

以超穎介面和光子晶體面射型雷射的結構光實現單眼深度感知和 臉部識別

Metasurface- and PCSEL-Based Structured Light for Monocular Depth Perception and Facial Recognition

徐閔正、張家勳、洪瑜亨、郭浩中、黃耀緯

Wen-Cheng Hsu, Chia-Hsun Chang, Yu-Heng Hong, Hao-Chung Kuo, and Yao-Wei Huang

國立陽明交通大學電機學院光電工程學系

摘要

這裡介紹的新型深度感測系統透過採用超穎介面和光子晶體面射型雷射在單眼深度感測中實現高效的臉部識別，徹底改變了結構光技術。與依賴繞射光學元件和準直透鏡的傳統點投影儀不同，我們的系統從三百微米尺寸的超穎介面投射約四萬五千個紅外線點，比商用基於繞射光學元件的點陣投影儀更多（1.43 倍）且尺寸更小（233 倍）。此系統的測量視場為 158° ，點採樣角為 0.611° ，無透鏡且重量輕，且功耗比垂直腔表面發射雷射陣列更低，功率降低 5–10 倍。這項創新利用基於砷化鎵基板的超穎介面和簡化的光學架構，不僅解決了傳統點陣投影儀的缺點，而且還為微型化整合至可穿戴設備中開闢了途徑，在尺寸、功效和廣泛採用的潛力方面提供了顯著的優勢。

關鍵字：超穎介面、光子晶體面射型雷射、結構光、深度感測

Abstract

The novel depth-sensing system presented here revolutionizes structured light (SL) technology by employing metasurfaces and photonic crystal surface-emitting lasers (PCSELs) for efficient facial recognition in monocular depth sensing. Unlike conventional dot projectors relying on diffractive optical elements (DOEs) and collimators, our system projects approximately 45,700 infrared dots from a compact 297- μm -dimension metasurface, drastically generate more spots (1.43 times) and has smaller area (233 times) than the DOE-based dot projector in an iPhone. With a measured field-of-view (FOV) of 158° and a 0.611° dot sampling angle, the system is lens-free and lightweight and boasts lower power consumption than vertical-cavity surface emitting laser (VCSEL) arrays, resulting in a 5–10 times reduction in power. Utilizing a GaAs-based metasurface and a simplified optical architecture, this innovation not only addresses the drawbacks of traditional SL depth sensing but also opens avenues for compact integration into wearable devices, offering remarkable advantages in size, power efficiency, and potential for widespread adoption.

Keywords : metasurfaces, photonic crystal surface-emitting laser, structured light, depth sensing

一、 簡介

拜半導體製造的技術快速演進，近幾年科技的發展可謂日新月異，而早期眾多在光學桌上示範的技術，也因此慢慢進入大眾的生活視野之中。整個技術藉由矽光子將同調通訊系統微縮至一個硬幣大小開始，到近年熱門的光達（light detection and ranging, Lidar）與共同封裝光學元件（Co-packaged Optics, CPO）都是將現有的技術與光學架構進行大幅度的微縮。而除了矽光子之外，超穎介面的概念也在最近被提出來，它使傳統光學元件的厚度能夠被大幅縮減並使元件結構平坦化。這樣的改變將使得被動光學元件更容易被整合進各種微系統中且藉由半導體技術大量生產製造。矽光子的核心是利進行光傳播路徑的用波導的概念操控，而超穎介面則是利用空間中相位的分布來操控光的波前的行為，藉此達到不同光學元件功能的效果。而藉由這些技術的提出，深度感測的應用也開始落地普及化如飛時測距（Time of Flight, TOF）、結構光（structured light, SL）[1]。

飛時測距利用射出脈衝光碰到待測物反彈後被偵測器偵測到訊號所花費的時間進行距離的估算，該方法常用於 Lidar 系統之中。它的優點是架構及計算簡單，然而近距離精密度較低因此常用於 5 公尺以外的感測場域。結構光則是投出離散的光點圖案，當照射至待測物時會使某些區域產生變形。因此我們可以利用攝影機拍攝變形後的光點圖案來與未變形之前的圖案做計算，透過在兩個圖像之間的每個區域計算其相關係數，可以知道變形區域的位移量並藉由系統的三角關係推算出深度的變化。雖然結構光應用距離在 5 公尺內，但因其有較高準確度，而被蘋果從微軟的體感遊戲系統 Kinect 收購並整合至手機的人臉辨識系統中。

從相關專利可以知道結構光需要一個點陣投影儀與攝影鏡頭組成，而點陣投影儀的微縮需要藉助垂直共振腔型面射型雷射陣列（vertical-cavity surface emitting laser array, VCSEL array）、準直透鏡組、摺疊波導與繞射光學元件（DOE）整合 [2]。這樣的光學架構，透過透鏡組將垂直共振腔型面射型雷射陣列上發光的影像由摺疊波導輸入繞射光學元件後產生多階繞射，使垂直共振腔型面射型雷射陣列上的發光影像重複拼接後形成離散的偽隨機光點圖案。在這種光學架構下，過去的文獻大多是針對繞射光學元件做更多高階繞射階數的設計與優化，或者是跟單一孔徑的垂直共振腔型面射型雷射做積體整合 [3, 4]。然而都難以達成整體系統上的微縮或是實際的應用價值。

實際上，我們可以利用全像投影的原理更簡單的達到這一光點圖案投射的目的並大幅簡化現有的光學架構，使點陣投影儀能夠被晶片化製造生產而不透過多次的封裝組合，如圖1。根據 GS（Gerchberg-Saxton）演算法可以知道所需投射圖案的相位分布，而藉由不同的柱子粗細或是旋轉角度可以達到不同程度的相位，我們可以把二者結合成具有投影出點陣圖案的超穎介面 [5, 6]。

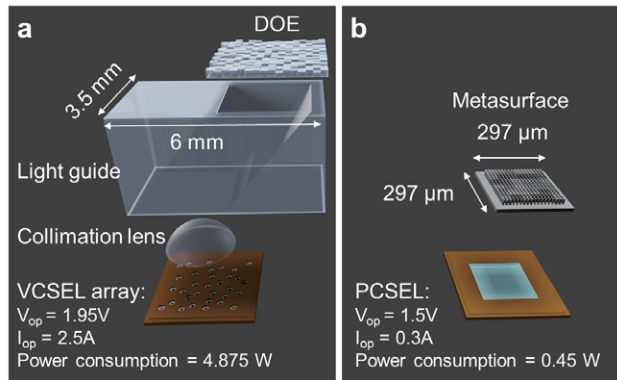


圖1. 光學架構比較。(a) iPhone 所使用點陣投影儀之光學架構。(b) 我們所提出之全像點陣投影光學架構。

該設計在利用 GS 演算法時，通常都是假設光源為平面波並將其強度與空間中的相位分布進行卷積的迭代運算，因此在實際的相關實驗與文獻中通常都需要將雷射光源通過準直透鏡形成發散角極小的平面光光源在打在超穎介面的樣品上。光子晶體面射型雷射 (photonic crystal surface-emitting lasers, PCSEL) 使用週期排列的均勻空氣洞形成共振腔，它利用光子晶體能帶上的共振，讓此種半導體雷射具有大面積發光、高功率與小於 1 度的超低發散角等特性 [7]。我們實驗用的 PCSEL 特性如圖2 所示 [5,6]。因此我們使用此種半導體雷射而不是 VCSEL array 作為我們系統的光源，不需要透鏡的同時大幅降低點陣投影儀的複雜架構。也因為高功率的關係，我們所提出的整個系統的功耗將比現有的架構還要來的更低，真正的實現了輕薄小體積與節能低功耗的目的。

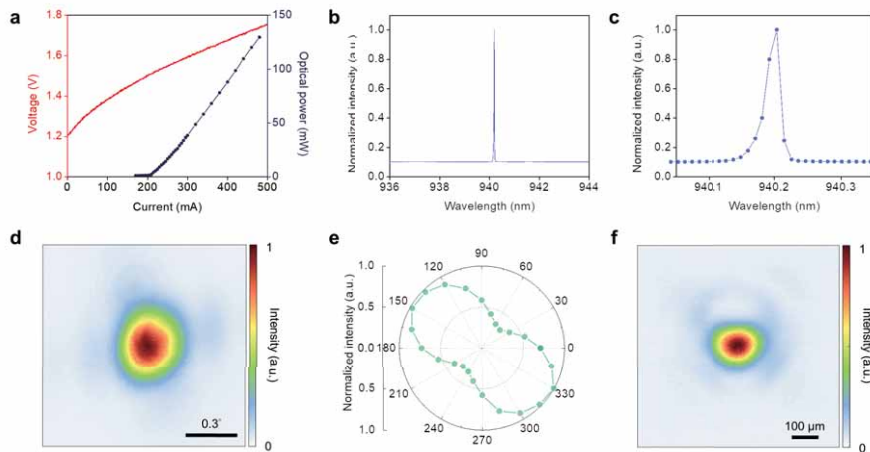


圖2. 光子晶體面射型雷射之相關光學特性量測結果 [5, 6]。(a) 光輸出功率、電流與電壓關係圖。(b) 雷射光譜。(c) 雷射單模的放大光譜。(d) 雷射光發散角。(e) 雷射極化特性。(f) 照射至樣品時的雷射光大小。

二、超穎介面點陣投影儀設計

在這節將提到點陣的設計與結構光深度感測的實驗架構。根據結構光相關文獻與蘋果公司專利所顯示的模組架構，結構光至少須包含一組點陣投影儀與攝影機才能進行深度相關的量測 [8,9]。因此，我們首先準備一 640×480 像素的攝影機並且計畫以

PCSEL 作為一平面光源搭配可投影出具有上萬點光點分布的全像圖的超穎介面合做為點陣投影儀。其中，實驗架構中的投影儀與攝影機相距 4 公分，在後續實驗中使用 iPhone XR 的結構光圖案做實驗對比時，我們直接將手機切換至相機的前鏡頭與人像模式，同時保持與攝影機相同的距離，實際架構如下圖 3 所示。

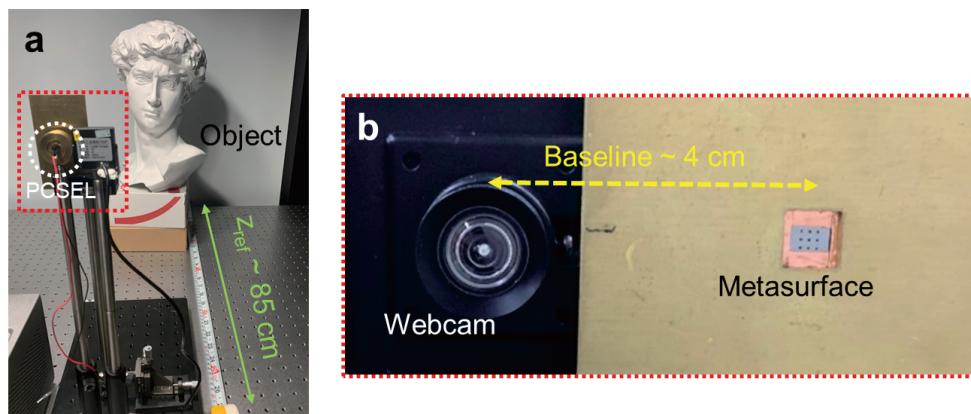


圖3. 實驗架構。(a) 深度感測實驗中的結構光架構，與待測物之情境。(b) 單目結構光正面架構圖。

超穎介面的設計首先以建立超穎原子的資料庫開始。考慮到光子晶體面射型雷射的極化特性，在這邊我們選擇正方形的超穎原子以避免極化方向造成的效率問題。不同大小的方形超穎原子會對應到不同的相位與穿透效率，在此我們盡量以 2π 光學相位的覆蓋率與盡可能高的穿透效率做為大小範圍的尋找目標。最終我們可以獲得在入射光波長為 940 奈米時，週期大小 330 nm，方形邊長 120 至 205 nm，柱子高度 800 nm 可有 2π 光學相位的覆蓋率與 65~85% 之間的穿透效率。

接著我們以 GS 演算法開始設計我們所需投影出點陣全像圖的相位分布。該方法將光源的強度分布與結構的相位分布做迭代的卷積運算，並根據與目標的誤差更新為新的相位分布。相當於保持相同光源下看不同相位時的遠場情況，並根據遠場結果代換近場相位分布，以達到所設計之遠場結果。特別的是，我們可以直接使用快速傅立葉轉換與反轉換做為運算核心，在誤差收斂後結束迭代運算。

此時我們的遠場結果將是會奠基於 Fraunhofer 繞射條件下的近似結果，意為投射出的全像圖，需要大於一個距離之後才能夠看到我們設計好的全像圖。該距離可由下列公式獲得： $L \gg A/\lambda$ ，其中 L 為繞射距離， A 為結構面積， λ 為光源波長。

我們根據 GS 演算法所獲得相位分布，與超穎原子資料庫比對，在超穎介面不同位置放置不同粗細的柱子，以對應所需的相位分布。

除此之外，在設計點陣圖時，需要考慮到雷射光源的發散角在點陣圖中光點與光點之間距離的影響。即雷射發散角越大時，光點與光點之間需要越大的距離才能彼此不受影響被分辨出來。在這裡所使用之光子晶體面射型雷射，具有 0.300 度的發散角與約 300 μm 的發光孔徑，而全像圖中的每個相位由 3×3 個單位晶胞組成，這將使的全像圖中每個像素具有約 0.189 度的解析度。因此光點與光點之間的距離至少要大於 3 個像素，才能有效被分辨出個別獨立的光點。上述關係如下圖 4 所示。

在完成設計之後，我們將超穎介面的結構利用半導體製程（如薄膜沉積、微影、乾蝕刻等）製作於砷化鎵基板上。樣品如圖 5 所示，其中可以藉由 SEM 看到製程完成之後的柱狀結構有達到我們所設計的尺寸，而與傳統點陣投影儀的面積也有 233 倍的面積差距。

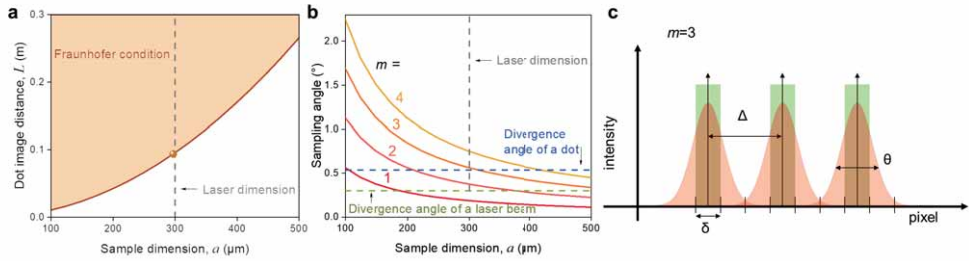


圖4. (a) Fraunhofer 繞射的成像距離。(b) 結構光光點間距與樣品大小及光點大小的曲線關係。(c) 雷射光發散角與可被分辨的結構光光點間距之關係示意圖。

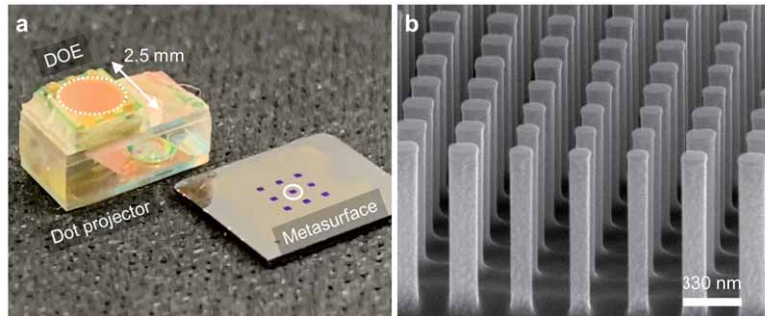


圖5. (a) 實際樣品與 iPhone 點陣投影儀之尺寸比較圖，左上為 iPhone 之點陣投影儀光學模組，右下為實際樣品在晶片上。共包含 9 種設計，個別都可以射出不同結構光樣式。(b) 實際樣品 SEM 圖。

三、深度感知

接著我們可以將光子晶體面射型雷射立於超穎介面樣品後方約 1 公分處，當雷射光源完整照射置樣品上時，可以看到設計好的全像圖被投影至遠處的屏幕上。我們預先設計三種結構光樣品包含較稀疏的規則光點陣列，較密集的規則光點陣列與較為密集的隨機分布光點陣列，個對應到樣品1、2 與 3。此時可以將該距離下的結構光圖像紀錄為參考圖像，然後再準備不同物品放置到攝影機畫面中紀錄為物件圖像。藉由計算參考圖像中的每個特徵區塊與在物件圖像中對應的特徵區塊平移量，我們可以獲得對應的像差圖 (disparity map)。最後根據這個像差圖與系統的近似三角關係，可以換算出二維深度圖與三維點雲圖。整個計算過程與步驟如下圖 6。在這邊我們用與人臉尺寸差不多大小的大衛像作為感測目標，且在不同位置皆可獲得對應的點雲圖。確認該實驗方法與架構沒問題之後，我們可以將 iPhone XR 開啟人臉模式以打出結構光，並在同樣的量測條件與方式下進行物件深度或是距離的探測計算。

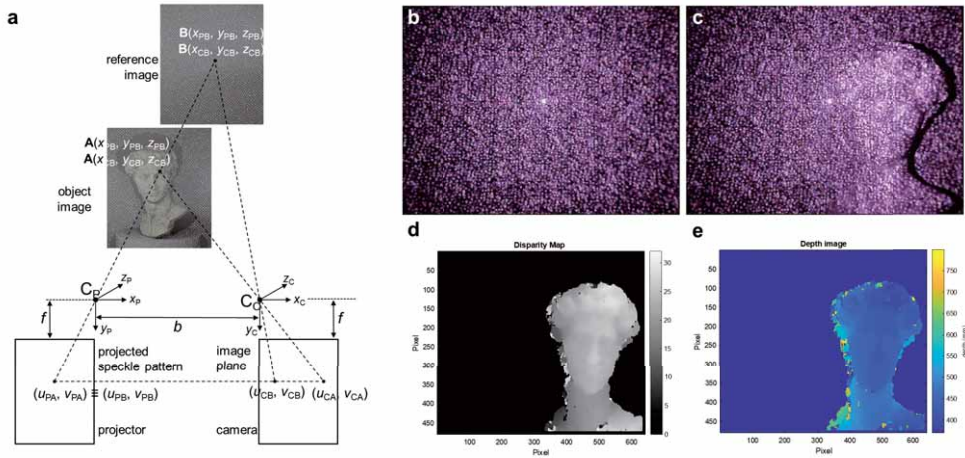


圖6. (a) 實驗中所使用的結構光計算深度方式與架構。(b) 於 85 cm 處的結構光影像，做為參考圖像。(c) 放上待測物大衛像之後的結構光影像，標記為物件圖像。(d) 藉由參考圖像與物件圖像來計算獲得的像差圖。(e) 藉由像差圖與系統中的三角關係所獲得的待測物深度資訊圖。

後續的實驗中，我們將針對設計好不同的結構光樣式與 iPhone XR 的點陣投影儀，對一小平板在不同距離下的實際位置與量測深度做比較，如圖7。我們總共設計了三種不同的結構光如圖7 (a-c)。精確度的比較實驗可以整理出如圖7 (e)，黑藍綠紅點分別對應了樣品 1、2、3 與 iPhone XR 對不同位置平板所量測到的距離平均值，陰影區域對應的是標準差，而黑色虛線則是代表沒有誤差的距離。在該圖中可以看出 30 至 80 公分的時候，所有結構光都可以落在虛線上，30 公分以內時開始出現較大的偏離且誤差開始增加。這是因為樣品 1、2 是規則光點陣列，再進行後續深度計算的演算法上容易出現誤判的狀況，而樣品 1 光點間的距離又比 2 更為疏離。樣品 3 則是與樣品 2 有一樣的光點間距，但是排列上呈現隨機分布的光點，因此在後續深度運算的演算法時，特徵kernel較不容易因為重複出現而誤判。而我們所設計的光點大小與 iPhone XR 所投射的光點大小都在 0.65 度左右，因此有相似的空間解析度與差不多的深度精確度結果。

在精確度的比較實驗中，可知樣品 3 與 iPhone XR 的結構光有相當的效果。因此我們在後續人臉辨識實驗裡，只做樣品 3 與 iPhone XR 的比較。

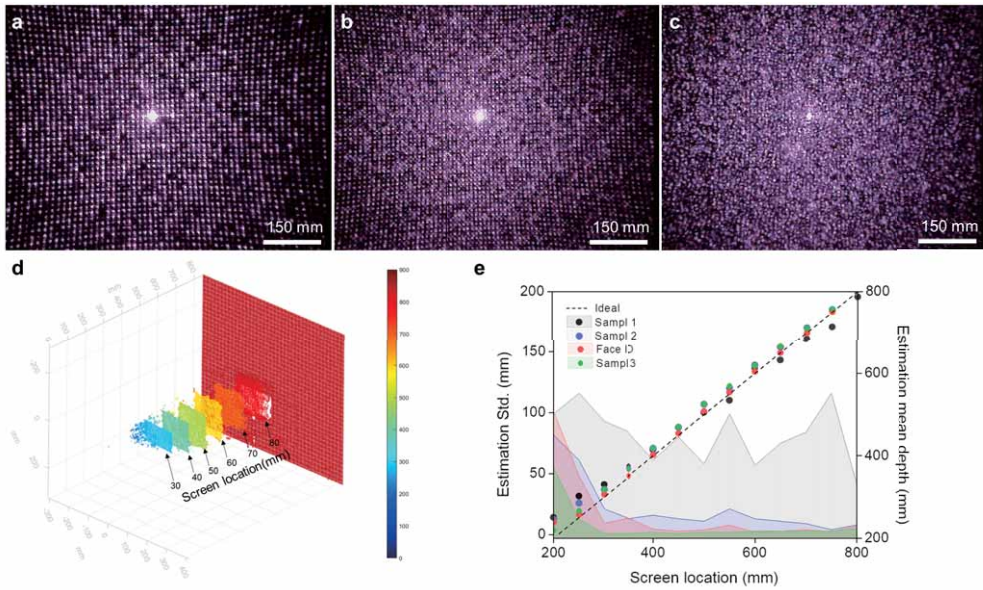


圖7. (a) 樣品一的結構光 (較疏的陣列光點)。 (b) 樣品二的結構光 (較密的陣列光點)。 (c) 樣品三的结构光 (較密的隨機光點)。 (d) 還原不同距離下的 3D 平板點雲圖。 (e) 所有設計樣品與 iPhone XR 的點陣投影儀, 在不同平板距離的平均量測距離與誤差範圍。

四、人臉辨識

我們的人臉辨識系統是基於開源的 OpenCV 人臉識別系統, 我們以此輔助與預先訓練總共包含了 6 種不同的人臉。其中大衛像的訓練集是收集網路上商品照片的素材作訓練, 目的在於確認系統能否將 3D 點雲圖重建後的三維模型圖像辨認出人臉特徵並被分類至正確的類別。

最後的實驗結果如圖8 顯示, 由我們所提出的點陣投影儀光學架構與目前主流的商业點陣投影儀都可以被辨認出人臉且歸類至大衛像, 而在 OpenCV 裡也提供了信賴度分數 (confidence score) 來判斷辨識結果的可信度可以發現二者也是都在約 50 分。

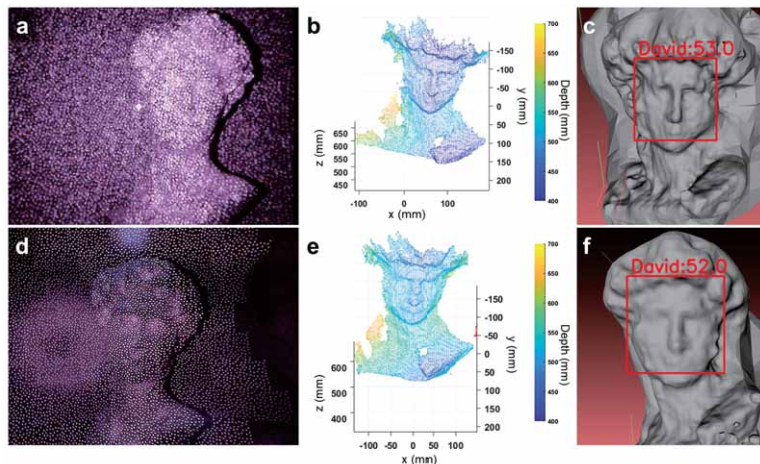


圖8. 重建模型的人臉辨識結果比較。 (a-c) 我們架構的結構光 (a)、對應獲得的點雲圖 (b), 與重建 3D 模型之後的人臉辨識結果 (c)。 (d-f) iPhone XR 的結構光 (d)、對應獲得的點雲圖 (e), 與重建 3D 模型之後的人臉辨識結果 (f)。

在 OpenCV 中，信賴度分數數值越低代表與訓練集中的目標越相似，且在 80 分以下都是屬於可信賴區間。因此代表了我們所提出的光學架構在大幅微縮且簡化的同時還能提供一樣的功能。而使用光子晶體面射型雷射的優勢除了在實驗室有更低的功耗之外，在後續主被動元件的單晶片整合製程中，無須透鏡的相位導致超穎介面與雷射光源之間沒有三維對位問題，這將使元件良率能夠大幅增加進而達到商品量產的目的。

在最後我們列出點陣投影儀的一些關鍵參數如表1，並根據參數的正反面影響性來定義 FoM，代表了綜合了所有參數的整體比較。

表 1. 結構光相關文獻與商品之關鍵參數整理。

	Face ID	Ref. [3]	Ref. [4]	Our sample 2
Diffraction order	-	±47	±1	±1
Dots number, N (#)	32K	10K	9	45.7K
Field of view, FOV (°)	74	180	20	158
Area, A (μm^2)	$\pi \times 1250^2$	500^2	$\pi \times 100^2$	297^2
Material	-	a-Si	GaAs	GaAs
Operation wavelength, λ (nm)	940	633	980	940
Sampling angle, Δ (°)	0.642	1.757	10	0.611
FoM	0.751	4.10	5.7×10^{-4}	134

$$FoM = \frac{N \times FOV}{A \times \Delta}$$

表中可以看到與相關文獻與 iPhone XR 相比，我們的樣品具有最高的 FoM。這是因為我們能夠在 $297^2 \mu\text{m}^2$ 的樣品面積投射出 45,700 個光點，且視場達到 158 度的同時還能有 0.611 度的光點取樣角度 (Sampling angle)。

五、 結論

在這篇文章中，我們提出了基於全像術的點陣投影儀架構，該架構比現行商業架構更簡單，同時利用了光子晶體面射型雷射與超穎介面的優勢讓元件面積縮小 233 倍。在縮小簡化的同時，我們還能在 158 度的視場投射出 45,700 個光點，並且藉由點雲圖展示出探測物件的形貌。我們也是第一個藉由 OpenCV 的人臉辨識系統來進行三維模型重建後的效果判定，結果顯示出我們的架構與 iPhone XR 的點陣投影儀有相當的效果，且結果在可信賴區間之內。而使用砷化鎵基板進行超穎元件的製作，也讓日後可以藉由覆晶製程完成主被動元件的單晶片整合。

六、 致謝

我們的工作獲得國家科學及技術委員會 (計畫編號：110-2622-8-A49-008-SB、111-2622-8-A49-021-SB 與 110-2112-M-A49-034-MY3) 的補助。我們感謝台灣教育部玉山青年學者計畫的支持也同時感謝鴻海研究院半導體所的技术支援。這項工作有部分

在國立陽明交通大學奈米中心、奈米科技中心完成。我們也要感謝蔡定平教授、盧廷昌教授、余沛慈教授和國家實驗研究院台灣半導體研究中心（TSRI）進行相關技術討論。

七、參考資料

- [1] Yang, S.-P.; Seo, Y.-H.; Kim, J.-B.; Kim, H.; Jeong, K.-H. Optical MEMS Devices for Compact 3D Surface Imaging Cameras. *Micro Nano Syst. Lett.* 2019, 7(1), 8.
- [2] Mor, Z. Integrated Structured-Light Projector Comprising Light Emitting Elements on a Substrate. US9825425B2, November 21, 2017.
- [3] Kim, G.; Kim, Y.; Yun, J.; Moon, S.-W.; Kim, S.; Kim, J.; Park, J.; Badloe, T.; Kim, I.; Rho, J. Metasurface-Driven Full-Space Structured Light for Three-Dimensional Imaging. *Nat. Commun.* 2022, 13(1), 5920.
- [4] Wang, Q.; Ni, P.; Xie, Y.; Kan, Q.; Chen, P.; Fu, P.; Deng, J.; Jin, T.; Chen, H.; Lee, H. W. H.; Xu, C.; Genevet, P. On-Chip Generation of Structured Light Based on Metasurface Optoelectronic Integration. *Laser Photonics Rev.* 2021, 15(3), 2000385.
- [5] Hsu, W.-C.; Chang, C.-H.; Hong, Y.-H.; Kuo, H.-C.; Huang, Y.-W. Metasurface- and PCSEL-Based Structured Light for Monocular Depth Perception and Facial Recognition. *Nano Letters* 2024, 24(5), 1808-1815.
- [6] Hsu, W.-C.; Chang, C.-H.; Hong, Y.-H.; Kuo, H.-C.; Huang, Y.-W. Compact Structured Light Generation Based on Meta-Hologram PCSEL Integration. *Discovery Nano* 2023, 18(1), 87.
- [7] Hirose, K.; Liang, Y.; Kurosaka, Y.; Watanabe, A.; Sugiyama, T.; Noda, S. Watt-Class High-Power, High-Beam-Quality Photonic Crystal Lasers. *Nat. Photonics* 2014, 8(5), 406-411.
- [8] Wang, G.; Yin, X.; Pei, X.; Shi, C. Depth Estimation for Speckle Projection System Using Progressive Reliable Points Growing Matching. *Appl. Opt.* 2013, 52(3), 516.
- [9] Donlic, M.; Petkovic, T.; Pribanic, T. On Tablet 3D Structured Light Reconstruction and Registration. *IEEE Int. Conf. Comput. Vision Workshops* 2017, 2462-2471.