ acquisition speed (Hz)	Camera exposure time (ms)	3D imaging speed (fps)	Depth resolution (mm)
Translating engineered box	0.5	0.3 × 0.2	500	2	10.4	1.3
Rotating vase	2.0	1.0 × 1.5	1000	1	20.8	3.7
Full human body	2.8	1.1 × 1.7	500	2	10.4	4.7

MS-BLIP contributes several conceptual advantages to PSFPP. First, its illumination scheme could be readily adapted to other multi-frequency PSFPP systems. Holding a strict imaging relationship between the DMD and the 3D objects, MS-BLIP preserves high accuracy and high contrast of grayscale fringe patterns at different frequencies. Despite being demonstrated only with six selected frequencies, MS-BLIP can be easily adapted to accommodate varying frequencies, arbitrary sets of patterns, and different working distances and FOVs for specific studies. Second, MS-BLIP employs a nanosecond pulsed laser as the illumination source. Since the laser pulse width is much shorter than the DMD's display time, no photons would be wasted during pattern switching, which is otherwise unavoidable for a continuous-wave laser. This advantage also sheds light on implementing MS-BLIP using high-speed cameras (Liang & Wang, 2018; Liu *et al.*, 2021) to increase 3D imaging speeds while maintaining signal-to-noise ratios. Third, a dove prism is used to adjust the orientation of the FOV. In existing PSFPP techniques, the aspect ratio of the projected area is fixed by the DMD's shape. Changing the aspect ratio limits the DMD's active region, which, however, reduces the overall light throughput. The dove prism increases the flexibility in the orientation of the projection region, which thus helps MS-BLIP fit different experimental scenarios.

The future work will be carried out in the following aspects. First, we could further improve MS-BLIP's imaging speed by adopting multiple cameras, a faster DMD, and a more powerful light source. Moreover, we plan to implement online feedback to adaptively adjust the intensity of projected patterns (Chen *et al.*, 2018), which would optimize MS-BLIP's dynamic range for different working conditions. Furthermore, to take into account possible measurement inaccuracy induced by laser speckles, we could implement a superluminescent diode and a rotating diffuser

in the MS-BLIP system as well as apply a filtering algorithm in image reconstruction (Liu *et al.*, 1999). Finally, a graphic processing unit could be used to accelerate the computation of phase maps for real-time display of the 3D reconstruction results. Besides technical improvement, we will continue exploring new applications of MS-BLIP, including automated industrial inspection (So *et al.*, 2013), archaeology (Menna *et al.*, 2011), and human-computer interaction (Zhu *et al.*, 2020).

5.5 Future perspectives

To address safety concerns related to high-powered laser illumination and to broaden BLIP's utility in fields like automated industrial inspection (So et al., 2013), archaeology (Menna et al., 2011), and human-computer interaction (Cui et al., 2012; Zhu et al., 2020), we're transitioning BLIP to use light-emitting diodes (LEDs). As illustrated in Fig. 5.7(a), LED-BLIP features a compact projection module integrating three RGB LEDs with a relatively tight spectral bandwidth (~30nm). Guided by a beam splitter and relay optics, the light is directed to the DMD via a customized total internal reflection (TIR) prism at an angle of ~24°. Subsequently, a band-limited system filters out high-frequency noise and spectral dispersion. The operational procedure will involve triggering the LEDs to sequentially project phase-shifting fringe patterns, which a color camera will capture through spatial multiplexing. After extracting images from distinct color channels, BLIP will reconstruct 3D data based on phase-shifting methods (Flores *et al.*, 2015; Pan *et al.*, 2006). Fig. 5.7(b) and (c) offer visuals of the projection module and its 3D-printed mounting base. The next step will be to merge the remaining BLIP components with the projection module, ensuring that LED-BLIP remains versatile and suited to diverse application scenarios.

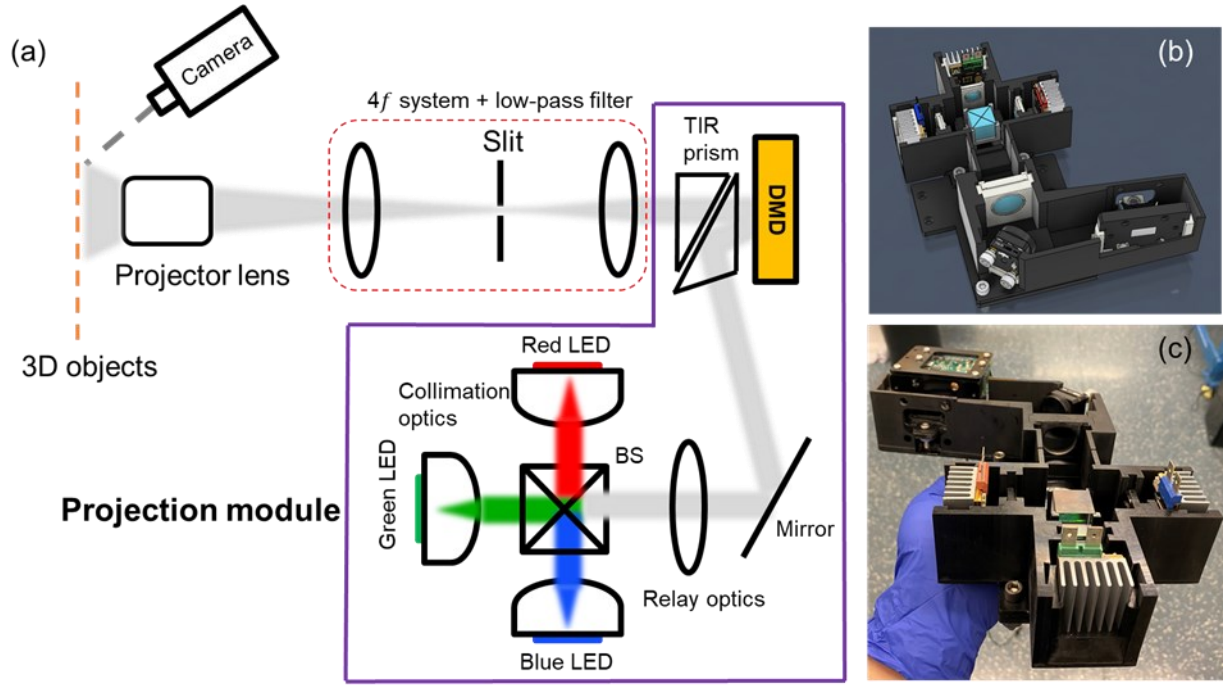

Figure 5.7 System schematic and the construction of of LED-BLIP.

(a) System schematic. BS: beam splitter. TIR prism: total internal reflection prism. (b)-(c) The construction of projection module: 3D modelling of the projection module (b) and the assembled system with 3D printed mounting base.

5.6 References

Alben S & Shelley MJ (2008) Flapping states of a flag in an inviscid fluid: bistability and the transition to chaos. *Physical Review Letters* 100(7):074301.

Bouguet J-Y (*Camera Calibration Toolbox for Matlab.* http://www.vision.caltech.edu/bouguetj/calib_doc/

Burden R, Faires JD & Reynolds A (2010) Numerical analysis, brooks/cole. *Boston, Mass, USA*.

Chang C, Liang J, Hei D, Becker MF, Tang K, Feng Y, Yakimenko V, Pellegrini C & Wu J (2013) High-brightness X-ray free-electron laser with an optical undulator by pulse shaping. *Optics express* 21(26):32013-32018.

Chen C, Gao N, Wang X & Zhang Z (2018) Adaptive pixel-to-pixel projection intensity adjustment for measuring a shiny surface using orthogonal color fringe pattern projection. *Measurement Science and Technology* 29(5):055203.

Cui Y, Chang W, Nöll T & Stricker D (2012) Kinectavatar: fully automatic body capture using a single kinect. *Asian Conference on Computer Vision.* Springer, p 133-147.

Flores JL, Ferrari JA, Torales GG, Legarda-Saenz R & Silva A (2015) Color-fringe pattern profilometry using a generalized phase-shifting algorithm. *Applied Optics* 54(30):8827-8834.

Geng J (2011) Structured-light 3D surface imaging: a tutorial. *Advances in optics and photonics* 3(2):128-160.

Ghiglia DC & Romero LA (1994) Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods. *Journal of the Optical Society of America A* 11(1):107-117.

Gilles Mn, Fayad H, Miglierini P, Clément J-F, Scheib S, Cozzi L, Bert J, Boussion N, Schick U & Pradier O (2016) Patient positioning in radiotherapy based on surface imaging using time of flight cameras. *Medical physics* 43(8Part1):4833-4841.

Hall-Holt O & Rusinkiewicz S (2001) Stripe boundary codes for real-time structured-light range scanning of moving objects. *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001.* IEEE, p 359-366.

Heikkila J & Silvén O (1997) A four-step camera calibration procedure with implicit image correction. *Proceedings of IEEE computer society conference on computer vision and pattern recognition.* IEEE, p 1106-1112.

Huang PS, Zhang S & Chiang F-P (2005) Trapezoidal phase-shifting method for three-dimensional shape measurement. *Optical Engineering* 44(12):123601.

Hyun J-S, Li B & Zhang S (2017) High-speed high-accuracy three-dimensional shape measurement using digital binary defocusing method versus sinusoidal method. *Optical Engineering* 56(7):074102.

Hyun J-S & Zhang S (2017) Superfast 3D absolute shape measurement using five binary patterns. *Optics and Lasers in Engineering* 90:217-224.

Ishii I, Yamamoto K, Doi K & Tsuji T (2007) High-speed 3D image acquisition using coded structured light projection. *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems.* IEEE, p 925-930.

Jha A (2007) Class room attendance system using facial recognition system. *The International Journal of Mathematics, Science, Technology and Management* 2(3):4-7.

Jia P, Kofman J & English CE (2007) Two-step triangular-pattern phase-shifting method for three-dimensional object-shape measurement. *Optical Engineering* 46(8):083201.

Jiang C, Jia S, Xu Y, Bao Q, Dong J & Lian Q (2015) The application of multi-frequency fringe projection profilometry on the measurement of biological tissues. *Bio-Medical Materials and Engineering* 26:S395-S403.

Jiang C, Kilcullen P, Lai Y, Ozaki T & Liang J (2020a) High-speed dual-view band-limited illumination profilometry using temporally interlaced acquisition. *Photon. Res.* 8(11):1808-1817.

Jiang C, Kilcullen P, Liu X, Gribben J, Boate A, Ozaki T & Liang J (2020b) Real-time high-speed three-dimensional surface imaging using band-limited illumination profilometry with a CoaXPress interface. *Optics Letters* 45(4):964-967.

Jiang C, Kilcullen P, Liu X, Lai Y, Ozaki T & Liang J (2021) High-speed three-dimensional surface measurement using band-limited illumination profilometry (BLIP). *Emerging Digital Micromirror Device Based Systems and Applications XIII.* International Society for Optics and Photonics, p 116980V.

Kosmopoulos D & Varvarigou T (2001) Automated inspection of gaps on the automobile production line through stereo vision and specular reflection. *Computers in Industry* 46(1):49-63.

Laser Institute of America O, Florida (2014) ANSI Z136.1-2014 Laser Safety Standard for the Safe Use of Lasers.

Lau DL & Arce GR (2018) *Modern digital halftoning.* CRC Press,

Li B, Wang Y, Dai J, Lohry W & Zhang S (2014) Some recent advances on superfast 3D shape measurement with digital binary defocusing techniques. *Optics and Lasers in Engineering* 54:236-246.

Li B & Zhang S (2017) Superfast high-resolution absolute 3D recovery of a stabilized flapping flight process. *Optics Express* 25(22):27270-27282.

Li X, Gong M, Zhang Z, Zhang Z & Yang C (2017) Calibration method for fringe projection profilometry with a binary defocusing technique. *Optical Review* 24(4):495-504.

Liang J (2020) Punching holes in light: recent progress in single-shot coded-aperture optical imaging. *Reports on Progress in Physics* 83(11):116101.

Liang J & Wang LV (2018) Single-shot ultrafast optical imaging. *Optica* 5(9):1113-1127.

Liang J, Wu S-Y, Kohn RN, Becker MF & Heinzen DJ (2012) Grayscale laser image formation using a programmable binary mask. *Optical Engineering* 51(10):108201.

Liang J, Zhou Y, Winkler AW, Wang L, Maslov KI, Li C & Wang LV (2013) Random-access optical-resolution photoacoustic microscopy using a digital micromirror device. *Optics letters* 38(15):2683-2686.

Liu H, Lu G, Wu S, Yin S & Francis T (1999) Speckle-induced phase error in laser-based phase-shifting projected fringe profilometry. *JOSA A* 16(6):1484-1495.

Liu X, Skripka A, Lai Y, Jiang C, Liu J, Vetrone F & Liang J (2021) Fast wide-field upconversion luminescence lifetime thermometry enabled by single-shot compressed ultrahigh-speed imaging. *Nature Communications* 12(1):6401.

Mazzetto M, Southier LF, Teixeira M & Casanova D (2019) Automatic classification of multiple objects in automotive assembly line. *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA).* IEEE, p 363-369.

Menna F, Nocerino E & Scamardella A (2011) Reverse engineering and 3D modelling for digital documentation of maritime heritage. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 38(5):W16.

Pan J, Huang PS & Chiang F-P (2006) Color phase-shifting technique for three-dimensional shape measurement. *Optical Engineering* 45(1):013602-013602-013609.

Parthiban V, Kohn Jr RN, Liang J & Becker MF (2016) Experimental demonstration of precise holograms using complex light modulation. *Emerging Digital Micromirror Device Based Systems and Applications VIII.* International Society for Optics and Photonics, p 97610M.

Qian J, Feng S, Tao T, Hu Y, Liu K, Wu S, Chen Q & Zuo C (2019) High-resolution real-time 360° 3D model reconstruction of a handheld object with fringe projection profilometry. *Optics letters* 44(23):5751-5754.

Saldner HO & Huntley JM (1997) Temporal phase unwrapping: application to surface profiling of discontinuous objects. *Applied optics* 36(13):2770-2775.

So EWY, Munaro M, Michieletto S, Antonello M & Menegatti E (2013) Real-time 3d model reconstruction with a dual-laser triangulation system for assembly line completeness inspection. *Intelligent Autonomous Systems 12,* Springer. p 707-716.

Song Z (2013) *Handbook of 3D machine vision : optical metrology and imaging.*

Tian J, Peng X & Zhao X (2008) A generalized temporal phase unwrapping algorithm for three-dimensional profilometry. *Optics and Lasers in Engineering* 46(4):336-342.

Wang Y, Zhao H, Jiang H & Li X (2018) Defocusing parameter selection strategies based on PSF measurement for square-binary defocusing fringe projection profilometry. *Optics express* 26(16):20351-20367.

Wu Z, Guo W, Pan B, Kemao Q & Zhang Q (2021) A DIC-assisted fringe projection profilometry for high-speed 3D shape, displacement and deformation measurement of textured surfaces. *Optics and Lasers in Engineering* 142:106614.

Yu W & Xu B (2010) A portable stereo vision system for whole body surface imaging. *Image and vision computing* 28(4):605-613.

Zhang S (2010) Recent progresses on real-time 3D shape measurement using digital fringe projection techniques. *Optics and lasers in engineering* 48(2):149-158.

Zhang S (2018) High-speed 3D shape measurement with structured light methods: A review. *Optics and Lasers in Engineering* 106:119-131.

Zhang S & Huang PS (2006) Novel method for structured light system calibration. *Optical Engineering* 45(8):083601.

Zhang S & Yau S-T (2009) High dynamic range scanning technique. *Optical Engineering* 48(3):033604.

Zhong K, Li Z, Zhou X, Li Y, Shi Y & Wang C (2015) Enhanced phase measurement profilometry for industrial 3D inspection automation. *The International Journal of Advanced Manufacturing Technology* 76(9-12):1563-1574.

Zhu L, Rematas K, Curless B, Seitz SM & Kemelmacher-Shlizerman I (2020) Reconstructing nba players. *European Conference on Computer Vision.* Springer, p 177-194.

Zuo C, Huang L, Zhang M, Chen Q & Asundi A (2016) Temporal phase unwrapping algorithms for fringe projection profilometry: A comparative review. *Optics and lasers in engineering* 85:84-103.

Zuo C, Tao T, Feng S, Huang L, Asundi A & Chen Q (2018) Micro Fourier Transform Profilometry (μFTP): 3D shape measurement at 10,000 frames per second. *Optics and Lasers in Engineering* 102:70-91.

6 SPIRIT OF LWIR IMAGING THERMOMETRY WITH PASSIVE DETECTION

Abstract

Efficient and reliable measurement of human body temperature has gained paramount importance, particularly in the wake of the COVID-19 pandemic. Conventional temperature measurement methods grapple with limitations due to their invasive or semi-invasive nature. Although recent advancements, like 2D thermography, aim to surmount these limitations, they are hampered by inefficient utilization of pixel resources in pinpointing body temperature indicators—specifically, the inner canthi—and encounter inaccuracies due to background blending. Single-pixel imaging (SPI) employs coded wide-field illumination and samples via a single-pixel detector, ensuring targeted area focus and diminishing background blending impacts. Nevertheless, SPI's application to human body temperature measurement has been impeded due to spectral ranges that exceed the range of optimized reflectivity of coatings of spatial encoders like digital micromirror devices (DMDs). In response to these challenges, we develop Single-Pixel infrared imaging thermometry (SPIRIT) with passive detection. Utilizing cyclic S-matrices, together with diagonal aggregation to generate masking patterns fabricated onto a physical mask, SPIRIT enables the first-time thermal imaging of the inner canthi by SPI with a frame size of 11 × 13 pixels. With statistical analysis, SPIRIT has performed temperature screenning of human subjects and investigated body temperature fluctuations and the impact of prolonged eyeglass usage.

6.1 Introduction

The COVID-19 pandemic has affected all over the world, amassing approximately 120 million confirmed cases and resulting in more than 2.5 million fatalities since its inception (Perpetuini *et al.*, 2021). Subsequent waves linked to the Omicron variant underscore the significant influence new strains can exert on global health (Martín Sánchez *et al.*, 2023; Reiner *et al.*, 2023). Early diagnosis of COVID-19 positive cases is essential to keep the infected person isolated and lower the risk of spreading the virus to others in the community. Clinical presentation of the disease ranges from no symptoms to severe pneumonia, in which the most commonly reported clinical symptom is fever (Team, 2020). Consequently, a reliable and efficient method to measure body temperature is of paramount importance.

Existing temperature sensing techniques can be generally grouped into three categories based on the level of contact between the medium of interest and the measurement device (Childs, 2001; Childs *et al.*, 2000). The first category is the direct-contact techniques, such as thermocouples (Bradley & Matthews, 1968; Heitor & Moreira, 1993; Kasap, 2001) and thermistors (Hughes-Riley *et al.*, 2017; Mosbach & Danielsson, 1974). They are crafted leveraging thermoelectricity and the temperature-dependent resistance variance of electrical conductors. Despite their utility, these approaches present multiple challenges, including entangled wires during movement, compromised subject comfort, sluggish responses, and dislodged skin sensors due to perspiration in prolonged heat exposure (Buono *et al.*, 2007). The second category contains minimally-invasive methods. As an example, thermochromic liquid crystals are applied to the target surface as a chemical indicator. These crystals undergo color alterations in response to temperature fluctuations, which can be monitored by a visible detector (Abdullah *et al.*, 2010; Liu *et al.*, 2021; Sage, 2011; Smith *et al.*, 2001). However, the necessity of spraying the chemical indicator poses inconvenience, is time-intensive, and raises potential issues in clinical settings (Al Raisi *et al.*, 2020).

The last category—non-contact techniques—offer label-free and non-invasive advantages (Ng *et al.*, 2005; Saegusa & Tabata, 2003). For example, the infrared thermometers (often referred to as “forehead guns”), collect and analyze the radiation emitted by an object for temperature calculation based on Stefan-Boltzmann law, are widely used for non-contact temperature measurement (Diwanji *et al.*, 2020; Sollai *et al.*, 2016; Zhang, 2018). A laser is typically used to help the thermometerts to identify the targeting area (Furstenberg *et al.*, 2008). In addition, photoacoustic techniques are used for temperature measurement in biomedical tissues (Pramanik & Wang, 2009; Shah *et al.*, 2008; Wang *et al.*, 2009). When pulsed or

modulated light is absorbed by a tissue or material, the amplitude and velocity of the acoustic signals can be correlated to the temperature of the medium. Furthermore, the temperature dependence of Raman shifts and Raman linewidths can be exploited for temperature measurements using coherent Anti-Stokes Raman spectroscopy (CARS) (Pealat *et al.*, 1985; Stricker *et al.*, 2003). Nonetheless, these methodologies necessitate active illumination, thereby posing safety concerns.

To eliminate this limitation, imaging thermography has been developed to monitor body temperature in a passive mode (Lahiri *et al.*, 2012; Meola & Carlomagno, 2004; Turner, 2001). This approach uses a radiometer to transform the thermal radiation from objects into an electronic video signal (Carlomagno & Cardone, 2010; Meola *et al.*, 2004). Though it sidesteps the need for active illumination, which avoids the safety concerns and mitigates the requirement for extra excitation sources, this approach is confronted with challenges in practicality and measurement accuracy. First, the equipment necessary for 2D thermography, such as infrared cameras, has limited availability or is prohibitively expensive (Gade & Moeslund, 2014). Second, imaging thermography is often configured to image the entire human face, which does not leverage its expansive sensor to concentrate on the key areas indicative of core body temperature, specifically the inner canthi (Ng *et al.*, 2004; Ring & Ammer, 2015; Teunissen & Daanen, 2011b). In addition, given the extensive FOV, the pixel count allocated to smaller, critical regions like the inner canthi is comparatively low. Consequently, the background can blend with the area of interest. Considering the temperature-signal correlation, any perturbation in the signal level inevitably compromises the accuracy of temperature readings (Avdelidis & Moropoulou, 2004; Rajic, 2002; Usamentiaga *et al.*, 2014).

Single-pixel imaging (SPI) with passive detection emerges as a cost-effective solution compared to conventional thermographic methods. This technique encodes deterministic spatial patterns onto the imaging beam, with subsequent data acquisition performed by a single-pixel detector (Gibson *et al.*, 2020). During the subsequent process of image reconstruction, these encoding patterns are used as prior knowledge, aiding in the accurate retrieval of spatial information (Edgar *et al.*, 2015; Gibson *et al.*, 2020; Sun *et al.*, 2016). By obviating the necessity for a 2D sensor, such as CCD or CMOS, SPI leverages the advantages of single-pixel detectors in their wider availability in the sensing spectrum and lower noise levels to fill the void where the 2D-sensor-based counterparts are not available or impractical (Duarte *et al.*, 2008; Kilcullen *et al.*, 2022). This niche has propelled SPI in non-visible spectral ranges, including short-wavelength infrared (Edgar *et al.*, 2015; Gibson *et al.*, 2017; Tao *et al.*, 2022), mid-wavelength infrared (Dong

et al., 2023; Ebner *et al.*, 2023; Wang *et al.*, 2023), terahertz (Stantchev *et al.*, 2016; Stantchev *et al.*, 2020), and photoacoustic (Liang *et al.*, 2014; Torke *et al.*, 2022) bands. Despite these attractive advantages, to date, SPI has not yet been demonstrated for human body thermometry. A pivotal challenge lies in the functionality of digital micromirror devices (DMDs), which are conventionally employed as signal encoders in SPI. These devices do not operate effectively in the long-wavelength infrared (LWIR) range due to exceeding the range of transmittance window of the protection glass and the optimized reflectivity of coatings, a spectrum that aligns with the blackbody radiation emitted by objects at typical human body temperatures (Vatansever & Hamblin, 2012). This limitation of DMDs in the LWIR spectrum underscores the imperative to investigate alternative modulation technologies and strategies. Such exploration is crucial to unlock the potential of SPI in temperature measurement and infrared thermography.

To surmount these limitations, we develop single-pixel imaging infrared thermometry (SPIRIT). Diagonally aggregated masking patterns constructed from cyclic S-matrices are used to generate 2D encoding masks with a compressed-sensing-compatible linear scan. The incident thermal radiation in the LWIR spectral range is encoded by these masks and then integrated by a cooled HgCdTe single-pixel detector. Meanwhile, to enhance the temperature sensing accuracy, SPIRIT separates its FOV to target the left and right human inner canthi. Through the integration of image reconstruction and temperature calibration, SPIRIT is empowered to deliver precise thermometry on the human inner canthi.

6.2 Method

6.2.1 Cyclic S-matrices

SPIRIT uses a cyclic S-matrix of order n=143 as the basis for its encoding scheme. An S-matrix of order n can be defined as any matrix within the class of $\{0,1\}$-valued $n \times n$ matrices that maximizes the determinant(Cohn, 1963; Sloane & Harwit, 1976). It can be shown that in order for an S-matrix of order *n* to exist, n must be either equal to 1, or of the form $4k-1$ for some positive integer k. Moreover, it can be shown that for any S-matrix S of order n,

$$S^{-1} = \frac{2}{n+1}(2S^T - J), \tag{6.1}$$

where J denotes the all-ones matrix with size $n \times n$, and S^T denotes the matrix transpose of S.

Cyclic S-matrices possess an additional cyclic structure whereby the initial row determines all subsequent rows via left-wise circular shifts. In particular, letting $S_{i,j}$ denote the element of a cyclic S-matrix S with row index i and column index j, the elements of S satisfy

$$S_{i,j} = S_{0,\,i+j}\,, \tag{6.2}$$

where $i, j = 0, \ldots, n-1,$ and $i + j$ is interpreted modulo n. Various methods are known for the construction of cyclic S-matrices. For imaging, it is desirable that n be factorizable into parts of approximately equal size. For the design of the masks in SPIRIT, the 'twin-prime construction method' was used (Kilcullen *et al.*, 2022; Sloane & Harwit, 1979; Stanton & Sprott, 1958). Given any pair of twin primes p and $q = p + 2$, it is possible to construct a cyclic S-matrix of order $n = pq$ as follows. An integer x is called a quadratic residue (mod y) if $x \neq 0$ (mod y) and there exists another integer z such that $x \equiv z^2$ (mod y). The following functions can then be defined:

$$f(j) = \begin{cases} +1 & \text{if } j \text{ is a quadratic residue (mod } p) \\ 0 & \text{if } j \equiv 0 \pmod{p} \\ -1 & \text{otherwise} \end{cases} \quad and \tag{6.3}$$

$$g(j) = \begin{cases} +1 & \text{if } j \text{ is a quadratic residue (mod } q) \\ 0 & \text{if } j \equiv 0 \pmod{q} \\ -1 & \text{otherwise} \end{cases}\,, \tag{6.4}$$

from which the elements of S are computed by

$$S_{i,j} = \begin{cases} 0 & \text{if } [f(i+j) - g(i+j)]\,g(i+j) = 0 \\ +1 & \text{otherwise} \end{cases}\,, \tag{6.5}$$

where $i, j = 0, \ldots, pq - 1$.

6.2.2 Data interpolation and image reconstruction

The image reconstruction procedure of SPIRIT leverages the fact that multiplication of a vector by a cyclic S-matrix of appropriate order can be viewed as an operation of two-dimensional discrete convolution (Gopalsami *et al.*, 2012; Kilcullen *et al.*, 2022). In particular, if the order n of S can be factored as $n = pq$ and the rows of S are deployed as $p \times q$ arrays via row-major ordering, the following equivalence of indexed sums may be observed for expressing the matrix-vector product of an underlying image x with S:

$$y_u = \sum_{v=0}^{n-1} S_{u,v}\, x_v = \sum_{k=0}^{p-1} \sum_{l=0}^{q-1} s(i+k, j+l)\, x(k,l) = y(i,j)\,, \tag{6.6}$$

where we have used the notation $x(k,l)$ as equivalent to x_{qi+j} for each vector (i.e. showing the conversion between 2D and 1D indexing schemes via row-major ordering), with all indices interpreted as modulo each vector's length, and abbreviated the initial row of S with the vector s via $s_i = S_{0,i}$. The salient feature of Eq. (9) is that, since both horizontal and vertical index shifts of ± 1 result in a high degree of similar spatial overlap between encoding patterns and the underlying image, it exposes how bucket signals collected from SPI carried out with cyclic S-matrices can be expected to exhibit 2D smoothness when organized according to row-major ordering. In terms of linear indexing, this implies that for any individual bucket signal y_i, a high degree of correlation should be expected between values such as of $y_{i\pm 1}$ and $y_{i\pm b}$ that closely neighbour y_i after 2D reshaping.

This expected 2D smoothness gives rise to a redundancy that was exploited to achieve a reduction in the size of physical encoding masks used in SPIRIT (Gopalsami *et al.*, 2012; Kilcullen *et al.*, 2022). The physical masks were designed in such a way that a linear scan produced the bucket signals $y(0,0)$, $y(1,1)$, $y(2,2)$, …, $y(m,m)$. This choice allows for the data collected by a continuous linear scan to acquire the elements of y in a diagonal fashion within a 2D array, while also exhibiting the convenient property that a premature stop to data collection (i.e. with $m < n$) would result in measured elements of y distributed in an approximately uniform manner well-suited for the estimation of non-measured elements via 2D interpolation. More details about the design procedure used to select the desired matrix order n and scan size m are provided in Supplementary Note 4.

For the estimation of non-measured data, the method of discrete Laplace interpolation (al., 2007)was used that was capable of accommodating irregularly distributed data points in the 2D array of bucket signal data. Given a set of measured data $\{y_i | i \in K\}$ labeled with linear indices $i \subset \{0, \dots, n-1\}$ as well as a set of indices $U = \{0, \dots, n-1\}/K$ corresponding to the linear indices of non-measured encoding patterns in an experiment, a system of $|U|$ equations in $|U|$ unknowns may be formed as follows:

$$y_i = \frac{1}{4}[y_{i+1} + y_{i-1} + y_{i+b} + y_{i-b}] \ , \tag{6.7}$$

for each index $i \in U$, and where each index is interpreted modulo n. Solving this system of equations recovers estimates of the non-measured values $\{y_i | i \in U\}$ allowing for a complete n-element interpolated data vector $\tilde{y}$ to be constructed as a combination of all bucket signal values both measured and estimated. From the full interpolated data vector $\tilde{y}$, a final recovered image $\tilde{x}$ is obtained via inversion of the full-rank measurement matrix:

$$\tilde{x} = S^{-1}\tilde{y} \ , \tag{6.8}$$

Finally, the vector x may be reshaped according to both row-major ordering, as well as separation and rotation in order to produce imagery concurring with the two FOVs defined by the mask window.

6.3 Results

6.3.1 System setup

A schematic of the SPIRIT system is depicted in Fig. 6.1a. Thermal radiation emitted from the human face is first imaged by a ZnSe lens L1 (39-532, Edmund Optics) onto the intermediate image plane with a 4:1 magnification ratio. As shown in Fig. 6.1b, a mask plate and a window plate are placed on the intermediate image plane. The mask plate, attached to a motorized translation stage (MTS25-Z8, Thorlabs) operating in stepper mode, contains two encoded stripes. Given the mean value of inner canthal distance is ~30 mm (Hayat *et al.*, 2019), the two encoded sripes are separated with a distance of 7.5 mm. These stripes contain aggregated masking patterns, with 11 × 7 and 11 × 6 encoding pixels (250 μm × 250 μm in size) in an individual pattern, respectively. Depicted in Fig. 6.1c, the window plate contains two open areas of 2.75 mm × 1.75 mm and 2.75 mm × 1.50 mm in size, which are rotated -45° and +45°, respectively. They have a center-to-center distance same as the two encoded stripes. As delineated in Fig. 6.1d, the window crops individual masking patterns from the mask, while the full encoding is performed through vertical movements of the mask. Subsequently, the encoded thermal radiation navigating through two open areas is collected by a ZnSe camera lens L2 (307B/1.0, Yoseen Infrared), passing through an optical chopper (MC2000B, Thorlabs), and finally focused onto an HgCdTe amplified single-pixel detector (PDAVJ10, Thorlabs) cooled by a Peltier thermoelectric cooler. In this configuration, the field of view on the human face contains two regions of 11 mm × 7 mm (-45°-rotated) and 11 mm × 6 mm (-45°-rotated) with a distance of 30 mm, which covers both inner canthus of most human (Hayat *et al.*, 2019). Meanwhile, each encoding pixel samples a 1 mm × 1 mm region in the targeted inner canthus. This configuration aligns with the standards set by IEC80601-2-59 (2023) (Anonyme, 2023), advocating a minimum of 1 pixel per mm for optimal febrile temperature screening via thermographs.

In operation, after system calibration, a color camera (16-544-RCD-05P, Edmund Optics) is used to coincide the inner canthi of a human subject with the FOV of the SPIRIT system. The optical chopper provides a modulation of 850 Hz. The acquired signals are sent to a lock-in

amplifier (SR830, Stanford Research) to enhance the signal-to-noise ratio (SNR) before being stored in a digitizer (ATS9625, Alazartech). The data acquisition is synchronized with the vertical scan of the encoding mask. The amassed data points, termed “bucket signals,” are then relayed to a computer for subsequent image reconstruction. Each scan, collecting 95 bucket signals (detailed in the next section), takes 15 seconds.

The full details of the mask-window alignment and the system synchronization are discussed in Supplementary Notes 1–2. A discussion of the performance comparison with different attempts at SNR enhancement is provided in Supplementary Note 3.

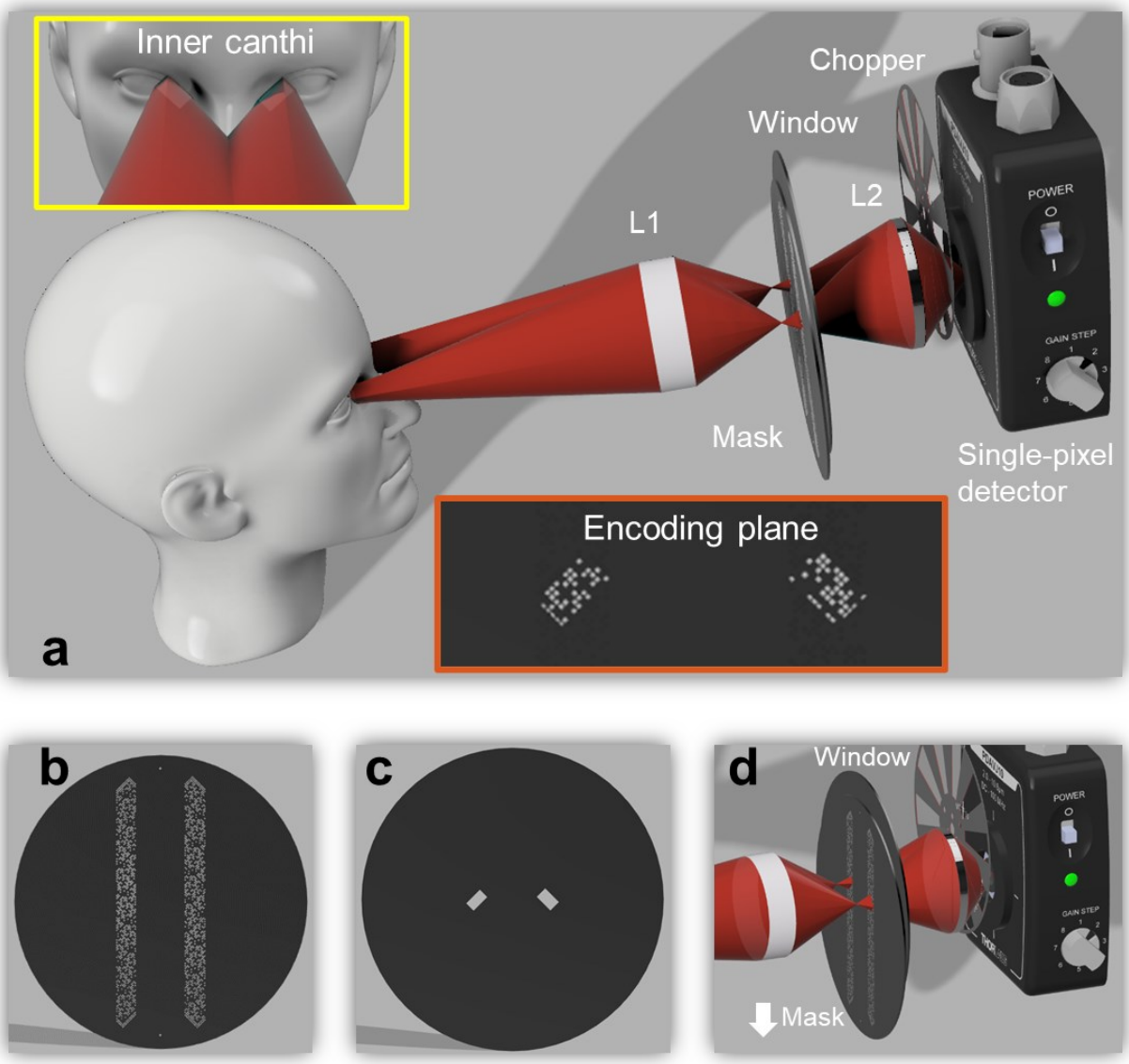

Figure 6.1 Schematic of SPIRIT.

a System illustration showing the encoding and collection of the thermal radiation of the human subject. L1-L2, lenses. Closeup images show the targeted FOVs of inner canthi (red-boxed) and the active pixels at the encoding plane (orange-boxed). Red cylinders: thermal radiations from the targeted FOVs. b - d Closeups of the mask(b), window (c), and the working manner of encoding (d). The white arrow in (d) indicates the scsan direction of the mask.

6.3.2 Coding strategy

The bucket signals, represented by an m-element vector y, can be interpreted as the inner products computed optically between an image comprising n pixels, represented by x, and the collection of encoding patterns, denoted as $\{s_i, i = 1,\dots,m\}$. This procedure can be succinctly articulated as an equation

$$y = Sx, \tag{6.9}$$

where the measurement matrix *S* incorporates each encoding pattern, represented in row form, and aligns with the sequential order of the bucket signals designated in y.

SPIRIT employs cyclic S-matrices to derive encoding patterns, a choice driven by several advantages. Firstly, S-matrices inherently optimize multiplex advantage for {0, 1}-valued encoding designs (Wehlburg *et al.*, 2001; Yue *et al.*, 2018), coupled with the benefit of possessing easily computable analytical inverses. In addition, the inherent cyclic property facilitates the amalgamation of multiple masking patterns into adjacent positions, rendering them apt for efficient placement on a physical substrate. Furthermore, this property also fosters highly efficient data processing, as the storage of merely the initial encoding is ample for computing both the set of encoding patterns and the matrix transform employed in image reconstruction.

Specifically, SPIRIT employs a 'twin-prime construction method' to generate cyclic S-matrices, detailed further in the section Supplementary Note 4. As depicted in Fig. 6.2a, SPIRIT operationalizes cyclic S-matrices with twin-prime factors of $p = 11, q = 13,$ yielding a resultant $n = pq = 143$. Individual encoding pattern, featuring dimensions of $p \times q$, emerges from the 2D transformation of the rows of S using row-major sequencing. Leveraging the 2D reshaping of the rows enables the transformation of S to construct a $(2p-1) \times (2q-1)$ pattern that precisely encapsulates all 2D encoding patterns as determined by its rows. Fig. 6.2b illustrates an instance of such restructuring, highlighting that each $p \times q$ sub-region exemplifies a 2D reshaped row of S, with the corresponding row index distinctly marked in the top-leftmost corner element.

Moreover, the diagonal aggregation method enables the cropped $p \times q$ sub-regions to execute a linear scan compatible with compressive sensing principles. Consequently, as illustrated in Fig. 6.2c, this methodology accomplishes a uniform fill level of 63.6% by employing

$m = 91$ patterns. In Fig. 6.2d, a separation and rearrangement of the encoding pattern is depicted for imaging the inner canthi of human subjects. In each scan, a $p \times q$ sub-region of the restructured matrix pattern is initially segmented into two distinct regions, with dimensions of $p \times [(q+1)/2]$ and $p \times [(q-1)/2]$. Subsequently, to accommodate variations in the inner canthal distance among different individuals, each of the two separated regions undergoes a 45º rotation.

The diagonal aggregation of the $p \times q$ sub-region then enables a sequential pack of the encoding patterns in a compact way to form two encoding paths, as illustrated in the zoom-in view of Fig. 6.2d. In addition to the encoding patterns, four runway patterns are designed at the commencement and conclusion of the encoding path, serving the purpose of verifying the alignment between the mask and the window. Subsequently, as delineated in the blue-boxed region of Fig. 6.2d, each encoding pattern is cropped and operated by a vertically scannable window.

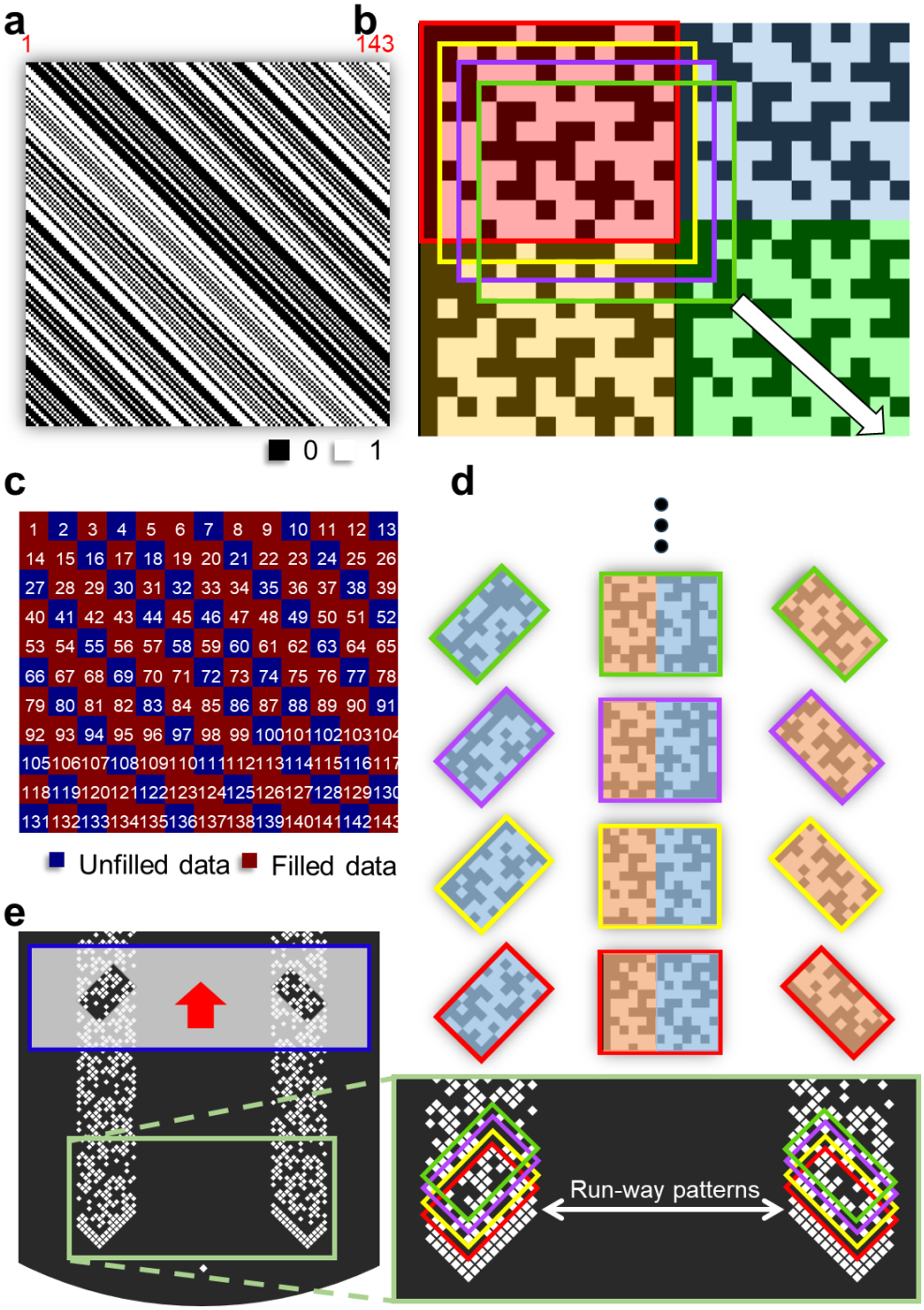

Figure 6.2 Coding strategy of SPIRIT.

a Binary cyclic S-matrix of order n= 143 (i.e., $p = 11$ and $q = 13$). b Restructured matrix patterns formed from the matrix shown in (a). Shaded colors highlighting four tiled copies of the reshaped initial row of (a). The color-boxed regions show a section of individual masking patterns used to generate the mask, with each region corresponding to a distinct row in (a) (Red: Row 1; Yellow: Row 15; Purple: Row 29; Green: Row 43). White arrow indicates the diagonal aggregation direction. c The encoded signal matrix collected by SPIRIT with a compressed-sensing-compatible linear scanning. d Separation and rearrangement of specific encoding patterns. The color-boxed regions (also appearing in (b)) highlight the role of a single encoding pattern. The

rotated left and right boxes with shaded colors show the two distinct encoding regions. e Mask with 91 encoding patterns based on the arrangement in (d) and four run-way patterns (two on the top and two on the bottom). The blue-boxed region indicates the window which meticulously crops the mask, ensuring only a specific encoding pattern is operational within the system. The red arrow indicates the direction of vertical scanning of the window to perform sequential encoding. A zoomed-in view offers detailed insights into the mask's starting region, with colored boxes underscoring the presence of the encoding patterns previously observed in (b) and (d).

6.3.3 Proof-of-concept demonstration and calibration of SPIRIT

To prove the concept of SPIRIT, we executed imaging of structured features, emphasizing thermal contrast. The experimental procedures were undertaken indoors, maintaining an ambient temperature and relative humidity of approximately 22°C and 40% respectively, thereby complying with the environmental considerations stipulated in IEC80601-2-59 (2023) (Anonyme, 2023) for infrared inspection endeavors. To further refine the testing environment, curtains were deployed to encircle the experimental setup, reducing airflow from ventilation ducts and thereby minimizing any forced cooling or heating of the target. Additionally, we applied a spray to the surfaces of system elements with high thermal radiation reflectivity, aiming to decrease reflectivity and thus mitigate interference from additional infrared radiation sources surrounding the experimental setup, such as incandescent and halogen lamps.

A plane black body radiation source (YSHT-35, Yoseen Infrared) served as a uniform background target within SPIRIT's FOVs, maintaining a stable temperature of 37°C. To establish thermal contrast with structured features, depicted in Fig. 6.3a, plastic plates possessing low thermal conductivities were placed to partially obstruct certain window plates. The relative positions between these plastic plates and the window, highlighted by yellow dashed-line boxes, illustrate the obstruction and allowance of thermal radiation. The bottom row showcases reconstructed images, while the averaged line profiles are presented in Fig. 6.3b, revealing that SPIRIT successfully observes a high-quality thermal contrast.

Furthermore, as depicted in Fig. 6.3c, plastic bars, each 2 mm in width, were utilized to generate structured features, with the reconstructed images displaying notable correspondence. According to the 4:1 imaging condition through L1, each plastic bar should be sampled with approximately 2 pixels in the width direction. This is aptly mirrored in the reconstruction images, substantiating the capability of SPIRIT to robustly image fine features with thermal contrast.

After validating the imaging capacity of SPIRIT, we performed a temperature calibration to establish a correlation between reconstructed intensity and temperature. The black body radiation source was deployed to generate temperatures ranging from 32°C to 38°C, incremented by 0.3°C intervals. For each temperature point, SPIRIT acquired and reconstructed five groups to

mitigate potential variations and instabilities from the black body radiation source over time. Subsequently, for each set of reconstructed images, the intensity profile was averaged across a 13 × 11-pixel grid, and the standard deviation was calculated.

In accordance with the Stefan-Boltzmann law (Wellons, 2007), which asserts that the thermal radiated power P of an object at temperature T can be expressed as

$$P = A\varepsilon\sigma T^4, \tag{6.10}$$

where ε is the emissivity of the object, σ denotes the Stefan-Boltzmann constant, valued at $5.67 \times 10^{-8}\, Wm^{-2}K^{-4}$, and A indicates the size of surface area. Given the black body radiation source and human body emissivity, ε is set at 0.98 (Charlton *et al.*, 2020).

Meanwhile, considering the light-to-current conversion of the single-pixel detector, the signal intensity of the final recovered image $\tilde{x}$, calculated in Volts, is given by

$$\tilde{x} = CG \times P = C\varepsilon\sigma T^4, \tag{6.11}$$

where CG is the conversion gain from Watt to Volt, and C is the constant to be calibrated. The averaged pixel intensity profiles, reconstructed at varying temperatures, are linearly fitted based on Eq. (6.11) to extract constant C.

Considering system noise, the measured temperature can be described as

$$\tilde{x} = C\varepsilon\sigma T_{\text{measured}}^4 = C\varepsilon\sigma T_{\text{set}}^4 + \Delta, \tag{6.12}$$

where T_{measured} and T_{set} are the SPIRIT measured temperature and the pre-set temperature of the black body radiation source, respectively. And Δ represents the system noise. Thus, the measured temperature is related to the pre-set temperature as

$$T_{\text{measured}} = \sqrt[4]{\frac{C\varepsilon\sigma T_{\text{set}}^4 + \Delta}{C}}. \tag{6.13}$$

Accordingly, Fig. 6.3d delineated the measured temperature T_{measured} aligns well with the set temperature T_{set}, which indicates SPIRIT temperature sensitivity maintains an accuracy of ±0.3°C.

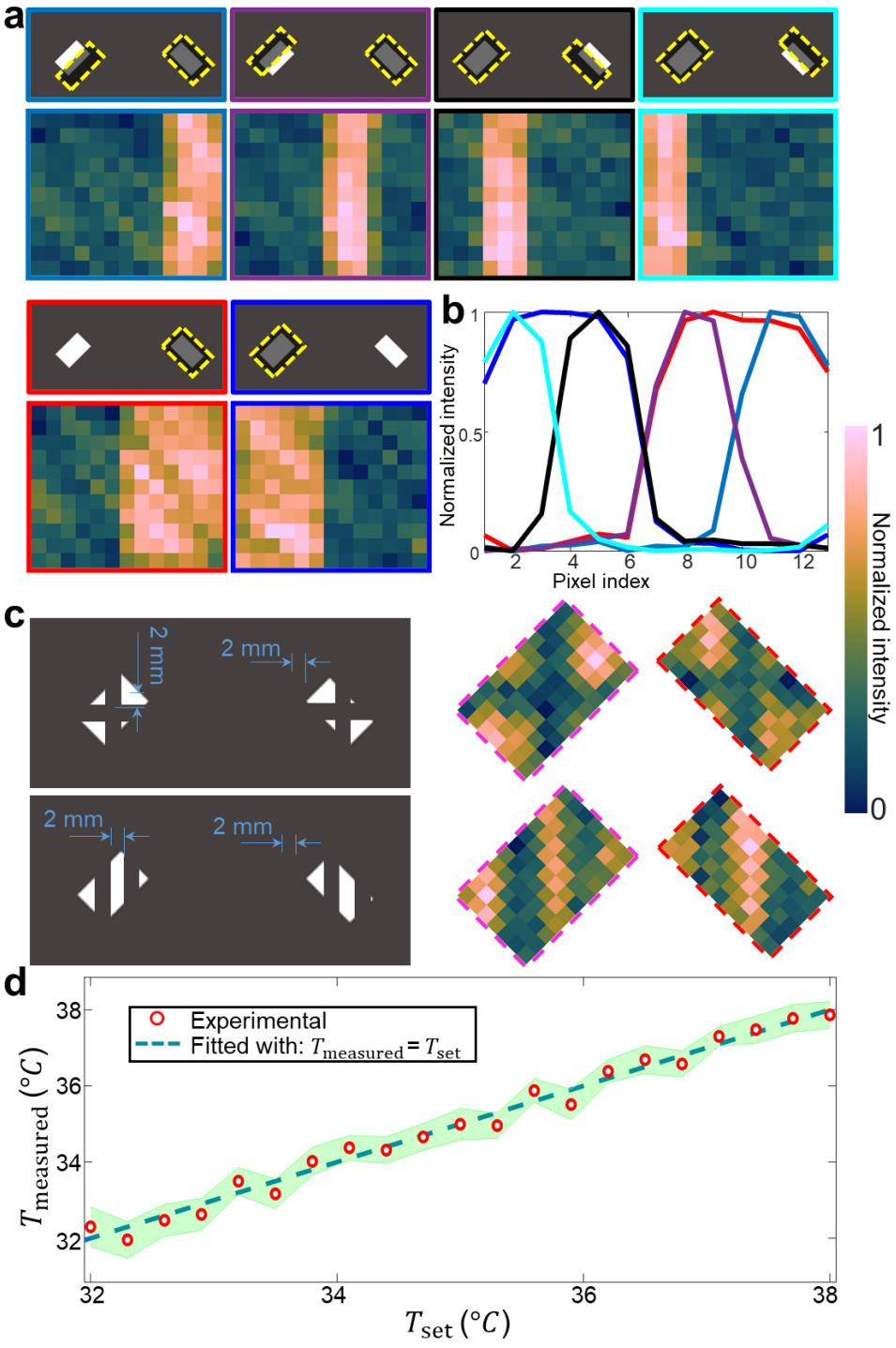

Figure 6.3 Proof-of-concept demonstration and temperature calibration of SPIRIT.

a Experimental configuration and reconstructed images of SPIRIT with a partially obstructed window: plastic plate arrangements (top) and corresponding reconstructed images (bottom) are displayed. The yellow dashed-line box highlights the plastic plates. b Averaged line profiles corresponding to (a). Colors indicate the correlation to different reconstructions in (a). c Experimental arrangement and reconstructed images of SPIRIT utilizing plastic bars: presentations of the plastic bar arrangements (left) and reconstruction images (right). d

Temperature calibration of SPIRIT with a linearly fitting as $T_{\text{measured}} = T_{\text{set}}$. The green-area represents the standard deviation.

6.3.4 Temperature screening using SPIRIT

To demonstrate SPIRIT of the human inner canthi, we recruited 14 healthy volunteers with no fever symptoms. The experimental condition remains consistent with the environmental settings outlined in the Section - Proof-of-Concept demonstration and calibration of SPIRIT. Participants were instructed to remove their eyeglasses and clear any eye makeup, which might alter the thermal properties of the inner canthi. All procedures, were executed in strict adherence to the protocol (Project CÉR-20-578), as approved by the Human Ethics Research Committee at the Institut National de la Recherche Scientifique, Université du Québec.

Initially, the FOVs of SPIRIT were registered with the assistance of a visible camera. A full-face human mask was employed as a simulated subject, accompanied by two Peltier coolers. The hot sides of the coolers were utilized to generate thermal contrast, aiding in identifying the FOVs through the reconstruction results (additional details can be found in Supplementary Note 5).

After identifying the FOVs, two markers were placed on the control PC screen, serving as guides to assist volunteers in precisely positioning their heads on a chinrest. Despite varying structures of the inner canthi and differences in inner canthal distances, SPIRIT's deployment of two encoding masks adeptly accommodates the targeted inner canthi. Furthermore, volunteers were advised to close their eyes for a brief rest prior to testing, and to maintain their eyes open during the approximately 15-second data acquisition period.

Figure 6.4a shows the system setup of SPIRIT for human subjects. A pre-calibrated thermal imaging camera (Yoseen M384D) was deployed to monitor the faces of human subjects during SPI measurements. Specifically, the average temperatures of the left and right inner canthi were ascertained as the ground truth.

As displayed in Fig. 6.4b, the eye contours of two representative volunteers, extracted from images taken by a visible camera, are aligned with the SPIRIT reconstructions at the inner canthi regions. The average temperatures within the inner canthi were calculated and recorded for the left and right FOVs, respectively.

Figure 6.4c demonstrates that all 14 human subjects exhibited inner canthi temperatures ranging from 34.67 °C to 36.12 °C, aligning well with the ground truth. The temperatures of the left and right inner canthi for male subjects were measured to be 35.41±0.42 °C (mean ± standard

deviation) and 35.38±0.50 °C, respectively, while the counterparts of the female subjects were 35.51±0.48 °C and 35.62±0.57 °C. These measurements reveal a temperature difference between male and female subjects of 0.17 ºC. These results align well with findings from previous clinical studies (Teunissen & Daanen, 2011a; Vardasca, 2017). Moreover, considering the recommended threshold temperature for fever screening based on inner canthi to be 37.1°C (Foster *et al.*, 2021), all participants were classified as fever-negatives.

6.3.5 Temperature monitoring using SPIRIT

We further investigated the temperature variation of the inner canthi. In the first experiment, we imaged four subjects at four time points (i.e., 10 AM, 12 PM, 2 PM, and 4 PM) in a single day. As illustrated in Fig. 6.4d, across all four subjects, temperatures increase ~1 °C over 6 hours with a fluctuation cycle of 4 hours. These findings align well with previous clinical studies, affirming that a healthy individual's body temperature varies throughout the day—typically manifesting lower temperatures in the morning and higher readings in the late afternoon—in response to the body's varying needs and activities (Walker *et al.*, 1990).

In the second experiment, we examined the influence of glasses on the inner canthus temperature. Glasses can establish a microclimate between the inner canthus and the lens (Foster *et al.*, 2021). Nonetheless, to our best knowledge, this phenomenon has not been specifically explored, and its impact remains undetermined. To fill out this void, SPIRIT was employed to image eight human subjects, four of whom had never worn glasses while the remaining wore glasses daily. To eliminate the influence of acquisition time on thermal behavior, subjects were divided into four groups, each being measured at approximately the same time. As depicted in Figure 6.4e, the subjects who do not wear glasses exhibit an average inner canthus temperature of 35.74±0.21 °C, while the glass-wearing subjects, measured immediately after glass removal, have an inner canthus temperature of 35.44±0.23 °C. The 0.30 °C-difference between the two groups is likely attributed to the increased airflow in the tunnel regions formed by the nose pads of glasses, situated near the inner canthi.

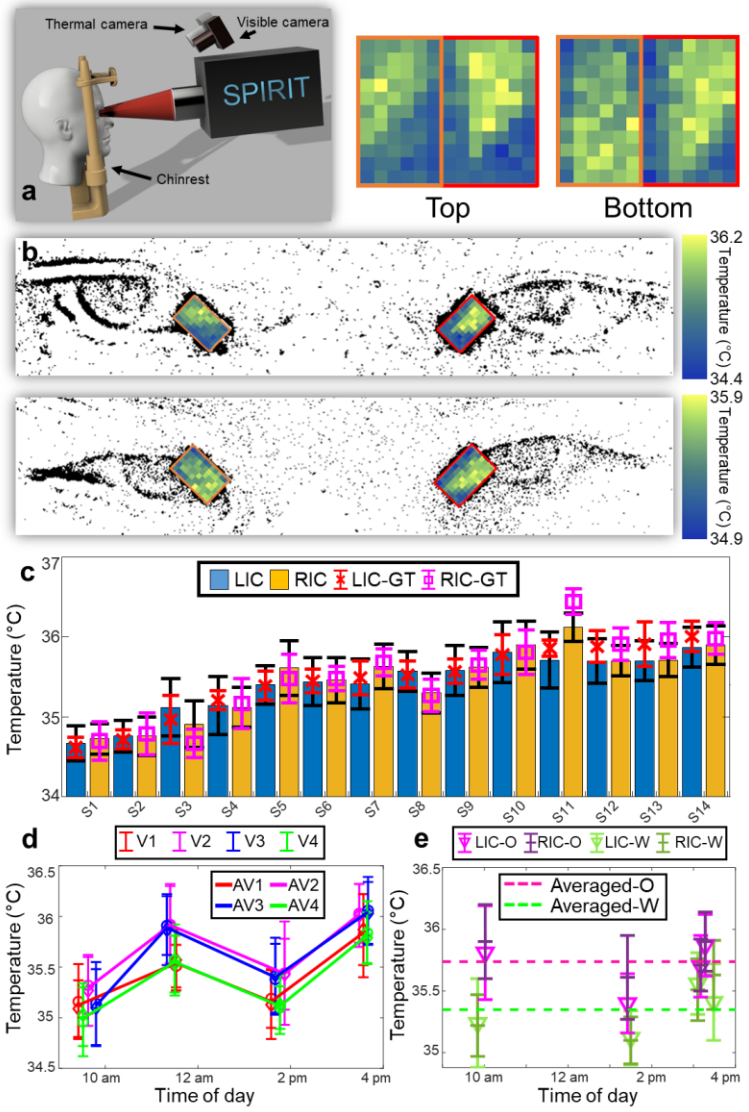

Figure 6.4 Human inner canthi infrared thermography of SPIRIT.

a System arrangement of SPIRIT for human subjects. b Representative reconstruction results for two human subjects (top right). Photographs of the human subjects, aligned with the rearranged reconstruction results to the inner canthi regions. c Comparison of reconstructed temperature information with ground truth. S1-S14: individual human subjects. LIC/RIC: measured temperature of the left/right inner canthus of subjects. LIC-GT/RIC-GT: ground truth for left/right inner canthus temperature. d Daytime variation in reconstructed temperature information. A1-A4: Human subjects, each denoted by a different color. Diamond and circle markers individually indicate left and right inner canthi. AV1-AV4: averaged temperatures for subjects V1-V4, with lines connecting averaged temperatures at various times to display temperature trajectory. e Temperature reconstruction comparison based on glass-wearing habits. LIC-O/RIC-O: left/right inner canthi temperatures of individuals who do not wear glasses. LIC-W/RIC-W: left/right inner canthi temperatures of long-term eyeglass wearers (> one year). Averaged temperatures for the two groups are indicated by purple and green dashed lines, respectively. Note: In (b), (c), and (d), error bars represent standard deviations of temperature across inner canthi pixels.

6.4 Discusson and conclusions

We have developed SPIRIT. Utilizing diagonal aggregation of cyclic S-matrices to compactly assemble masking patterns, SPIRIT implements a 1D scanning for spatial encoding. Leveraging the 2D smoothness of the encoded data matrix, condensed encoding matrices with a data interpolation are used for an efficient reconstruction. As the first SPI system operating in the LWIR spectral range, SPIRIT passively detects the thermal radiation from human inner canthi for temperature mapping. We demonstrated SPIRIT's effectiveness for inner canthi thermometry in human subjects. Statistical analysis revealed no significant difference in inner canthi temperatures between the left and right canthus for individual personnel or between male and female subjects. SPIRIT was also used to examine the daytime temperature variation of human inner canthi, aligning well with previous clinical studies. Moreover, SPIRIT investigated a temperature differential of 0.3°C between two subject groups: those who wore eyeglasses and those who did not.

SPIRIT possesses a number of technical advantages and practical merits. First, it avoids the spectral sensing limitation imposed by the spatial light modulators—e.g., digital micromirror device (DMD), which extends SPI's operation to the LWIR range. Moreover, with passive detection, it address the safety concern by eliminating additional active laser illumination. Finally, in contrast to existing 2D thermography techniques, SPIRIT designates FOVs on the inner canthi—a key area indicative of human's core body temperature. This configuration eliminates the superfluous information, intrinsically reduces background blending.

While currently limited by the specifications of core components, SPIRIT performance could be further enhanced by optimizing the mask scanning mechanism and software. The imaging speed of the current SPIRIT system is limited by the maximum mechanical scanning speed of the motorized *z*-translational stage (i.e., 2.4 mm/s), which results in a total data acquisition time of approximately 15 s. By integrating a motorized stage that boasts a quicker scanning rate, SPIRIT's imaging speed could be further enhanced (Song *et al.*, 2010). Finally, the currently deployed interpolation method for SPIRIT is lightweight. Undoubtedly, other sophisticated interpolation methods could also be explored to potentially improve reconstruction quality while maintaining fast reconstruction (Chai *et al.*, 2020a; Chai *et al.*, 2020b).

As a SPI imaging modality for far-wavelength IR, SPIRIT holds promises in terms of broad application scope. Besides the thermal imaging of human inner canthi demonstrated in this work, the principle of SPIRIT is readily applicable to identifying inflammatory conditions, circulatory disorders, or detecting breast cancer by perceiving abnormal heat patterns (Ng & Etehadtavakol,

2017; Nur, 2014), monitoring activities even in obscured conditions like fog or smoke (Akula *et al.*, 2011; Wong *et al.*, 2009), and identifying heat signatures of living beings in rescue operations (Andrea *et al.*, 2018). In addition, SPIRIT could be adopted for astronomical and atmospheric observations, like investigating cool celestial bodies and imaging the black body radiation in the observation of interstellar gases (Haworth *et al.*, 2018). Finally, SPIRIT may find applications in material characterization by investigating the thermal properties of materials, particularly in semiconductor research (Kamble *et al.*, 2013; Kölzer *et al.*, 1996).

6.5 Supplementary information

6.5.1 Supplementary Note 1: Details of alignment between the mask and window

In the SPIRIT methodology, generating the individual encoding pattern involves utilizing a window to selectively expose a specific masking pattern to thermal radiation, necessitating meticulous alignment between the mask and window to ensure measurement precision.

As shown in Fig. 6.5a and b, the mask and window were fabricated on two-inch round plates via laser drilling. A close-up view of Fig. 6.5a reveals that although the pixel was designed to be 250 μm × 250 μm, to avoid ‘island pixel’, the actual fabricated pixel size was reduced to 200 μm × 200 μm to preclude ‘island pixel’ issues, which could jeopardize successful fabrication.

The alignment process commences with a vertical scanning check of the mask, aiming to confirm the mask’s scanning motion is perpendicular to the encoding beam path. Specifically, two vertical alignment markers, each of which contains one pixel, were designed and fabricated on both the upper and lower sides of the encoding stripes, highlighted as red boxes in Fig. 6.5a. Utilizing a thermal imaging camera (Yoseen M384D), the horizontal positions of the alignment markers were recorded during the mask's vertical scanning and subsequently employed to guide the mask's rotational adjustment. Upon achieving horizontal alignment of the markers, the rotational mount was mechanically secured to prevent any subsequent rotational movement.

Post vertical scanning check, the window—affixed to a separate rotational mount on top of a two-axis translation stage—was cautiously moved towards the mask until they almost touch. The thermal imaging camera continuously monitored the region cropped through the window. Fig. 6.5c-f illustrates that the exact one-to-one correspondence between the predetermined encoding patterns (Fig. 6.5c and e) and the images acquired by the thermal imaging camera (Fig. 6.5d and

f) affirms the precision in the alignment of the window and mask. In a similar fashion, a systematic check of all 91 encoding patterns was performed.

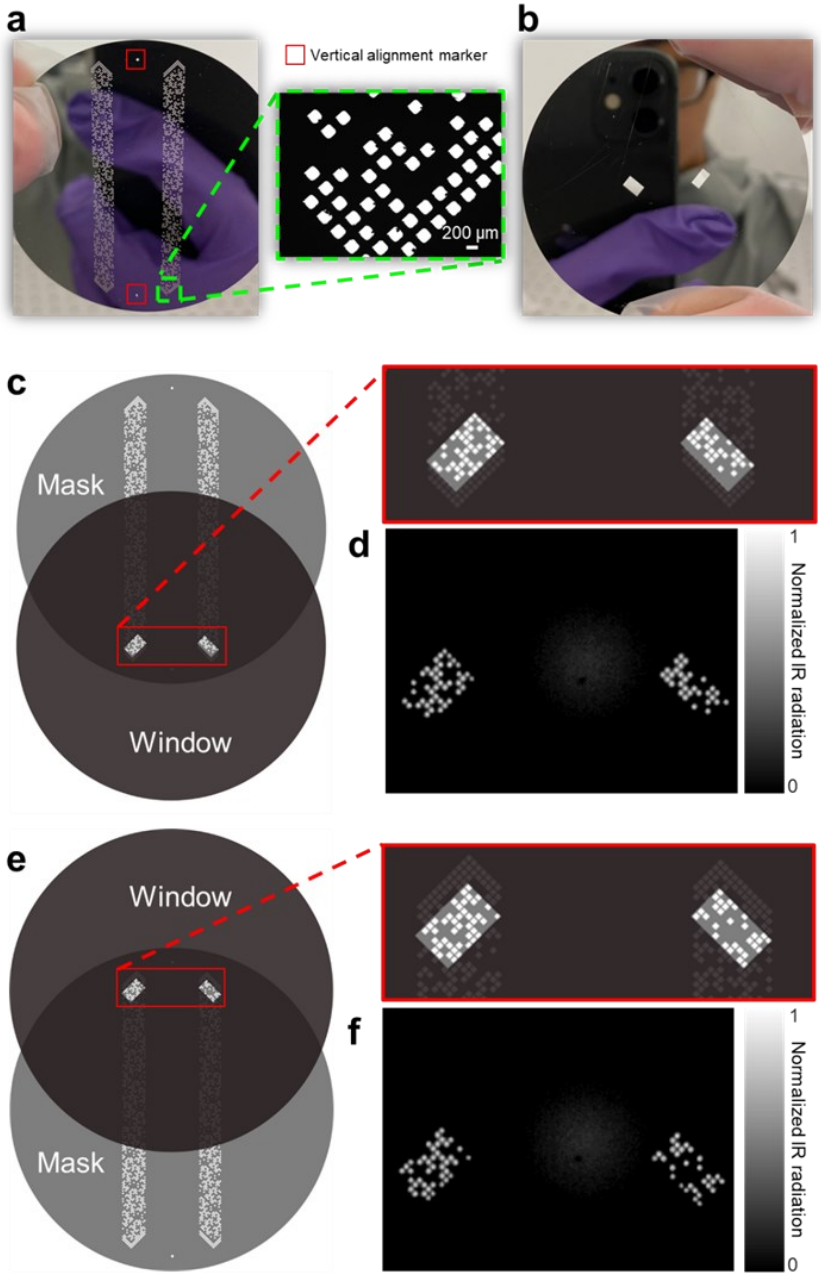

Figure 6.5 Details of alignment check between the mask and window.

a Photograph of the mask plate. Close-up view of a microscopy inspected image is shown in the dashed box. The red boxes cropped the vertical alignment markers on the mask plate. b Photograph of the window plate. c-f Two selective alignment check. c and e: the predetermined alignment setups, with close-up views shown in the red boxes. d and f: the images captured by the thermal imaging camera.

6.5.2 Supplementary Note 2: System synchronization and data acquisition

The SPIRIT measurement procedure involves a vertical mask scanning facilitated by a motorized z-translation stage (MZT) in stepper mode, and the collection of bucket signals via a single-pixel detector, in conjunction with an optical chopper and a lock-in amplifier (LA). The acquisition of data from LA used a digitizer card (ATS9625, AlazarTech), which interfaces with a personal computer (PC), and is stored as waveforms sampled at a rate of 20 MHz. The bucket signals undergo discrete sampling of the LA waveform and subsequent averaging. Fig. 6.6 illustrates the signal path and timing diagrams.

The LA waveform is registered with the MZT through pre-calibrated position signals. Following the verification of alignment between the mask and window, as detailed in Supplementary Note 2, each encoding pattern specifies a particular vertical scan position, readable from the MZT. These pre-stored positions are activated as the MZT transitions to respective positions. Fig. 6.6b shows instances of LA, MZT waveforms, and the sampling procedure in the PC, corresponding with registration patterns.

During vertical scanning, the MZT initially lowers the mask to a position where the window is elevated 5 mm above the mask's first runway pattern. At this juncture, the window is obstructed, and no encoding patterns are exposed until the mask is lifted by the MZT. As depicted in Fig. 6.6b, the signal harvested by the LA initially remains at a low amplitude, then surges to a peak value, indicating that the window is cropping the first runway pattern of the mask (denoted as (i) in Fig. 6.6b). When the mask is scanned to the position of the first encoding pattern (labeled as (ii) in Fig. 6.6b), the MZT’s online position readout aligns with the pre-calibrated position signal. Subsequently, signals from the LA to the digitizer card are averaged and stored as the encoding bucket signal within the time interval until the MZT reports a differing position.

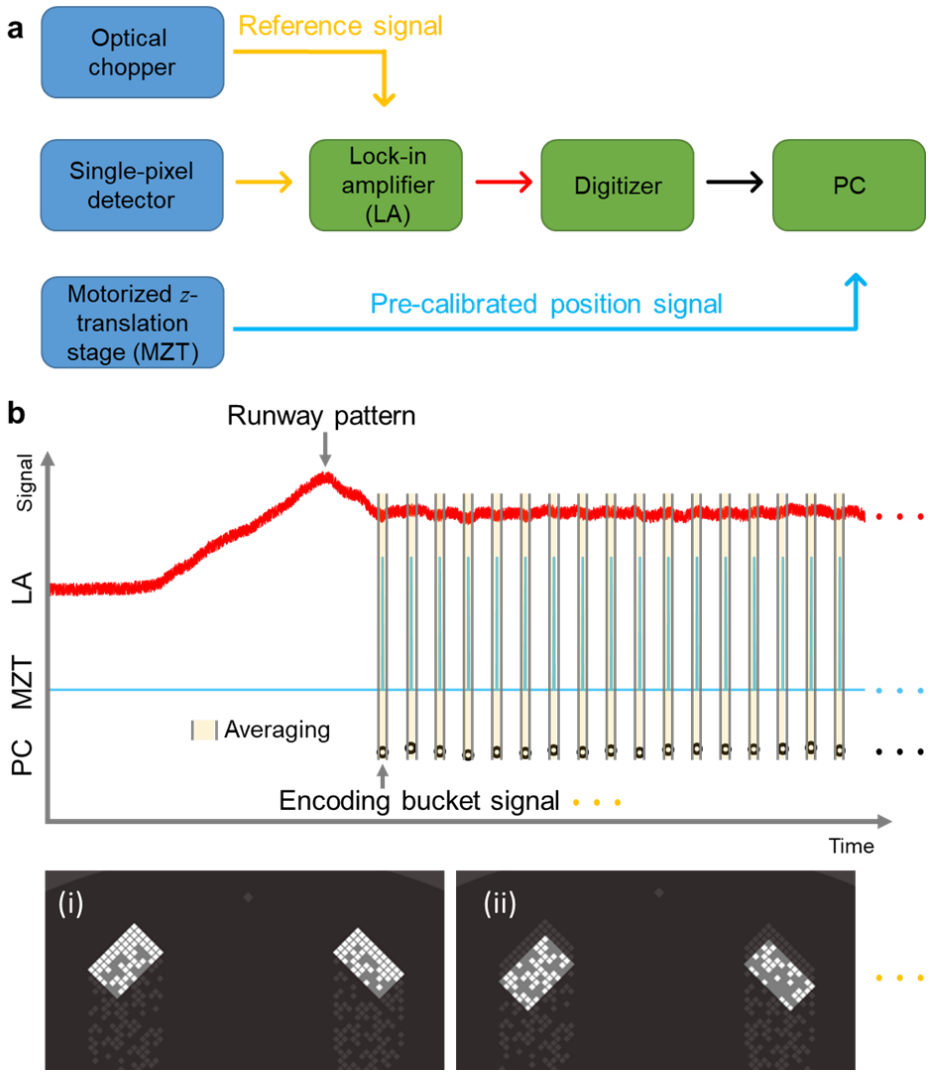

Figure 6.6 Illustration of synchronization and data registration in SPIRIT.

a Block diagram illustrating the paths of signals involved in timing and acquisition. b Schematic timing diagram illustrating synchronization and data registration. Insets show two initial mask patterns: (i) runway pattern, (ii) the first aggregate encoding pattern.

6.5.3 Supplementary Note 3: Comparison of different cooling methods with SPIRIT

Navigating the challenges of a notably low SNR of 0.007, as indicated by the blue dataset in Fig. 6.7, SPIRIT—an imaging modality utilizing passive detection—has sought strategies to mitigate systematic thermal noises originating from various sources such as ambient environment, equipment, and reflected thermal radiation from numerous surfaces. To pare down the system noise level, both thermal management and the implementation of a lock-in amplifier were tested to enhance the SNR.

Engaging in thermal management, the encoding and detection system of SPIRIT was encapsulated and relocated into a specially modified fridge, equipped with custom channels for control cables and an imaging window. Despite its partially open structure somewhat mitigating cooling efficacy, the fridge maintained an internal temperature of approximately -5°C. The impact of this adjustment is illustrated by the yellow dataset in Fig. 6.7: a reduced noise level and an elevated SNR of 1.01. Nevertheless, this method introduced an unforeseen variable; the presence of approximately 40% humidity in the experimental space led to the formation of water droplets and ice crystals, potentially obstructing encoding pixels. Consequently, a lock-in amplifier was employed to enhance SNR while maintaining the system at room temperature.

The lock-in amplifier, a robust instrument renowned in scientific research, is adept at extracting signals amid pervasive noise environments. This instrument harnesses the principle of synchronous detection and low-pass filtering to amplify a signal characterized by a known carrier wave amidst noise. Within the SPIRIT framework, an optical chopper, functioning at 850 Hz, was positioned in front of the single-pixel detector to lock the acquired signal. The time constant of the low-pass filter can be modulated to further curtail noise, while an enhanced SNR comes at the cost of a slower response time if a larger time constant is used. Thus, through empirical selection, a time constant and low-pass filter level were set at 30 ms and 12 dB respectively. Illustrated by the orange dataset in Fig. 6.7, the SNR is bolstered to 3.14, signaling a marked improvement compared to the thermal management method.

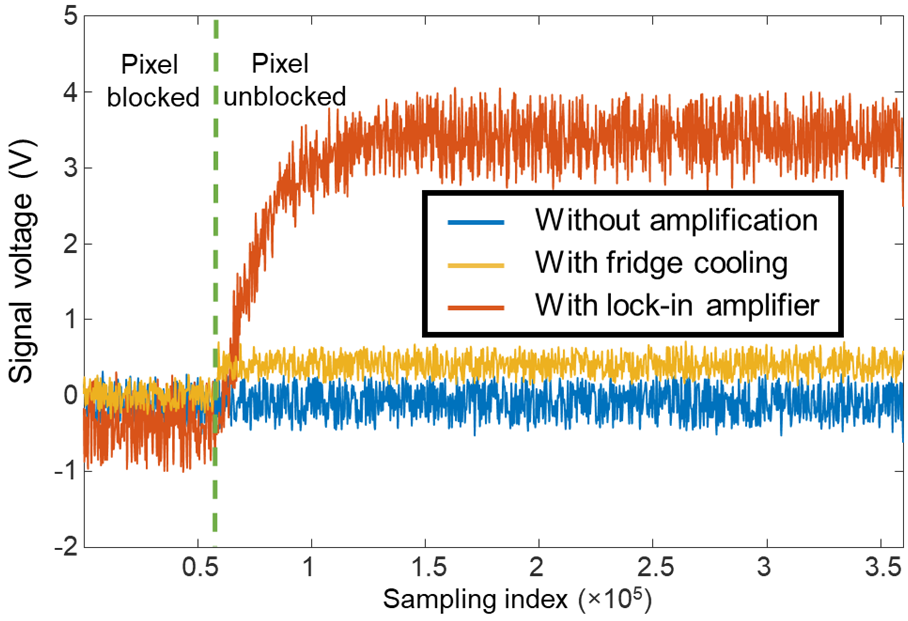

Figure 6.7 Comparison of signal-to-noise ratio with different cooling setups in SPIRIT.

A plate with low thermal conductivity is used to block and unblock the encoding pixels with a constant temperature background provided by a black body thermal radiation source.

6.5.4 Supplementary Note 4: Design search for construction of the physical masks

For the design of the physical masks used in SPIRIT, it was necessary to select n, the order of a constructable cyclic S-matrix, and m, the number of masking patterns to be aggregated into the physical mask. In order to make the best possible selection, a systematic computer search was carried out to survey possible values of n and m that suited the constraints of manufacturability and overall size while still providing a sufficient pixel count.

Although a wide variety of techniques can generate S-matrices broadly, only three construction techniques are currently known for cyclic S-matrices (Sloane & Harwit, 1979) corresponding to values of n of the form (i) $n = p$ for prime p of the form $4k - 1$, (ii) $n = 2^k - 1$ for certain $k \geq 2$, and (iii) $n = pq$ for p and $q = p + 2$ both prime. First, in order to accommodate encoding patterns with rectangular shape, n was required to be a composite number ($n = pq$), thus eliminating constructions of type (i), and some particular constructions of type (ii) (i.e. the Mersenne primes). Further, the condition $1/3 < p/q < 3$ for some factorization of n was introduced in order to constrain the aspect ratio of the system FOV. This condition, while always met by constructions of type (iii), provided a useful filter for cases of type (ii) constructions with highly composite n.

For a given n with factors p and q specifying a possible encoding pattern design, two further pieces of information were derived. First, it is desirable that diagonal scanning ultimately result in the deployment of all masking patterns contained within the measurement matrix. Given row-major reshaping of measurement matrix rows into encoding patterns of size $p \times q$, this is mathematically equivalent to the condition that the function $f(x) = x(q + 1)$ with $x \in \{0, \dots, n - 1\}$ range over all values $\{0, \dots, n - 1\}$ when considered modulo n. Alternatively, this condition is equivalent to

$$\gcd(q + 1, n) = 1, \tag{S1}$$

where $\gcd(x, y)$ denotes the greatest common divisor between integers x and y. Second, it is desirable that a given design allow for a truncation $m < n$ in the number of encoding patterns deployed by a physical mask while still producing a relatively uniform distribution of acquired data for the purposes of well-conditioned 2D interpolation as part of image reconstruction.

Mathematically, the distribution of data in the $p \times q$ array used for interpolation may be described for a given m as the set of row and column coordinates

$$\left\{ \left(\mathrm{mod}\left(\left\lfloor \frac{k(q+1)}{q} \right\rfloor, p \right),\ \mathrm{mod}(k(q+1), q) \right) \ \middle|\ k = 0 \ldots m-1 \right\}, \tag{S2}$$

where $\mathrm{mod}(x, y)$ denotes the value of the integer x taken modulo the positive integer y. As k increases the $p \times q$ array is gradually filled by a set of diagonal stripes that wrap around the top and bottom and left and right edges, respectively. Choosing values of m that produce complete stripes with left/right wrapping amounts to considering $m = ql$ for values of $l = 0 \ldots p-1$. In this case, the first (row) coordinates produced by equation (S2) provide sufficient information to evaluate the even-ness of the diagonal filling. This may be quantified by considering the set of values

$$A = \{\mathrm{mod}(i(q+1), p) | i = 0 \ldots l-1\}, \tag{S3}$$

that, as a consequence of equation (S1), consists of l distinct values. Writing these elements as a sequence $\tilde{p}_0, \tilde{p}_1, \ldots, \tilde{p}_{l-1}$ arranged in strictly ascending order, the largest cyclic 'gap' admitted by this distribution may be quantified as

$$d_l = \max\{\tilde{p}_1 - \tilde{p}_0, \tilde{p}_2 - \tilde{p}_1, \ldots, \tilde{p}_{l-1} - \tilde{p}_{l-2}, \tilde{p}_0 + p - \tilde{p}_{l-1}\}. \tag{S4}$$

The possible values of d_l can be shown to satisfy the lower bound $\left\lceil {}^{p}/_{l} \right\rceil \leq q_l$. Thus, we may identify values of m producing ideal distributions of data to be of the form $m = q\tilde{l}$ for values of $\tilde{l}$ equaling the lower gap bound $d_{\tilde{l}} = \left\lceil {}^{p}/_{\tilde{l}} \right\rceil$.

Various physical constraints determined an upper bound on the values of m compatible with the SPIRIT system. First, the 150 μm thickness of the metal mask substrate dictated the lower bound on the size of pixel apertures that could be both fabricated with acceptable accuracy via laser drilling, and compatible with large numerical aperture. Second, the use of 2 inch diameter rotational mounts dictated the total amount of space available for encoding mask fabrication. Together, these constraints dictate a rough theoretical limit of 350 for the number of encoding patterns that can be accommodated by SPIRIT's physical mask hardware.

Based on the above considerations, values of n in the range [1,350] were searched for designs compatible with the construction of cyclic S-matrices and the various constraints on the reshaping dimensions p and q. Results of this search are compiled into table S1. Factors that influenced the final design selection included the desire for additional space allowances, high

pixel count, and a moderate level of compression ratio m/n. Ultimately, the values chosen for the construction of the masks used in SPIRIT used a value of $n = 143$ corresponding to construction type (iii) with $p = 11$ and $q = 13$. This design admitted a truncation length of $m = 91$, corresponding to a uniform distribution of data with a sampling ratio of 63.6%.

Table 6.1 Design search results for the construction of physical masks used in SPIRIT based on cyclic S-matrices.

(Bold values on the highlighted line indicate the parameters chosen for actual fabrication)

n	Type	p	q	m	m/n (%)
15	(iii)	5	3	15, 12	100.0, 80.0
35	(iii)	5	7	35, 21, 14	100.0, 60.0, 40.0
35	(iii)	7	5	35, 30	100.0, 85.7
63	(ii)	7	9	63, 45, 27, 18	100.0, 71.4, 42.9, 28.6
63	(ii)	9	7	63, 56	100.0, 88.9
143	**(iii)**	**11**	**13**	143, **91**, 52	100.0, **63.6**, 36.4
143	(iii)	13	11	143, 132	100.0, 92.3
255	(ii)	17	15	255, 240	100.0, 94.1
323	(iii)	17	19	323, 209, 114	100.0, 64.7, 35.3
323	(iii)	19	17	323, 306	100.0, 94.7

6.5.5 Supplementary Note 5: Details of the FOV registration

The registration of the FOVs is pivotal to SPIRIT's capacity to accurately identify and image the correct targeted regions, accommodating for the anatomical variations inherent across different individuals. The precision in positioning, both in simulation and practical application, ensures the fidelity of SPIRIT's thermal imaging, safeguarding against data corruption or misrepresentation resulting from misaligned FOVs.

The frame sizes utilized in SPIRIT incorporate dimensions of 11 × 7 pixels and 11 × 6 pixels for the left and right encoding masks, respectively, aligning with targeted FOVs measuring 11 mm × 7 mm and 11 mm × 6 mm. As depicted in Fig. 6.8a, a human full-face mask, affixed to a chinrest and positioned opposite to SPIRIT, mimics actual human subject conditions. Two Peltier coolers, each measuring 8.8 mm × 8.8 mm, are adhered to the inner canthi regions of the

face mask, positioned 30 mm apart. The hot sides of the Peltier coolers achieve a temperature of approximately 70°C, generating a substantial thermal contrast relative to the ambient temperature, which hovers around 22°C. Additionally, a visible camera, mounted atop the single-pixel detector, captures the face mask and aids in pinpointing the two FOVs on the screen view.

Fig. 6.8b-d depict the operation of FOVs registration. SPIRIT's reconstructions (displayed in the right rows) act as indicators, facilitating the identification of the FOVs' positions on the image captured by the visible camera. To mitigate potential misinterpretations between the reconstruction and corresponding FOVs, the target (face mask) is manipulated, moving in specified directions and distances (indicated by blue arrows in Fig. 6.8b-d), until a congruence between the reconstructions and FOVs is achieved.

Ultimately, the finely registered FOVs are illustrated by two markers on the screen view of the visible camera, serving as guides for human subjects to accurately position their heads on the same chinrest, ensuring their inner canthi are precisely targeted on the FOVs of SPIRIT.

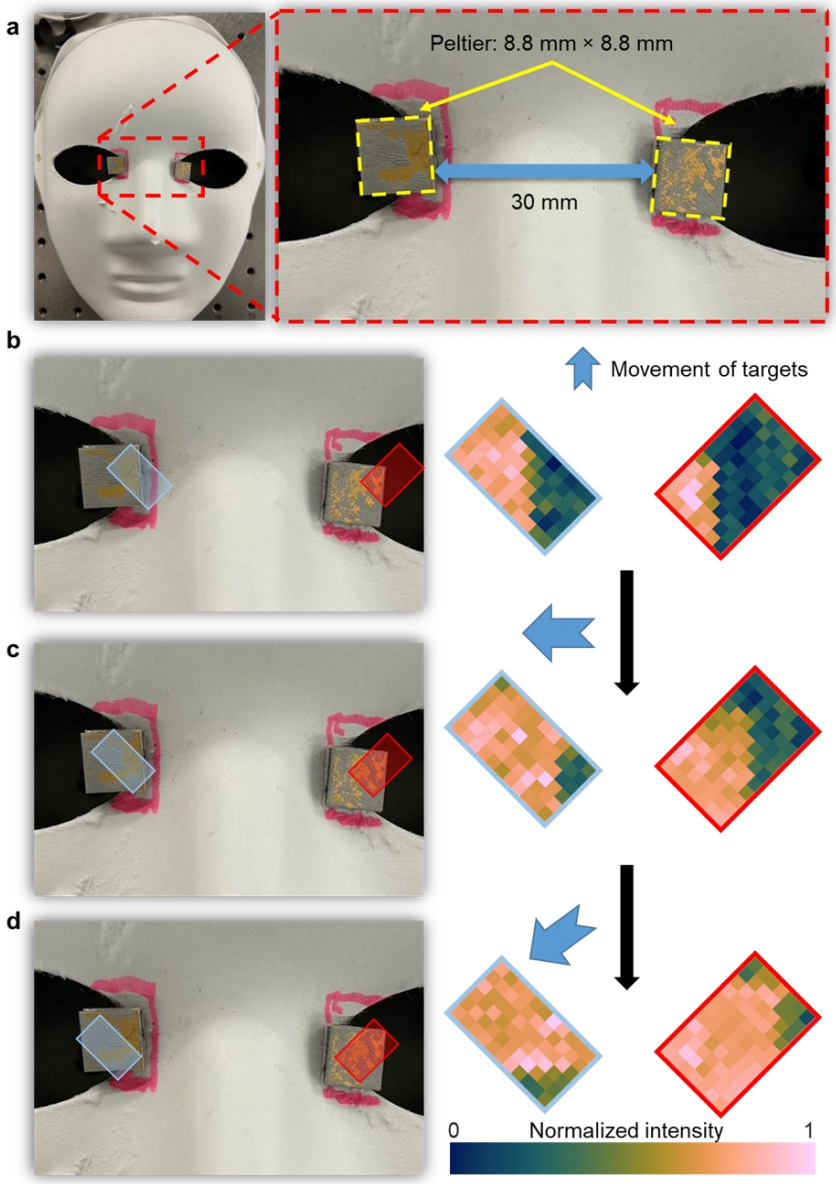

Figure 6.8 Illustration of fields of view registration in SPIRIT.

a Photograph of face mask complemented by Peltier coolers. A close-up view is shown in red-dashed box. b-d Registration procedure between SPIRIT reconstructions (presented in the right rows) and imagery captured by the visible camera (featured in the left rows). The light blue and red boxes indicate the left and right FOVs, respectively. The blue arrows represent the direction of the target's movements. The black arrows illustrate the operational steps, guiding through the process in a sequential manner.

6.6 References

Abdullah N, Talib ARA, Jaafar AA, Salleh MAM & Chong WT (2010) The basics and issues of thermochromic liquid crystal calibrations. *Experimental Thermal and Fluid Science* 34(8):1089-1121.

Akula A, Ghosh R & Sardana H (2011) Thermal imaging and its application in defence systems. *AIP conference proceedings.* American Institute of Physics, p 333-335.

Al Raisi SI, Pouliopoulos J, Barry MT, Qian P, Thiagalingam A, Swinnen J & Kovoor P (2020) Renal artery branch denervation: Evaluation of lesion characteristics using a thermochromic liquid crystal phantom model. *Heart, Lung and Circulation* 29(3):445-451.

al. WHPe (2007) Numerical Recipes: The Art of Scientific Computing. *Cambridge University Press* 3rd Ed.

Andrea CC, Byron JQ, Jorge PI, Inti TC & Aguilar WG (2018) Geolocation and counting of people with aerial thermal imaging for rescue purposes. *Augmented Reality, Virtual Reality, and Computer Graphics: 5th International Conference, AVR 2018, Otranto, Italy, June 24–27, 2018, Proceedings, Part I 5.* Springer, p 171-182.

Anonyme (2023) Medical electrical equipment — Part 2-59: Particular requirements for the basic safety and essential performance of screening thermographs for human febrile temperature screening — Amendment 1. *International Electrotechnical Commission* IEC 80601-2-59:2017/Amd 1:2023.

Avdelidis N & Moropoulou A (2004) Applications of infrared thermography for the investigation of historic structures. *Journal of Cultural Heritage* 5(1):119-127.

Bradley D & Matthews K (1968) Measurement of high gas temperatures with fine wire thermocouples. *Journal of mechanical engineering science* 10(4):299-305.

Buono MJ, Jechort A, Marques R, Smith C & Welch J (2007) Comparison of infrared versus contact thermometry for measuring skin temperature during exercise in the heat. *Physiological measurement* 28(8):855.

Carlomagno GM & Cardone G (2010) Infrared thermography for convective heat transfer measurements. *Experiments in fluids* 49:1187-1218.

Chai X, Gu H, Li F, Duan H, Hu X & Lin K (2020a) Deep learning for irregularly and regularly missing data reconstruction. *Scientific reports* 10(1):3302.

Chai X, Tang G, Wang S, Peng R, Chen W & Li J (2020b) Deep learning for regularly missing data reconstruction. *IEEE Transactions on Geoscience and Remote Sensing* 58(6):4406-4423.

Charlton M, Stanley SA, Whitman Z, Wenn V, Coats TJ, Sims M & Thompson JP (2020) The effect of constitutive pigmentation on the measured emissivity of human skin. *Plos one* 15(11):e0241843.

Childs P (2001) *Practical temperature measurement.* Elsevier,

Childs PR, Greenwood J & Long C (2000) Review of temperature measurement. *Review of scientific instruments* 71(8):2959-2978.

Cohn JH (1963) On the value of determinants. *Proceedings of the American Mathematical Society* 14(4):581-588.

Diwanji MM, Hisvankar SM & Khandelwal CS (2020) Temperature Measurement using Infrared Contactless Thermal Gun. *2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC).* IEEE, p 134-137.

Dong K, Li J, Zhang T, Gu F, Cai Y, Gupta N, Tang K, Javey A, Yao J & Wu J (2023) Single-pixel reconstructive mid-infrared micro-spectrometer. *Optics Express* 31(9):14367-14376.

Duarte MF, Davenport MA, Takhar D, Laska JN, Sun T, Kelly KF & Baraniuk RG (2008) Single-pixel imaging via compressive sampling. *IEEE signal processing magazine* 25(2):83-91.

Ebner A, Gattinger P, Zorin I, Krainer L, Rankl C & Brandstetter M (2023) Diffraction-limited hyperspectral mid-infrared single-pixel microscopy. *Scientific Reports* 13(1):281.

Edgar MP, Gibson GM, Bowman RW, Sun B, Radwell N, Mitchell KJ, Welsh SS & Padgett MJ (2015) Simultaneous real-time visible and infrared video with single-pixel detectors. *Scientific reports* 5(1):10669.

Foster J, Lloyd AB & Havenith G (2021) Non-contact infrared assessment of human body temperature: The journal Temperature toolbox. *Temperature* 8(4):306-319.

Furstenberg R, Kendziora C, Stepnowski J, Stepnowski S, Rake M, Papantonakis M, Nguyen V, Hubler G & McGill R (2008) Stand-off detection of trace explosives via resonant infrared photothermal imaging. *Applied Physics Letters* 93(22).

Gade R & Moeslund TB (2014) Thermal cameras and applications: a survey. *Machine vision and applications* 25:245-262.

Gibson GM, Johnson SD & Padgett MJ (2020) Single-pixel imaging 12 years on: a review. *Optics express* 28(19):28190-28208.

Gibson GM, Sun B, Edgar MP, Phillips DB, Hempler N, Maker GT, Malcolm GP & Padgett MJ (2017) Real-time imaging of methane gas leaks using a single-pixel camera. *Optics express* 25(4):2998-3005.

Gopalsami N, Liao S, Elmer TW, Koehl ER, Heifetz A, Raptis AC, Spinoulas L & Katsaggelos AK (2012) Passive millimeter-wave imaging with compressive sensing. *Optical Engineering* 51(9):091614-091614.

Haworth TJ, Glover SC, Koepferl CM, Bisbas TG & Dale JE (2018) Synthetic observations of star formation and the interstellar medium. *New Astronomy Reviews* 82:1-58.

Hayat N, Alkhairy S, Cheema A, Ehsan M & Khan MA (2019) Normal interpupillary, inner canthal distance and outer canthal distance in a normal population of Pakistan. *Pak J Med Sci* 35(1):50-54.

Heitor M & Moreira A (1993) Thermocouples and sample probes for combustion studies. *Progress in energy and combustion science* 19(3):259-278.

Hughes-Riley T, Lugoda P, Dias T, Trabi CL & Morris RH (2017) A study of thermistor performance within a textile structure. *Sensors* 17(8):1804.

Kamble VB, Bhat S & Umarji A (2013) Investigating thermal stability of structural defects and its effect on d ferromagnetism in undoped SnO2. *Journal of Applied Physics* 113(24).

Kasap S (2001) Thermoelectric effects in metals: thermocouples. *Canada: Department of Electrical Engineering University of Saskatchewan*.

Kilcullen P, Ozaki T & Liang J (2022) Compressed ultrahigh-speed single-pixel imaging by swept aggregate patterns. *Nature Communications* 13(1):7879.

Kölzer J, Oesterschulze E & Deboy G (1996) Thermal imaging and measurement techniques for electronic materials and devices. *Microelectronic engineering* 31(1-4):251-270.

Lahiri BB, Bagavathiappan S, Jayakumar T & Philip J (2012) Medical applications of infrared thermography: a review. *Infrared physics & technology* 55(4):221-235.

Liang J, Gao L, Li C & Wang LV (2014) Spatially Fourier-encoded photoacoustic microscopy using a digital micromirror device. *Optics letters* 39(3):430-433.

Liu X, Skripka A, Lai Y, Jiang C, Liu J, Vetrone F & Liang J (2021) Fast wide-field upconversion luminescence lifetime thermometry enabled by single-shot compressed ultrahigh-speed imaging. *Nature Communications* 12(1):6401.

Martín Sánchez FJ, Martínez-Sellés M, Molero García JM, Moreno Guillén S, Rodríguez-Artalejo FJ, Ruiz-Galiana J, Cantón R, De Lucas Ramos P, García-Botella A, García-Lledó A, Hernández-Sampelayo T, Gómez-Pavón J, González Del Castillo J, Martín-Delgado MC & Bouza E (2023) Insights for COVID-19 in 2023. *Rev Esp Quimioter* 36(2):114-124.

Meola C & Carlomagno GM (2004) Recent advances in the use of infrared thermography. *Measurement science and technology* 15(9):R27.

Meola C, Carlomagno GM & Giorleo L (2004) The use of infrared thermography for materials characterization. *Journal of materials processing Technology* 155:1132-1137.

Mosbach K & Danielsson B (1974) An enzyme thermistor. *Biochimica et Biophysica Acta (BBA)-Enzymology* 364(1):140-145.

Ng DK, Chan C-H, Lee RS & Leung LC (2005) Non-contact infrared thermometry temperature measurement for screening fever in children. *Annals of tropical paediatrics* 25(4):267-275.

Ng EY & Etehadtavakol M (2017) *Application of infrared to biomedical sciences.* Springer,

Ng EY, Kawb G & Chang W (2004) Analysis of IR thermal imager for mass blind fever screening. *Microvascular research* 68(2):104-109.

Nur R (2014) *Identification of thermal abnormalities by analysis of abdominal infrared thermal images of neonatal patients*. (Carleton University).

Pealat M, Bouchardy P, Lefebvre M & Taran J-P (1985) Precision of multiplex CARS temperature measurements. *Applied optics* 24(7):1012-1022.

Perpetuini D, Filippini C, Cardone D & Merla A (2021) An Overview of Thermal Infrared Imaging-Based Screenings during Pandemic Emergencies. *International Journal of Environmental Research and Public Health* 18(6):3286.

Pramanik M & Wang LV (2009) Thermoacoustic and photoacoustic sensing of temperature. *Journal of biomedical optics* 14(5):054024-054024-054027.

Rajic N (2002) Principal component thermography for flaw contrast enhancement and flaw depth characterisation in composite structures. *Composite structures* 58(4):521-528.

Reiner R, Collins JK, Team C-F & Murray CJ (2023) Forecasting the trajectory of the COVID-19 pandemic into 2023 under plausible variant and intervention scenarios: a global modelling study. *medRxiv* :2023.2003. 2007.23286952.

Ring EFJ & Ammer K (2015) The technique of infrared imaging in medicine*. in *Infrared Imaging* (IOP Publishing), p 1-1-1-10.

Saegusa Y & Tabata H (2003) Usefulness of infrared thermometry in determining body temperature in mice. *Journal of veterinary medical science* 65(12):1365-1367.

Sage I (2011) Thermochromic liquid crystals. *Liquid crystals* 38(11-12):1551-1561.

Shah J, Park S, Aglyamov S, Larson T, Ma L, Sokolov K, Johnston K, Milner T & Emelianov SY (2008) Photoacoustic imaging and temperature measurement for photothermal cancer therapy. *Journal of biomedical optics* 13(3):034024-034024-034029.

Sloane NJ & Harwit M (1976) Masks for Hadamard transform optics, and weighing designs. *Applied optics* 15(1):107-114.

Sloane NJ & Harwit M (1979) Hadamard Transform Optics. (Academic Press New York).

Smith C, Sabatino D & Praisner T (2001) Temperature sensing with thermochromic liquid crystals. *Experiments in fluids* 30(2):190-201.

Sollai S, Dani C, Berti E, Fancelli C, Galli L, de Martino M & Chiappini E (2016) Performance of a non-contact infrared thermometer in healthy newborns. *BMJ open* 6(3):e008695.

Song Y, Wang J, Yang K, Yin W & Zhu Y (2010) A dual-stage control system for high-speed, ultra-precise linear motion. *The International Journal of Advanced Manufacturing Technology* 48:633-643.

Stantchev RI, Sun B, Hornett SM, Hobson PA, Gibson GM, Padgett MJ & Hendry E (2016) Noninvasive, near-field terahertz imaging of hidden objects using a single-pixel detector. *Science advances* 2(6):e1600190.

Stantchev RI, Yu X, Blu T & Pickwell-MacPherson E (2020) Real-time terahertz imaging with a single-pixel detector. *Nature communications* 11(1):2535.

Stanton RG & Sprott DA (1958) A family of difference sets. *Canadian Journal of Mathematics* 10:73-77.

Stricker W, Lückerath R, Meier U & Meier W (2003) Temperature measurements in combustion—not only with CARS: a look back at one aspect of the European CARS workshop. *Journal of Raman Spectroscopy* 34(12):922-931.

Sun M-J, Edgar MP, Gibson GM, Sun B, Radwell N, Lamb R & Padgett MJ (2016) Single-pixel three-dimensional imaging with time-based depth resolution. *Nature communications* 7(1):12010.

Tao C, Zhu H, Zhang Y, Luo S, Ling Q, Zhang B, Yu Z, Tao X, Chen D & Li Q (2022) Shortwave infrared single-pixel spectral imaging based on a GSST phase-change metasurface. *Optics Express* 30(19):33697-33707.

Team EE (2020) Updated rapid risk assessment from ECDC on the novel coronavirus disease 2019 (COVID-19) pandemic: increased transmission in the EU/EEA and the UK. *Eurosurveillance* 25(10):2003121.

Teunissen L & Daanen H (2011a) Infrared thermal imaging of the inner canthus of the eye as an estimator of body core temperature. *Journal of medical engineering & technology* 35(3-4):134-138.

Teunissen LPJ & Daanen HAM (2011b) Infrared thermal imaging of the inner canthus of the eye as an estimator of body core temperature. *Journal of Medical Engineering & Technology* 35(3-4):134-138.

Torke PR, Nuster R & Paltauf G (2022) Photoacoustic computational ghost imaging. *Optics Letters* 47(6):1462-1465.

Turner TA (2001) Diagnostic thermography. *Veterinary Clinics of North America: Equine Practice* 17(1):95-114.

Usamentiaga R, Venegas P, Guerediaga J, Vega L, Molleda J & Bulnes FG (2014) Infrared thermography for temperature measurement and non-destructive testing. *Sensors* 14(7):12305-12348.

Vardasca R (2017) The influence of angles and distance on assessing inner-canthi of the eye skin temperature. *Thermology international* 27(4):130-135.

Vatansever F & Hamblin MR (2012) Far infrared radiation (FIR): its biological effects and medical applications. *Photonics Lasers Med* 4:255-266.

Walker HK, Hall WD & Hurst JW (1990) Clinical methods: the history, physical, and laboratory examinations.

Wang S-H, Wei C-W, Jee S-H & Li P-C (2009) Photoacoustic temperature measurements for monitoring of thermal therapy. *Photons Plus Ultrasound: Imaging And Sensing 2009.* SPIE, p 451-461.

Wang Y, Huang K, Fang J, Yan M, Wu E & Zeng H (2023) Mid-infrared single-pixel imaging at the single-photon level. *Nature Communications* 14(1):1073.

Wehlburg CM, Wehlburg JC, Gentry SM & Smith JL (2001) Optimization and characterization of an imaging Hadamard spectrometer. *Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VII.* SPIE, p 506-515.

Wellons M (2007) The Stefan-Boltzmann Law. *Physics Department, The College of Wooster, Wooster, Ohio* 44691:25.

Wong WK, Tan PN, Loo CK & Lim WS (2009) An effective surveillance system using thermal camera. *2009 international conference on signal acquisition and processing.* IEEE, p 13-17.

Yue J, Han J, Li L & Bai L-f (2018) Denoising analysis of spatial pixel multiplex coded spectrometer with Hadamard H-matrix. *Optics Communications* 407:355-360.

Zhang J (2018) Development of a non-contact infrared thermometer. *2017 International Conference Advanced Engineering and Technology Research (AETR 2017).* Atlantis Press, p 308-312.

7 CONCLUSIONS AND FUTURE PERSPECTIVES

In this dissertation, we present two orthogonality-based beam shaping methods: BLIP, utilizing sinusoidal fringe modulation, and SPIRIT, employing cyclic S-matrices encoding. These techniques are specifically applied for 3D surface imaging and LWIR imaging thermometry, respectively.

We have developed BLIP with three configurations: CI-BLIP, TIA-BLIP and MS-BLIP. From the perspective of structured pattern projection, akin to binary defocusing methods, BLIP's image acquisition speed can keep up with the DMD's refreshing rate. At the same time, BLIP shows better performance by always operating at the conjugate plane, which makes BLIP free of depth-dependent blurring effect and image distortion induced by DMD's uneven surface.

Regarding the detection, first, CoaXPress-interfaced high-speed image acquisition and GPU-based image reconstruction in BLIP overcome the limitation of data transmission of existing image acquisition devices. This high-speed PSFPP system has enabled, for the first time, real-time 3D position tracking of a single customer-selectable point at 1 kHz using CI-BLIP (Jiang *et al.*, 2020b). Meanwhile, CI-BLIP has empowered the 3D analysis of the wave propagation, gravity-induced phase mismatch, and asymmetric flapping motion of a non-rigid flag at 1 kHz. These experiments will pave the way for potential future studies of high-speed nonlinear flapping dynamics.

Second, TIA-BLIP has reached high-speed 3D imaging with a larger FOV without sacrificing the camera capturing speed. It achieves kHz-level 3D imaging speeds over an FOV of up to 180×130 mm^2 (corresponding to 1180×860 pixels) in captured images. This technique implements TIA in multi-view 3D PSFPP systems, which allows each camera to capture half of the sequence of phase-shifting fringes. At the same speed, TIA increased the FOV by 3.87 and 2.07 times compared to the same speed of conventional single-view PSFPP and dual-view PSFPP, respectively (Jiang *et al.*, 2020a). TIA also eliminates the redundant capture of fringe patterns in data acquisition. The roles of the main camera and the auxiliary camera are interchangeable. Despite demonstrated only with high-speed cameras, TIA-BLIP is a universal imaging paradigm easily adaptable to other multi-view PSFPP systems. Moreover, TIA reduces the workload for each camera employed in the multi-view systems. The freed capacity is used to enhance the technical specifications in PSFPP. In particular, at a certain frame rate, more pixels on the sensors of the deployed cameras can be used, which increases the imaging FOV. Alternatively, if the FOV is fixed, TIA supports these cameras to have higher frame rates, which

thus increases the 3D imaging speed. Both advantages shed light on implementing TIA-BLIP with an array of cameras to simultaneously accomplish high accuracy and high-speed 3D imaging over a larger FOV. TIA-BLIP has empowered the 3D visualization of glass vibration induced by sound, which could be further explored to function as an imaging accelerometer for vibration monitoring in rotating machinery (Randall, 2004).

Third, we have demonstrated MS-BLIP for robust video-rate multi-scale 3D imaging. MS-BLIP applies an iterative method for distortion compensation, which empowers a video-rate 3D visualization with a measurement volume of up to 1.5 m^3—over two orders of magnitude enhancement compared to existing systems. By adopting a dual-level intensity projection with multi-frequency phase unwrapping, MS-BLIP improves the imaging capability of BLIP by robustly resolving spatially isolated objects with varied reflectance. With an FOV of up to 1.7 m × 1.1 m and a working distance of up to 2.8 m, MS-BLIP captures vase rotation movements and full human-body movements at video rate. Compared to the state-of-the-art competing technique, MS-BLIP preserves high accuracy and high contrast of grayscale fringe patterns at different frequencies. MS-BLIP's use of a dove prism allows customizable FOV orientation and aspect ratio adjustments, making it adaptable for various projection specifications. In addition, MS-BLIP employs a nanosecond pulsed laser as the illumination source. Since the laser pulse width is much shorter than the DMD’s display time, no photons would be wasted during pattern switching, which is otherwise unavoidable for a continuous-wave laser. This advantage also sheds light on implementing MS-BLIP using high-speed cameras (Liang & Wang, 2018; Liu *et al.*, 2019; Liu *et al.*, 2021) to increase 3D imaging speeds while maintaining SNRs.

Finally, We have developed SPIRIT, which uses diagonal aggregation of cyclic S-matrices for compact aggregation of masking patterns, facilitating 1D spatial encoding. Leveraging the 2D smoothness of the encoded data matrix, compressed encoding matrices with data interpolation are used for efficient image reconstruction. Marking its uniqueness, SPIRIT is the first SPI system in the LWIR spectrum, passively mapping temperature by sensing thermal radiation from the human inner canthi. Our results confirmed its accuracy in inner canthi thermometry with statistical analysis revealling no significant difference between the left and right canthus for individual personnel or between male and female subjects. SPIRIT's observations of daytime temperature variations mirrored prior studies and identified a 0.3°C difference between eyeglass wearers and non-wearers. Technically, SPIRIT avoids the spectral sensing limitation imposed by DMDs, which extends SPI’s operation to the LWIR range and enhance safety through passive detection. Unlike traditional 2D thermography, SPIRIT designates FOVs on the inner canthi—a key area indicative

of human's core body temperature. This configuration eliminates the superfluous information, and intrinsically reduces background blending.

The future work will be carried out in the following aspects.

First, we plan to further improve BLIP's imaging speed by implementing a faster DMD with a more powerful laser source. we also plan to implement online feedback to adaptively adjust the intensity of projected patterns (Chen *et al.*, 2018), which would optimize BLIP's dynamic range for different working conditions. Furthermore, to take into account possible measurement inaccuracy induced by laser speckles, we could implement a superluminescent diode and a rotating diffuser in the BLIP system as well as apply a filtering algorithm in image reconstruction (Liu *et al.*, 1999). Moreover, we will continue exploring new applications of BLIP.

Finally, the performance of SPIRIT, as it stands, is bound by the constraints of its core components. However, there's potential for improvement by refining its mask scanning mechanism and the associated software. By integrating a motorized stage that boasts a quicker scanning rate, SPIRIT's imaging speed could be further enhanced (Song *et al.*, 2010). Moreover, the interpolation method in use right now is relatively basic. We believe that exploring more advanced interpolation techniques could further refine the reconstruction quality, without sacrificing speed (Chai *et al.*, 2020a; Chai *et al.*, 2020b).

As a SPI imaging modality for LWIR, SPIRIT holds promises in terms of broad application scope. Beyond its demonstrated capability for thermal imaging of the human inner canthi, SPIRIT's foundational principle can be extended to detect conditions like inflammation, circulatory issues, and even breast cancer by identifying abnormal thermal patterns (Ng & Etehadtavakol, 2017; Nur, 2014), monitor activities even in obscured conditions like fog or smoke (Akula *et al.*, 2011; Wong *et al.*, 2009), identify heat signatures of living beings in rescue operations (Andrea *et al.*, 2018), as wel as characterize thermal property of advanced materials (Kamble *et al.*, 2013; Kölzer *et al.*, 1996).

References

Akula A, Ghosh R & Sardana H (2011) Thermal imaging and its application in defence systems. *AIP conference proceedings.* American Institute of Physics, p 333-335.

Andrea CC, Byron JQ, Jorge PI, Inti TC & Aguilar WG (2018) Geolocation and counting of people with aerial thermal imaging for rescue purposes. *Augmented Reality, Virtual Reality, and Computer Graphics: 5th International Conference, AVR 2018, Otranto, Italy, June 24–27, 2018, Proceedings, Part I 5.* Springer, p 171-182.

Chai X, Gu H, Li F, Duan H, Hu X & Lin K (2020a) Deep learning for irregularly and regularly missing data reconstruction. *Scientific reports* 10(1):3302.

Chai X, Tang G, Wang S, Peng R, Chen W & Li J (2020b) Deep learning for regularly missing data reconstruction. *IEEE Transactions on Geoscience and Remote Sensing* 58(6):4406-4423.

Chen C, Gao N, Wang X & Zhang Z (2018) Adaptive pixel-to-pixel projection intensity adjustment for measuring a shiny surface using orthogonal color fringe pattern projection. *Measurement Science and Technology* 29(5):055203.

Jiang C, Kilcullen P, Lai Y, Ozaki T & Liang J (2020a) High-speed dual-view band-limited illumination profilometry using temporally interlaced acquisition. *Photonics Research* 8(11):1808-1817.

Jiang C, Kilcullen P, Liu X, Gribben J, Boate A, Ozaki T & Liang J (2020b) Real-time high-speed three-dimensional surface imaging using band-limited illumination profilometry with a CoaXPress interface. *Optics Letters* 45(4):964-967.

Kamble VB, Bhat S & Umarji A (2013) Investigating thermal stability of structural defects and its effect on d ferromagnetism in undoped SnO2. *Journal of Applied Physics* 113(24).

Kölzer J, Oesterschulze E & Deboy G (1996) Thermal imaging and measurement techniques for electronic materials and devices. *Microelectronic engineering* 31(1-4):251-270.

Liang J & Wang LV (2018) Single-shot ultrafast optical imaging. *Optica* 5(9):1113-1127.

Liu H, Lu G, Wu S, Yin S & Francis T (1999) Speckle-induced phase error in laser-based phase-shifting projected fringe profilometry. *JOSA A* 16(6):1484-1495.

Liu X, Liu J, Jiang C, Vetrone F & Liang J (2019) Single-shot compressed optical-streaking ultra-high-speed photography. *Optics Letters* 44(6):1387-1390.

Liu X, Skripka A, Lai Y, Jiang C, Liu J, Vetrone F & Liang J (2021) Fast wide-field upconversion luminescence lifetime thermometry enabled by single-shot compressed ultrahigh-speed imaging. *Nature Communications* 12(1):6401.

Ng EY & Etehadtavakol M (2017) *Application of infrared to biomedical sciences.* Springer,

Nur R (2014) *Identification of thermal abnormalities by analysis of abdominal infrared thermal images of neonatal patients*. (Carleton University).

Randall RB (2004) State of the art in monitoring rotating machinery-part 1. *Sound and vibration* 38(3):14-21.

Song Y, Wang J, Yang K, Yin W & Zhu Y (2010) A dual-stage control system for high-speed, ultra-precise linear motion. *The International Journal of Advanced Manufacturing Technology* 48:633-643.

Wong WK, Tan PN, Loo CK & Lim WS (2009) An effective surveillance system using thermal camera. *2009 international conference on signal acquisition and processing.* IEEE, p 13-17.

8 SOMMAIRE

8.1 L’introduction

La précision dans la mise en forme des faisceaux lumineux a été profondément explorée au fil des ans. Les techniques basées sur l'orthogonalité, utilisant des fonctions mathématiques orthogonales telles que les modes Hermite-Gaussien (HG) et Laguerre-Gaussien (LG), sont notables pour la personnalisation des profils lumineux en raison de leur comportement indépendant dans des domaines définis. Ces modes sont essentiels dans les systèmes de communication et l'imagerie super-résolue. Deux fonctions représentatives incluent la fonction sinus et la matrice de Hadamard, toutes deux largement étudiées et offrant diverses applications.

La mise en forme du faisceau sinusoïdal module la lumière en un motif sinusoïdal, essentiel pour la profilométrie de projection de franges (FPP) en imagerie 3D en raison de sa résistance aux changements de réflectivité de surface. Cette technique, avec des dispositifs à micromiroirs numériques (DMD) produisant des motifs sinusoïdaux en niveaux de gris, offre une imagerie haute résolution. Cependant, les systèmes actuels ne peuvent pas produire des motifs de franges sinusoïdaux à des taux de kilohertz, nécessaires pour l'imagerie 3D haute vitesse. Bien qu'il existe des méthodes pour y remédier, comme les techniques de défloutage binaire, elles présentent des défis tels qu'une profondeur de détection limitée et des distorsions d'image potentielles.

La matrice de Hadamard, ainsi que sa dérivation, la S-matrice cyclique, essentielles pour l'imagerie à pixel unique (SPI), possèdent des caractéristiques orthogonales qui garantissent une robuste résistance au bruit et une reconstruction efficace du signal. Dans le SPI, des modulateurs spatiaux de lumière comme les DMD affichent rapidement des matrices de masquage, modulant l'information de la cible pour une récupération rapide des données spatiales. Associé à la détection compressée, le SPI offre une alternative économique pour l'imagerie à travers diverses longueurs d'onde, y compris l'imagerie thermique pour le dépistage de la température humaine.

Les méthodes traditionnelles de dépistage de la température, soit invasives, soit dépendantes de l'éclairage actif, posent des problèmes. En alternative non invasive, la thermographie 2D a été introduite. Cependant, sa vue étendue peut entraîner des inexactitudes en raison de la fusion d'arrière-plan. Un SPI correctement configuré peut résoudre ce problème en se concentrant sur des zones cruciales, comme les canthus internes, réduisant ainsi les coûts et garantissant la précision.

Cependant, le déploiement du SPI dans la mesure de la température corporelle rencontre des défis, en particulier l'inefficacité des DMD dans le domaine infrarouge à longue longueur d'onde (LWIR), où les lectures de température corporelle sont essentielles. Par conséquent, trouver des mécanismes de modulation alternatifs est impératif pour exploiter le potentiel du SPI pour l'imagerie thermique.

8.2 Objectif de la these

Cette thèse propose des solutions aux défis de la mise en forme sinusoïdale des faisceaux dans la FPP et à l'application du SPI en thermométrie d'imagerie LWIR.

Nous introduisons la profilométrie d'illumination à bande limitée (BLIP), une méthode qui crée un motif de frange sinusoïdal en niveaux de gris sur le plan conjugué du DMD à l'aide d'un système d'imagerie 4f et d'un filtre optique passe-bas. Cette approche est polyvalente et peut être intégrée à diverses configurations.

Pour une imagerie de surface 3D en temps réel au niveau kHz, nous présentons le BLIP interfacé CoaXPress (CI-BLIP). Cette technique projette des motifs de franges sinusoïdaux à 5 kHz. Boosté par l'interface CoaXPress, il atteint des taux de transmission de données allant jusqu'à 25 Gbps. Un GPU facilite la reconstruction d'images 3D en ligne, atteignant des reconstructions 3D en temps réel à 1 kHz. Nous prévoyons d'appliquer le CI-BLIP pour étudier la dynamique des fluides des drapeaux non rigides, ce qui pourrait conduire à des découvertes sur la dynamique du battement à haute vitesse.

Pour améliorer davantage la capacité d'imagerie 3D, nous suggérons un BLIP à double vue avec acquisition temporellement entrelacée (TIA). Cette méthode optimise la bande passante d'acquisition de données en capturant des trames alternées avec deux caméras, accélérant ainsi l'acquisition ou élargissant l'activation des pixels. TIA-BLIP vise à améliorer le champ de vision d'imagerie sans compromettre la vitesse. Son application prévue est d'étudier les comportements dynamiques du verre lors des tests de sécurité.

De plus, nous proposons le BLIP Multi-échelle (MS) pour fournir une imagerie 3D polyvalente à différents niveaux de réflectance et distances de travail. Il combine des projections à double intensité avec une projection de frange à multi-fréquence pour mieux reconstruire des informations de surface 3D détaillées. Nous espérons utiliser le MS-BLIP pour la numérisation d'œuvres d'art complexes et la documentation du mouvement humain en 3D.

Enfin, pour une mesure efficace de la température corporelle, nous proposons la thermométrie d'imagerie infrarouge à pixel unique (SPIRIT). Cette approche détecte passivement le rayonnement thermique provenant de zones clés du corps, en particulier les canthus internes. Au lieu d'un DMD, elle utilise un masque physique avec un codage de matrices S-matrices cycliques et une numérisation unidimensionnelle. La configuration comprend une agrégation unique de motifs de masquage centrée sur les canthus internes et des techniques pour améliorer la qualité du signal. Après étalonnage, l'image acquise est convertie en une carte de température. Nous avons pour objectif d'utiliser SPIRIT pour le dépistage de la température corporelle humaine et la surveillance des variations de température.

8.3 Résultats

8.3.1 CI-BLIP

Pour surmonter les limitations liées à l'acquisition et au traitement d'images haute vitesse dans les systèmes FPP existants, nous avons développé le CI-BLIP (Jiang *et al.*, 2020b). En tant que FPP basé sur DMD, le CI-BLIP [Fig. 8.1(a)] présente deux innovations techniques par rapport à l'état de l'art. Premièrement, le BLIP peut projeter un motif de frange en niveaux de gris à partir d'un seul motif binaire DMD au niveau des kHz. Le motif en niveaux de gris est généré sur le plan conjugué du DMD, indépendamment de la période de frange. Ainsi, le BLIP évite l'effet de flou lié à la profondeur et la distorsion d'image dans les techniques précédentes (Song, 2013). Deuxièmement, le CI offre une interface haute vitesse de 25 Gbps pour la diffusion continue de données. Avec un logiciel basé sur GPU, le CI-BLIP permet un suivi de position 3D en temps réel à 1 kHz.

Le flottement des drapeaux dans le vent est un phénomène courant. Les mécanismes qui gouvernent ces phénomènes ont suscité l'intérêt scientifique dans diverses études (Banerjee *et al.*, 2015). En raison de la complexité des dynamiques, étudier l'impact induit par la gravité dans l'interaction fluide-drapeau reste un défi. Motivés par cet obstacle technique, nous avons appliqué le CI-BLIP à l'observation du mouvement 3D de l'interaction fluide-drapeau pour la première fois [voir Fig. 8.1(b) et Visualisation]. Les historiques temporels des déplacements 3D d'un point sélectionné [indiqué sur la Fig. 8.1(b)] montrent une différence de phase [Fig. 8.1(c)], attribuée à l'effet d'affaissement induit par la gravité. La relation de phase $y - z$ [illustrée sur la Fig. 8.1(d)] montre une forme elliptique dans les résultats expérimentaux et ajustés, démontrant que des ondes sinusoidales monofréquence dominent le mouvement. Enfin, les courbes de profondeur de la ligne médiane du drapeau [Fig. 8.1(e)] montrent un mouvement asymétrique, indiquant les

forces inégales sur la surface du drapeau. Cette observation détaillée fournira des preuves directes pour vérifier les modèles théoriques de l'interaction fluide-drapeau. Avec la capacité de diffuser des données en continu, le CI-BLIP pourrait ouvrir la voie à l'imagerie de dynamiques de flux imprévues ou non répétables.

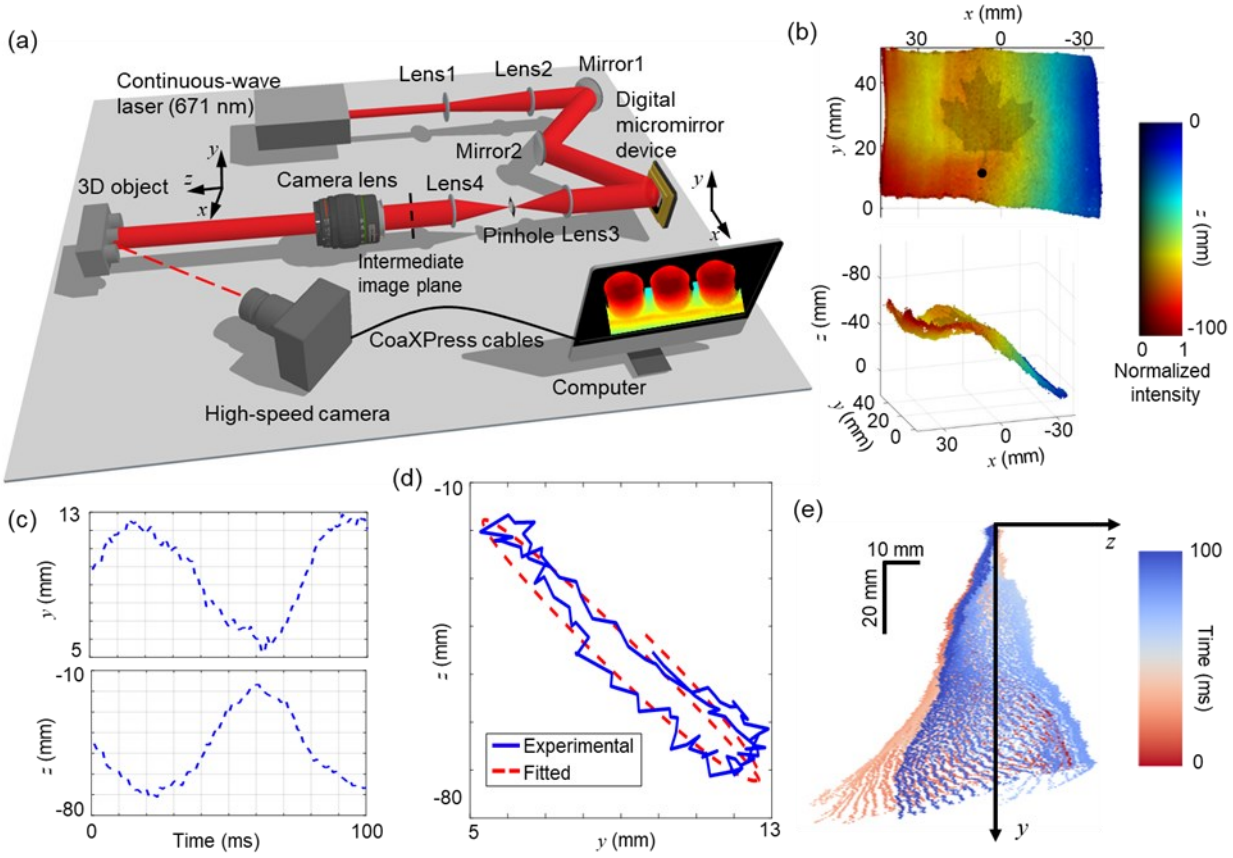

Figure 8.1 Capture du mouvement 3D d'un drapeau flottant en temps réel à l'aide de CI-BLIP.

(a) Schéma du système CI-BLIP. (b) Deux vues en perspective d'une image 3D représentée du drapeau flottant. (c) Évolution des positions 3D d'un point sélectionné marqué en (b). (d) Relation de phase des positions y-z de ce point. (e) Lignes centrales superposées du drapeau au fil du temps.

8.3.2 TIA-BLIP

En tant que système PSFPP à vue unique avec dépliage de phase spatial, le CI-BLIP présente fondamentalement une limitation pour améliorer la vitesse globale d'imagerie 3D en raison d'un motif de calibration supplémentaire inévitable. Pour renforcer davantage la capacité d'imagerie 3D du CI-BLIP, nous avons développé le BLIP à double vue avec acquisition temporellement entrelacée (Jiang *et al.*, 2020a).

Comparé au CI-BLIP, les motifs de frange sont capturés alternativement par deux caméras haute vitesse [Fig. 8.2(a)]. Le TIA permet à chaque caméra de capturer la moitié de la séquence des motifs décalés en phase, réduisant la charge de transfert de données de chaque caméra de 50% [Fig. 8.2(b)]. Cette capacité libérée est utilisée soit pour transférer des données provenant de davantage de pixels sur le capteur de chaque caméra, soit pour soutenir l'utilisation de fréquences d'images plus élevées pour les deux caméras. Exploitant les caractéristiques indiquées dans la condition de correspondance d'intensité, l'algorithme nouvellement développé applique des contraintes géométriques et de phase pour trouver la paire de points correspondants dans les caméras principales et auxiliaires, et guide le dépliage de phase pour extraire l'information de profondeur sans le motif de calibration. De plus, les deux caméras sont placées aussi près que possible du même côté du projecteur, ce qui atténue largement la différence d'intensité et les effets d'ombre. Avec ces avantages, le TIA-BLIP atteint une vitesse d'imagerie 3D allant jusqu'à 1,2 kHz et un champ de vision allant jusqu'à 180 × 130 mm^2. Par rapport au CI-BLIP, le TIA-BLIP augmente le champ de vision de 3,87 fois tout en maintenant une vitesse d'imagerie 3D au niveau des kHz. Le TIA-BLIP a permis une imagerie 3D haute vitesse de la rupture du verre par un marteau [Fig. 8.2(c)], qui pourrait être davantage mise en œuvre pour étudier la caractérisation dynamique du verre dans son interaction avec les forces externes lors d'analyses de tests de sécurité non répétables (Bedon *et al.*, 2019; Haldimann *et al.*, 2008; Ramos *et al.*, 2013).

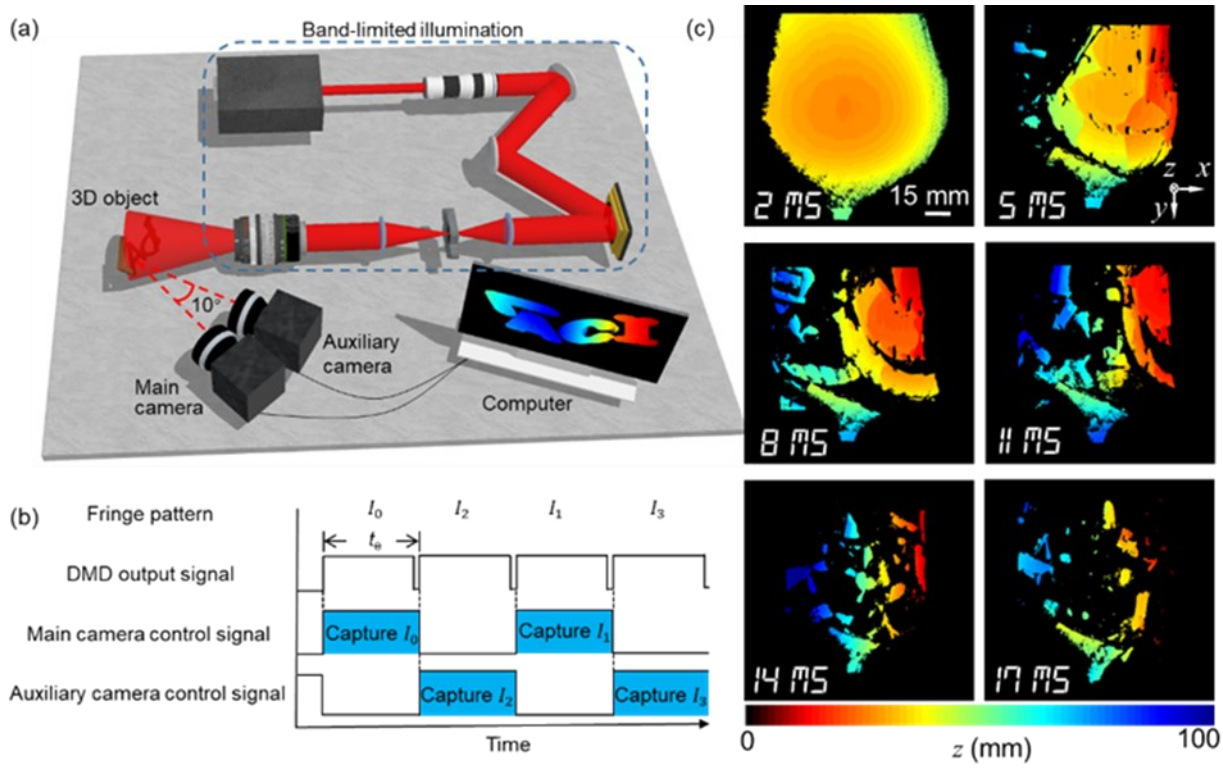

Figure 8.2 Capture du mouvement 3D de la rupture de verre par un marteau à l'aide de TIA-BLIP.

(a) Schéma du système TIA-BLIP. (b) Diagramme de temporisation et séquence d'acquisition. t_e : temps d'exposition de la caméra. (c) Six images 3D reconstruites montrant une tasse en verre brisée par un marteau.

8.3.3 MS-BLIP

En générant le motif sinusoïdal en niveaux de gris sur le plan conjugué du DMD, BLIP élimine la distorsion causée par la surface inégale du DMD et maintient le contraste des motifs de frange projetés. Il est également intrinsèquement compatible avec la projection de franges multi-fréquences, des distances de travail variables et différents champs de vision (FOV). Malgré ces avantages attrayants, BLIP n'a pas encore été démontré pour l'imagerie d'objets 3D qui sont soit spatialement séparés, soit possèdent des surfaces hautement réfléchissantes. Par ailleurs, limité par la puissance des sources lumineuses et les focales des lentilles sur le projecteur et la caméra dans l'appareil existant, le FOV de BLIP n'a pas encore atteint le niveau du mètre carré (m2).

Pour surmonter ces limitations, nous avons développé le BLIP multi-échelle (MS) pour une imagerie robuste des surfaces 3D à la cadence vidéo [Fig. 8.3(a)]. Le MS-BLIP met en œuvre une projection de frange multi-fréquence avec le dépliage de phase associé, ce qui permet une imagerie 3D robuste d'objets spatialement isolés. Par ailleurs, le MS-BLIP adopte une projection

d'intensité à double niveau pour renforcer sa gamme dynamique, ce qui permet de récupérer des informations 3D à partir de surfaces à haute réflectance. De plus, MS-BLIP utilise une méthode itérative pour compenser la distorsion, ce qui améliore la qualité de la reconstruction 3D sur un FOV de niveau m2. De plus, un prisme de Dove est utilisé pour ajuster l'orientation du FOV, aidant ainsi le MS-BLIP à s'adapter à différents scénarios expérimentaux. Enfin, le MS-BLIP utilise un laser pulsé à la nanoseconde comme source d'éclairage. Étant donné que la largeur d'impulsion du laser est bien plus courte que le temps d'affichage du DMD, aucun photon ne serait gaspillé pendant la commutation de motif, ce qui serait autrement inévitable pour un laser à onde continue. Ces supports ont permis au MS-BLIP de réaliser une visualisation 3D dynamique à une vitesse d'imagerie allant jusqu'à 20,8 images par seconde et un volume de mesure allant jusqu'à 1,5 m3 – soit une augmentation de trois ordres de grandeur par rapport aux systèmes BLIP existants. MS-BLIP a été utilisé pour capturer les mouvements complets du corps humain [Fig. 8.3(b)], ce qui pourrait ouvrir la voie à des applications dans l'interaction homme-machine (Zhu *et al.*, 2020).

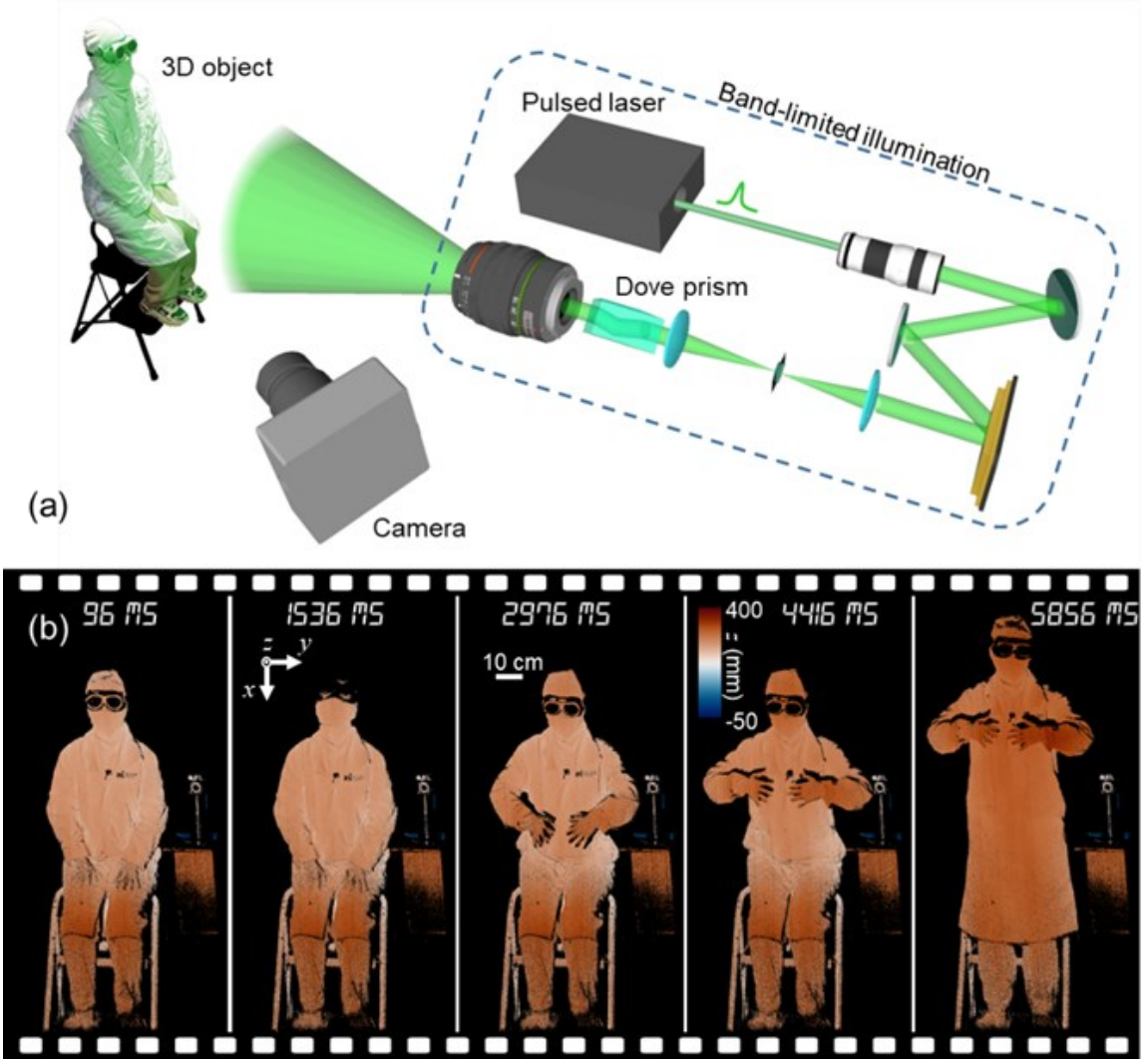

Figure 8.3 Capture du mouvement 3D du corps humain à l'aide de MS-BLIP.

(a) Schéma du système MS-BLIP. (b) Cinq images 3D reconstruites montrant les mouvements du corps humain.

8.3.4 SPIRIT

L'intérêt pour la thermométrie d'imagerie LWIR a augmenté rapidement, en particulier pendant la pandémie. Les méthodes traditionnelles comme les pistolets thermiques peuvent être invasives ou dangereuses. La thermographie 2D est non invasive mais coûteuse et ne cible pas les canthi internes, essentiels pour des lectures de température précises, ce qui peut entraîner des imprécisions dues à l'interférence du fond. Se concentrer sur les canthi internes avec SPI peut optimiser les ressources de thermographie 2D et éviter les interférences. Cependant, utiliser SPI pour LWIR présente des défis car les gammes spectrales dépassent les limites de dispositifs tels que les DMD, ce qui affecte le dépistage efficace de la température.

Face à ce défi, comme le montre la Figure 8.4(a), nous développons la Thermométrie d'Imagerie Infrarouge à Pixel Unique (SPIRIT) avec détection passive dans la gamme spectrale LWIR. SPIRIT utilise des S-matrices cycliques pour les motifs d'encodage en raison de plusieurs avantages. Les S-matrices optimisent l'encodage {0, 1} et ont des inverses facilement calculables. Leur nature cyclique permet la combinaison de plusieurs motifs de masquage et leur agrégation efficace sur des substrats. Cela simplifie également le traitement des données, car seul l'encodage initial est nécessaire pour calculer les motifs et les transformations matricielles pour la reconstruction d'images.

Au lieu d'utiliser un DMD, nous utilisons un masque physique [Fig. 8.4(b)–(c)]. Pour une numérisation optimale des matrices d'encodage, une numérisation unidimensionnelle (1D) est privilégiée. Pour rationaliser l'assemblage des motifs de masquage et réduire les procédures de numérisation, une numérisation diagonale est développée pour agréger les motifs de masquage sur un masque physique [Fig. 8.4(d)]. Profitant de la douceur 2D de la matrice de signal encodé, une procédure de numérisation compressée et une interpolation de données sont mises en œuvre pour une reconstruction d'image efficace [Fig. 8.4(e)]. Ces motifs de masquage sont ensuite restructurés pour s'adapter aux sous-FOVs doubles centrés sur les canthi internes. Ensuite, une autre lentille collecte le rayonnement thermique encodé, le dirigeant vers le détecteur à pixel unique. Pour augmenter le rapport signal/bruit (SNR), un amplificateur de verrouillage et un hacheur optique sont intégrés au système. Avec étalonnage, l'image récupérée subit une transformation en une carte thermique. SPIRIT permet pour la première fois l'imagerie thermique des canthi internes par SPI avec détection passive. Grâce à l'analyse statistique, SPIRIT a effectué un dépistage de la température des sujets humains [Fig. 8.4(f)] et a étudié les fluctuations de la température corporelle et l'impact de l'utilisation prolongée de lunettes.

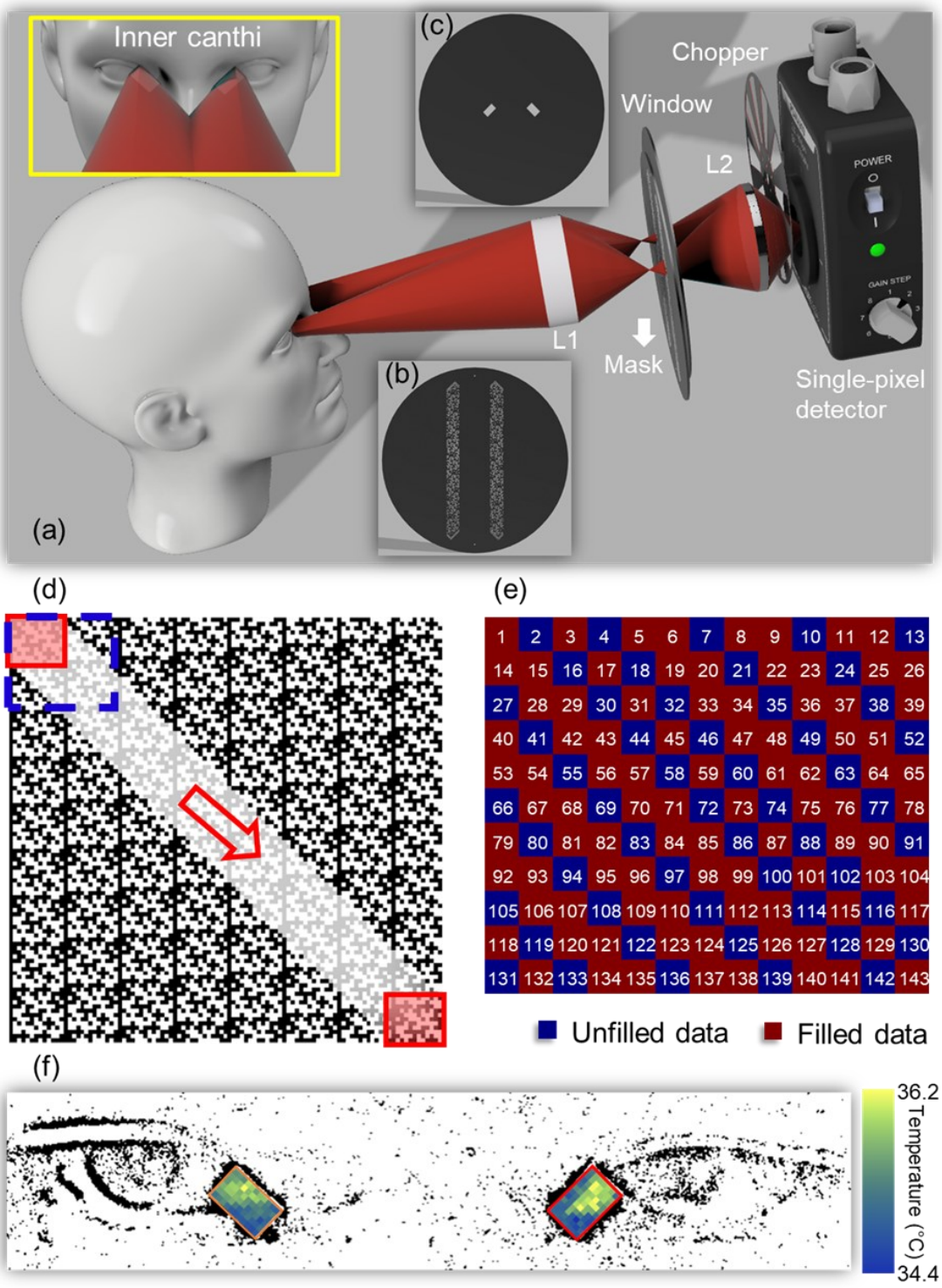

Figure 8.4 Mesure de la température corporelle humaine à l'aide de SPIRIT.

(a) Schéma du système SPIRIT. L'encadré jaune indique les FOVs ciblés des canthi internes humains. (b)-(c) Image du masque (b) et de la fenêtre (c). (d) Illustration du balayage diagonal pour générer les motifs de masquage agrégés avec un grand carrelage des matrices S cycliques redimensionnées. Les boîtes rouges indiquent les motifs de masquage individuels, et la flèche rouge représente la direction du balayage. (e)

Agencement de la matrice de signal encodé de SPIRIT avec un balayage diagonal compressé. (f) Reconstruction représentative de la température des canthi internes.

8.4 Conclusions et perspectives d'avenir

Cette dissertation introduit deux méthodes de façonnage de faisceau, BLIP et SPIRIT, pour l'imagerie 3D et la thermométrie infrarouge.

BLIP, exploitant la modulation de frange sinusoidale, propose trois configurations : CI-BLIP, TIA-BLIP et MS-BLIP. Ces variantes abordent des défis tels que les limites de transmission de données et permettent un suivi en temps réel de la position 3D à des vitesses alignées sur le taux de rafraîchissement du DMD. Par exemple, CI-BLIP permet une analyse à haute vitesse de mouvements dynamiques comme le flottement d'un drapeau.

TIA-BLIP élargit le champ de vision (FOV) sans compromettre la vitesse de la caméra. Cette méthode double efficacement le taux de capture en ayant deux caméras partageant la séquence de franges en décalage de phase. En conséquence, elle augmente le FOV, supporte des fréquences d'images plus élevées et réduit la redondance dans la capture de données. Ses applications pratiques comprennent la visualisation des vibrations de verre induites par le son.

MS-BLIP se concentre sur l'imagerie 3D multi-échelle robuste à des vitesses vidéo. En utilisant une méthode unique de compensation de distorsion itérative, elle atteint des volumes de mesure expansifs et peut distinguer des objets à réflectance variée. Sa capacité à ajuster l'orientation et le rapport d'aspect du FOV, combinée à l'utilisation efficace d'un laser pulsé à la nanoseconde, promet adaptabilité et efficacité dans divers scénarios d'imagerie.

D'autre part, SPIRIT utilise des matrices S cycliques avec balayage diagonal pour le codage. Cette méthode se distingue comme la première à fonctionner dans le spectre infrarouge à longue onde avec détection passive, ciblant la mesure de la température corporelle humaine. Le système se concentre de manière unique sur les canthus internes, garantissant des lectures précises sans interférence d'arrière-plan superflue. Nos tests vérifient sa précision, mettant en lumière les variations de température diurnes et soulignant même de légères différences entre les porteurs de lunettes et les non-porteurs. La conception de SPIRIT surmonte les limitations spectrales traditionnelles, en se concentrant sur les zones clés indiquant la température corporelle centrale, minimisant ainsi les erreurs et les redondances.

Des efforts futurs se concentreront sur plusieurs améliorations :

Améliorations de BLIP: Nous visons à augmenter la vitesse d'imagerie de BLIP en utilisant un DMD plus rapide et un laser plus puissant. Nous explorons également les diverses

applications de BLIP. Reconnaissant le potentiel de BLIP dans la chirurgie assistée par vision 3D, nous collaborons pour intégrer BLIP dans les chirurgies assistées par robot. Nous utilisons un laser proche infrarouge, avec un design compact pour l'adaptabilité. Après l'installation et la calibration, le système sera testé sur des mesures 3D de tissus en temps réel et intégré à un système chirurgical robotisé pour un retour en temps réel.

Sécurité & Diversification de BLIP: Compte tenu des préoccupations de sécurité liées aux lasers, nous faisons évoluer BLIP vers les LEDs. LED-BLIP, doté d'un module de projection compact avec trois LEDs RGB, utilisera un séparateur de faisceau, des optiques de relais et un prisme TIR pour diriger la lumière vers le DMD. Le système filtrera le bruit à haute fréquence et dispersera les spectres. En utilisant des LEDs pour projeter des motifs de franges à déphasage capturés par une caméra couleur, BLIP reconstruira les données 3D. L'objectif est de garantir l'adaptabilité de LED-BLIP à travers des applications telles que l'inspection industrielle, l'archéologie et l'interaction homme-machine.

Optimisations de SPIRIT: Les performances de SPIRIT peuvent être améliorées en perfectionnant son mécanisme de balayage de masque et son logiciel. L'intégration d'une scène motorisée plus rapide peut augmenter sa vitesse d'imagerie. De plus, l'adoption de méthodes d'interpolation avancées peut affiner la qualité de la reconstruction de l'image sans compromettre la vitesse.

Élargissement des Applications de SPIRIT: SPIRIT, une modalité d'imagerie thermique pour LWIR, a de vastes applications potentielles. Au-delà de l'imagerie thermique du canthus interne de l'homme, elle peut détecter des affections telles que l'inflammation ou même le cancer du sein à travers des motifs thermiques anormaux, opérer dans des conditions obscurcies comme le brouillard, identifier les signatures thermiques d'êtres vivants lors de missions de sauvetage et analyser les propriétés thermiques de matériaux avancés.

8.5 Références

Banerjee S, Connell BS & Yue DK (2015) Three-dimensional effects on flag flapping dynamics. *Journal of Fluid Mechanics* 783:103-136.

Bedon C, Fasan M & Amadio C (2019) Vibration analysis and dynamic characterization of structural glass elements with different restraints based on Operational Modal Analysis. *Buildings* 9(1):13.

Haldimann M, Luible A & Overend M (2008) *Structural use of glass.* Iabse,

Jiang C, Kilcullen P, Lai Y, Ozaki T & Liang J (2020a) High-speed dual-view band-limited illumination profilometry using temporally interlaced acquisition. *Photon. Res.* 8(11):1808-1817.

Jiang C, Kilcullen P, Liu X, Gribben J, Boate A, Ozaki T & Liang J (2020b) Real-time high-speed three-dimensional surface imaging using band-limited illumination profilometry with a CoaXPress interface. *Optics Letters* 45(4):964-967.

Ramos A, Pelayo F, Lamela M, Canteli AF, Huerta C & Acios A (2013) Evaluation of damping properties of structural glass panes under impact loading. *COST Action TU0905 Mid-Term Conference on Structural Glass; Belis, J., Louter, C., Mocibob, D., Eds.*

Song Z (2013) *Handbook of 3D machine vision : optical metrology and imaging.*

Zhu L, Rematas K, Curless B, Seitz SM & Kemelmacher-Shlizerman I (2020) Reconstructing nba players. *European Conference on Computer Vision.* Springer, p 177-194.