

壹、前言

一、研究動機

隨著元宇宙的話題持續發燒，同時也帶動了相關產業技術的積極投入。目前常見的VR裝置操控方式還是以手持控制器或手勢辨識為主，然而實際生活中若能像電影《鋼鐵人》（Iron Man）中使用眼球來操控螢幕選項，進而操控VR中呈現事物的科幻情節，已隨著「眼球追蹤（Eye Tracking）」之技術開發而成為可能。為了能更多了解這項工程技術的原理及可能的應用，我們開始了相關的資料檢索與分析，並特別關注近兩、三年國際會議上的研究題目及未來發展方向。

二、研究目的

- （一）瞭解眼球追蹤技術的歷史與原理。
- （二）探討眼球追蹤技術的研究現況與方向。
- （三）探索在元宇宙的虛實整合世界中如何應用眼球追蹤技術。

貳、文獻探討

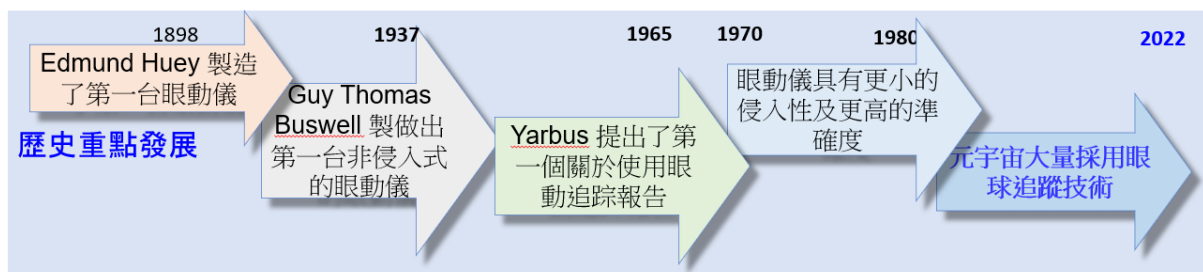
一、歷史重點發展

1898 年，Edmund Huey 使用了具有人工瞳孔接觸鏡，進而製造出了第一台眼動儀。

1937 年，巴斯韋爾（Guy Thomas Buswell）製做出第一台非侵入式的眼動儀，此眼動儀也首次在膠片上錄製了眼球運動。他使用了膠捲記錄從眼睛反射回來的光，而這光的來源則是經過設計從特定方向打到眼睛表面，在當時完成了一組有200名參與者所進行之文字閱讀和瀏覽照片實驗，記錄受測者當時的眼動狀況，並做了系統的研究。

1965 年，Yarbus 在他的著作《眼動與視覺》中提出了第一個關於眼球追蹤的報告，在這本書中，Yarbus 描述了眼睛觀察文字或圖像的注視動作，以及相關的研究結果。

圖一：眼球追蹤技術歷史發展重點



（資料來源：作者綜合整理自：Zohreh, Sharafi et al. (2020). “A practical guide on 10conducting eye tracking studies in software engineering”. *Empirical Software Engineering*. 25. 10.1007/s10664-020-09829-4.)

在 1970 年代和 1980 年代，眼球追蹤的相關研究更是蓬勃發展。眼動儀尺寸變得更小、更不具侵入性，以及具有更高的準確度。心理學家也制定了不同的理論來連結眼睛注視的狀態與認知過程(Jacob and Karn 2003)，並且研究領域也擴展到各種商業和科學等研究上。(Sharafi, Zohreh et al., 2020)。

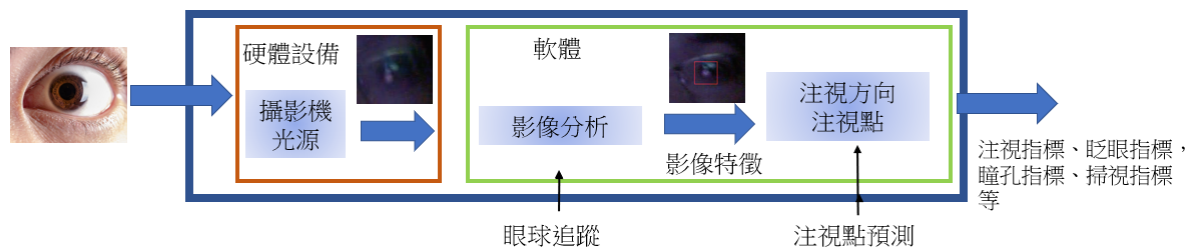
而時至如今，因著COVID-19疫情阻隔了很多實體活動的進行，在元宇宙情境下之使用方式也被熱鬧的探討著，這些眼球追蹤技術歷史發展重點如圖一所示。

二、眼球追蹤原理分析

心理學研究者常使用眼動追蹤儀(Eye Tracker)記錄實驗過程中，受試者眼睛所注視之位置與時間，並且以時間序列的方式記錄，成為做為探討認知行為的基本資料，這樣的方法目前已廣泛被運用於認知功能與情意動機反應等相關研究領域之中（吳昭容，2019）。

眼動儀機身上除了攝像鏡頭外，另有外置的紅外光LED光源，主要原理是瞳孔的部份幾乎不會反射紅外光，因此眼動儀便可以捕捉到一個暗色的圓形部分。而眼球上瞳孔以外的部份則會呈現藉反光而取得之光點，經過電腦影像處理過這些反光光點的資料之後，便可計算眼球目前的位置，以達到眼球的追蹤的目的（陳學志等，2010），我們將整個取像及分析的過程繪製如圖二所示。眼動儀採集的最原始的數據基本上就是時間序列上對應的屏幕坐標和瞳孔直徑，使用者可以依照所需分析的不同認知行為情境加以選擇利用，如表一。

圖二：眼球追蹤技術之系統構建架構



（資料來源：作者綜合整理自： Villanueva, A. et al. (2009). Report on New Approaches to Eye Tracking. Summary of new algorithms. <https://reurl.cc/AOZ75d>)

表一：眼動儀之基本輸出資料

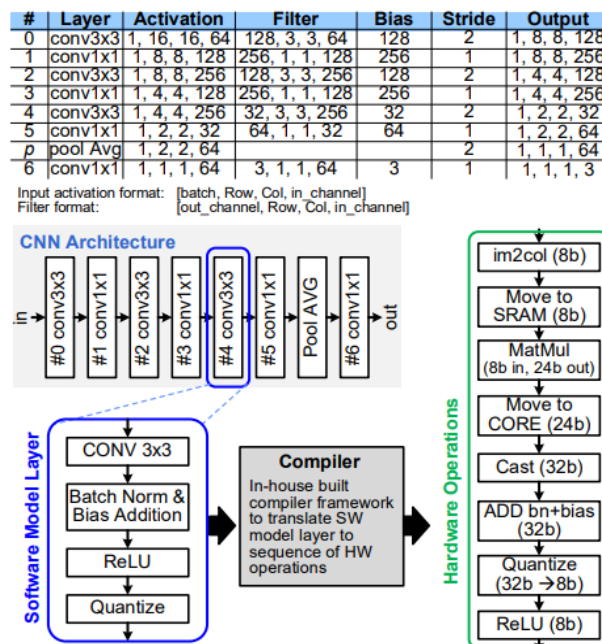
指標項目	指標意義
Gaze Point Left X, Y	左眼注視點的水平 (X) 和垂直 (Y) 座標
Gaze Point Right X, Y	右眼注視點的水平 (X) 和垂直 (Y)座標
Eye Position Left X, Y, Z	左眼相對於眼動儀的 3D 位置
Eye Position Right X, Y, Z	右眼相對於眼動儀的 3D 位置
左瞳孔直徑	左眼瞳孔的估計直徑，以毫米為單位
右瞳孔直徑	右眼瞳孔的估計直徑，以毫米為單位
Validity Code Left	表示眼動儀已正確識別左眼的置信度
Validity Code Right	表示眼動儀已正確識別右眼的置信度

(來源：作者綜合整理自：Tobii (2014), User's Manual : Tobii X2-30 Eye Tracker, July 22, 2022, <https://reurl.cc/O4NvgX>)

而使用紅外攝像感測器來跟踪眼球運動的VR裝置近年來也陸續出現，包含現在可以買到的Microsoft HoloLens 2、Pico Neo 2 Eye、HTC Vive Pro Eye和2022年9月新推出的VIVE Focus 3 眼球追蹤套件，甚至是預期不久的將來會出現一波新的 VR 頭戴裝置，包含Meta 名為 Project Cambria 的專業級 VR 頭戴裝置、PlayStation VR 2，或許還有Apple 裝置。

在眼球追蹤的影像分析演算法發展上，也漸漸從傳統以色彩或銳利度變化為主的方法，轉為使用訓練過的人工智慧模型來偵測，如圖三即是Meta 在2021 年所發表的眼球追蹤客制化晶片上所執行的模型，可以看到僅僅使用了結構簡單的卷積層(Conv)、池化層 (Pool)、激活層 (Relu) 等常用算子，利用大數據訓練得到的數值統計原理，即可找出眼球追蹤的特徵訊息。

圖三：眼球追蹤特徵提取之 CNN 模型尺寸和AI模型架構



資料來源：(H. E. Sumbul et al., "System-Level Design and Integration of a Prototype AR/VR Hardware Featuring a Custom Low-Power DNN Accelerator Chip in 7nm Technology for Codec Avatars," 2022 *IEEE Custom Integrated Circuits Conference (CICC)*, 2022, pp. 01-08)

叁、研究方法

運用文獻分析法，針對所檢索到的國際期刊論文、報告、文章和網路影片進行整理和分析，列出該技術的起源、發展、原理和未來應用方向。在查找資料的同時，除了加深和加廣眼球追蹤的知識外，也對名詞和現象進行更深入的討論，以期對研究成果做到更嚴謹的驗證。

肆、研究分析與結果

一、使用眼球追蹤研究學生在課堂的注意力狀況

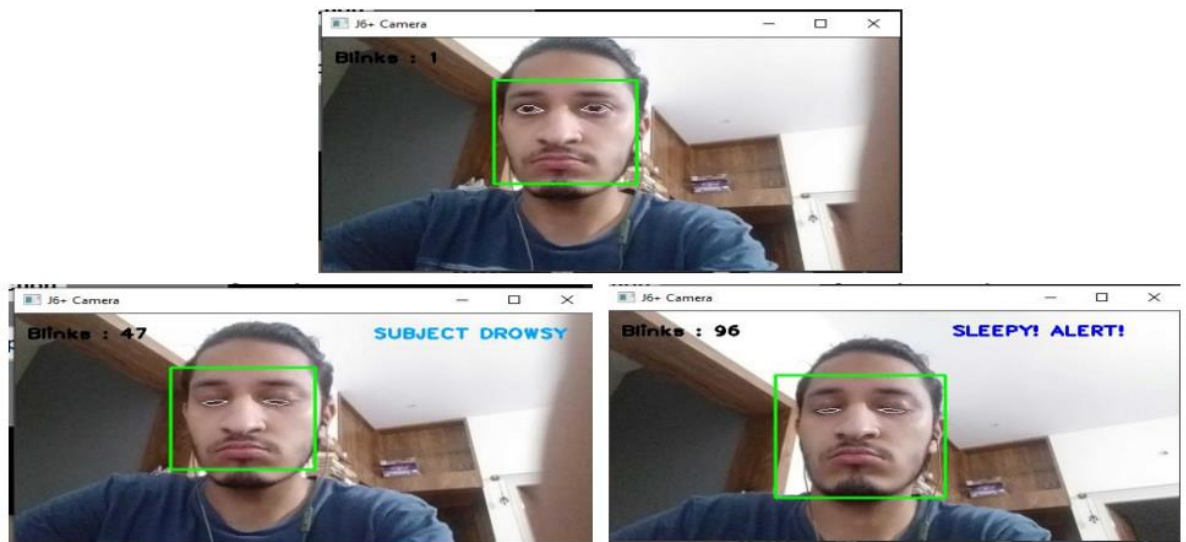
在2012年美國物理學會會議論文集中，有一篇來自肯尼索州立大學 (Kennesaw State University) 關於自然科學系教學過程中學生的注視模式的研究(Rosengrant, David et al. 2012)。在這篇研究中，除了可以針對有哪些因素能夠讓學生在課堂上中保持聽課狀態，也分析了是什麼因素分散了學生的注意力，以及最重要的在整個授課過程中學生的關注重點有那些。

二、商業化實例

(一) 駕駛者監控系統 (Driver Monitoring System)

2019年，歐盟要求所有在歐洲生產的車輛必須配備駕駛者監控系統 DMS (Driver Monitoring System)。如果其他國家汽車製造商的汽車銷往歐洲，也必須遵守同樣的規定。由於眼球追蹤技術可以輔助分析駕駛當前的認知和心理狀態，如圖四所示，包括是否足夠專注、是否有疲勞現象、是否心情不好等 (A.S, et al. 2021)，這樣的技術使用有望受到其他地區汽車廠商的青睞。

圖四：駕駛者監控系統 (Driver Monitoring System) 使用示意



資料來源：(MB, A. S, A. et al. (2021) "Drowsy Driver Detection using Eye-Tracking through Machine Learning," *2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 2021, pp. 1916-1923.

(二) 消費者行為調查解決方案

眼球追蹤技術製造商Tobii發表了基於VR的消費者行為調查解決方案，通過即時取得眼球追蹤分析，可以讓品牌、零售商和展場設計師能夠分析消費者在購買行為的決策關鍵影響因素，如圖五所示(Bonasio, Alice , 2018)。

圖五：基於VR的消費者行為調查解決方案



資料來源：Bonasio, Alice (2018). Looking Through Your User's Eyes With Virtual Reality Gaze Tracking, Retrieved Jun 21, 2018. from <https://reurl.cc/AOryej>

（三）輔助工具上的應用

霍金博士被公認為相當傑出的理論物理學家，但他過世前的日子卻飽受漸凍人症狀所苦。為了能讓霍金博士能順利的與外界溝通，Intel在2014年時採用當時最新的眼球追蹤技術，結合腦波儀（electroencephalogram）的功能做為操作電腦的媒介。這項新技術利用追蹤眼球運動的特性，讓霍金博士可以使用眼睛就像使用電腦滑鼠一樣“指向”電腦屏幕。霍金博士的眼鏡上並附有一個電腦晶片，藉由這個晶片的功能，霍金博士每分鐘可以表達五到十個單詞，這也比他之前使用的電腦輸入法快得多，如圖六。（Android Castle, 2017）

圖六：霍金博士與外界溝通的方式



資料來源：Android Castle (2017)。Collabroration of intel lab with Dr. Stephan Hawking on .NET based Assistive Technology〔影片〕。YouTube。 <https://www.youtube.com/watch?v=w9wXZIIK8UA>

三、傳達複雜的社會和情感狀態之研究

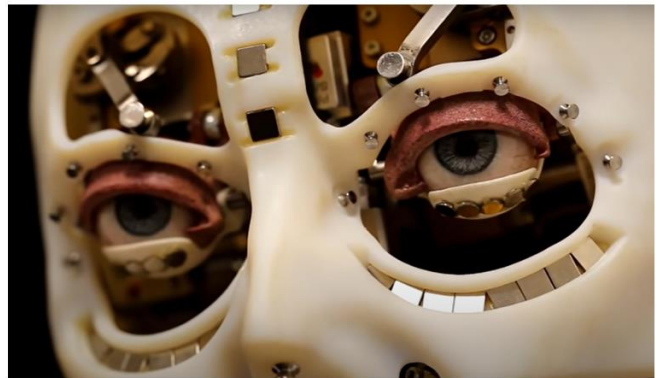
由於眼睛注視不止已被證明是一種關鍵的社會互動信號，注視也同時傳達著複雜的情感狀態。鑑於注視在社交互動中的重要性，在2020年智能機器人與系統國際會議中，迪士尼研究院的研究小組發表了一篇他們近期發明一種更逼真的機器人注視系統(Pan, et al., 2020)。科學家們用一個名為Audio-Animatronics® 的半身像演示了這項技術，如圖七。這項研究的目標主要是想開發一個系統來模擬類似人類的相互注視，為了使與機器人的互動更加逼真，研究小組不僅設計程式控制眼睛的動作，還設計程式可以控制脖子和眉毛，如圖八。研究小組使用這樣的交互式眼動機器人來模擬以注視的方式傳達情感狀態及塑造感知的能力，這也將會是一個重要的互動實驗驗證平台。

圖七：Audio-Animatronics® 半身像用於演示凝視系統



資料來源：M. K. X. J. Pan et al., "Realistic and Interactive Robot Gaze," 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2020, pp. 11072-11078

圖八：眼動機器人眼部細節



資料來源：DisneyResearchHub (2021)。Realistic and Interactive Robot Gaze [影片]。YouTube。
https://www.youtube.com/watch?v=D8_VmWWRJgE

伍、研究結論與建議

一、技術瓶頸

(一) 避免額外紅外線光源之使用

常見的眼球追蹤裝置如Tobii，目前是使用了外置紅外線光源對瞳孔和虹膜對紅外光吸收率的差異而取得的瞳孔影像，再經過軟體的影像處理而分析得到使用者的眼球活動。由於這樣的方式需使用紅外線光源，在功耗及裝置機構都是設計上所需要另外考慮的因素。

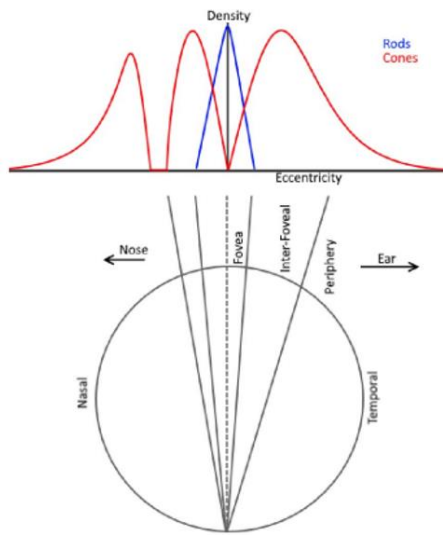
而由台灣大學電機系簡韶逸教授所領導的見臻科技團隊所做出的眼球追蹤模組，則是採用左右兩個一般的影像感測器，加上使用人工智慧的演算法，做成了微小化嵌入式的眼球追蹤模組，可整合進VR/AR裝置成為智慧眼鏡的關鍵模組。

(二) 注視點渲染技術 (foveated rendering) 之實現

由於未來VR眼鏡畫面可能到達 8K 大小的解析度，並且可能是360度全方位的VR裝置。假如保持場景全部畫出，勢必消耗相當大運算力去呈現畫面。然而人類眼球所能看到的範圍其實是有一定範圍，依使用者的不同僅僅有約10到20度視角範圍（如圖九）。這是因為我們的眼

睛無法看到眼睛正在接受的所有細節。事實上眼睛中心的中央凹部分會跳來跳去查看細節，我們實際上只會專注中央的部分，而中央的外圍就會漸模糊，最外圍會更模糊。同樣的技巧可以應用在眼球追蹤 VR 中，我們也就是搭配眼球追蹤技術，利用人眼的這個特性，降低了注視區域之外的影像解析度（如圖十），這樣就能節省繪圖處理器在渲染圖像時所需的功耗以增進其效能。

圖九：人眼可分為中央凹、中央凹間和周邊區域三個主要區域。



資料來源：Matthews, S. L. et al. (2020) "Rendering Optimizations for Virtual Reality Using Eye-Tracking," *2020 22nd Symposium on Virtual and Augmented Reality (SVR)*, 2020, pp. 398-405.

圖十：注視點渲染技術的效果示意



資料來源：VR Expert (2022). What is foveated rendering and what does it mean for VR?, Retrieved August 22, 2022. from <https://reurl.cc/RXVO7Z>

（三）眼球追蹤人工智慧模型的最佳化

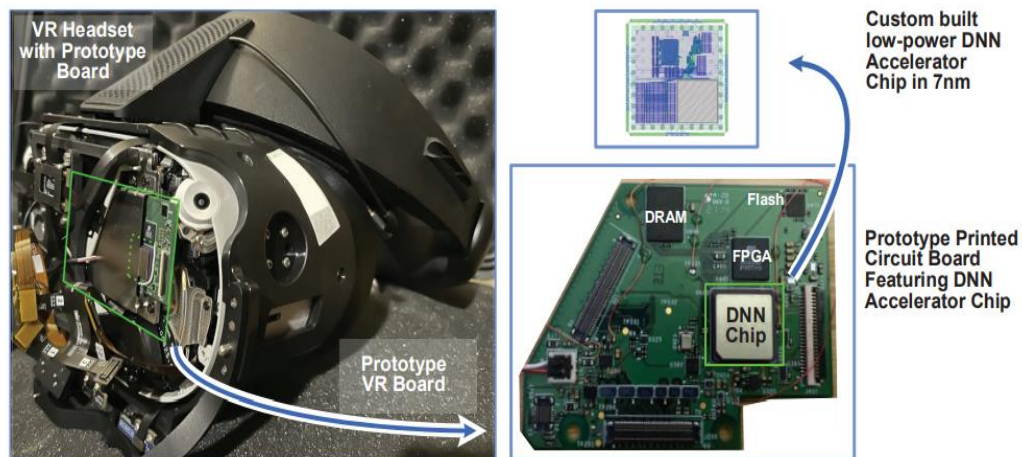
目前朝向把眼球追蹤人工智慧模型權重數降低，搭配資料位寬的縮小做到模型輕量化，以期用更少的計算資源達到眼球追蹤的效果。(Sumbul, et al. 2022)

（四）晶片功率消耗之降低

對於AR/VR這類可穿戴用戶來說，運行在超低功耗上的設備的優勢是很容易理解的。通過消耗更少的能量，電池供電的端點設備可以佩戴更長時間而無需充電。有了這樣的低功耗特性，製造商可以讓消費者使用容量及尺寸更小的電池，從而實現更薄的外形和更時尚的設計，因此眼球追蹤的資料取得及運算的部分也需要儘可能的省電。

Meta Reality Labs團隊在2021年打造出一款專為AI處理設計的客製化加速晶片，如圖十一，實測上在這個7 奈米製程的測試晶片上，可以達到在 2.56 mm² 晶片面積內達到 375 μ J/幀/眼的眼球追蹤能效。(Zhang, X. et al, 2021)

圖十一：VR 硬件原型的系統細節：包含修改過的 Oculus Quest 2 移動 VR 裝置(左方)、原型印刷電路板(右下)，以及特別定制7奈米低功耗 DNN 加速器測試芯片(右上)。



資料來源：(Zhang X. et al. (2021) "F-CAD: A Framework to Explore Hardware Accelerators for Codec Avatar Decoding," *2021 58th ACM/IEEE Design Automation Conference (DAC)*, 2021, pp. 763-768)

二、個人資料隱私問題

在元宇宙的虛實世界中，我們若將類似個人的注視數據、眼動方式、瞳孔大小及相關預測資料等和這些個人有關的資料傳上網路，就會面臨個人資料隱私問題。

而在近期的研究中，就有學者提出了一個資料保護框架(David-John, et al. 2021)，該框架透過應用程式 API 接口的設計，以及藉由軟件實現的隱私機制一起結合把關，即使個人資料標識將允許應用程序將用戶的個人 ID 與他們的工作 ID 關聯起來，然而藉由這樣的保護機制，將在虛擬現實中的自然觀看條件下，能夠減少用戶可能輕易被識別的風險。

三、元宇宙中人機互動方式的改變

(一) 以眼神控制 VR 中的事物

若在公共場合使用VR裝置，『手勢操作』等控制方式過於明顯，相較之下使用眼球追蹤的控制方式則有更好的沉浸感。在VR裝置上使用眼球追蹤技術，可以通過眼球直視而不是嘗試用手來控制 VR 中的事物，藉由這樣做法應該是進入未來具有更好看的 VR 圖形，和更小、更智能VR裝置的門戶。

(二) Codec Avatars 分身互動情境下之眼神情感交流

Meta Reality Labs 團隊在 2021 年打造出一款專為 AI 處理設計的 7nm 客製化加速晶片，可以耗費較少電力，在 VR 裝置呈現逼真的 Codec Avatars 分身。(Sumbul, H.E. et al. 2022)。而Meta也在2022年6月於臉書、Instagram及Messenger上推出 Avatars Store，可以讓使用者替自己的虛擬替身添購新裝。圖十二即為Meta 執行長祖克柏以自己與其妻子的虛擬替身作了換裝的示範(ABP News Bureau, 2022)。在未來的VR社交互動中，由於眼球追蹤的技術可以捕捉到使用者更

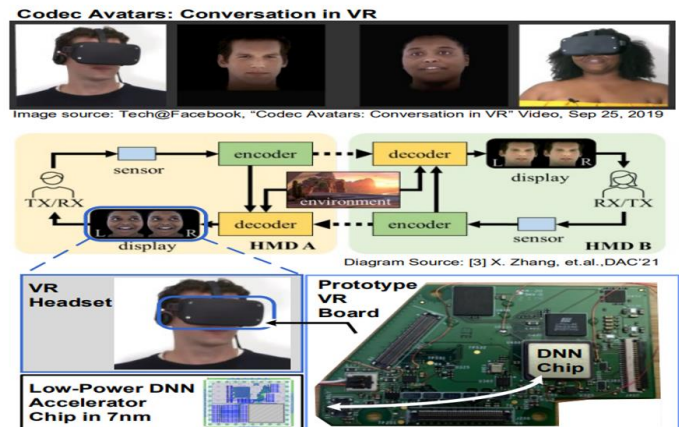
細膩的眼神變化，藉此可讓VR 化身在未來的元宇宙空間中與其他使用者之化身進行更像人類的眼神互動。這也將是 Meta 下一代 Oculus Quest 產品的亮點，如圖十三。

圖十二：祖克柏以自己與妻子的虛擬替身作了換裝的示範



資料來源：ABP News Bureau,(2022) “Mark Zuckerberg's Meta Wants You To Buy Designer Outfits For Your Digital Avatar, Launching Avatars Store” Retrieved 18 Jun 2022. from <https://reurl.cc/RXVOK9>

圖十三：以兩個使用者的分身模型互動應用場景為例，VR 裝置上使用了可以追蹤使用者眼球的定制 DNN 加速器測試晶片的原型硬件，如此做出逼真的分身互動演示



資料來源：(H. E. Sumbul et al., "System-Level Design and Integration of a Prototype AR/VR Hardware Featuring a Custom Low-Power DNN Accelerator Chip in 7nm Technology for Codec Avatars," 2022 IEEE Custom Integrated Circuits Conference (CICC), 2022, pp. 01-08)

（三）將逝者之眼神合成至生前相片達成對逝者的追思

我們若能將已故先人生前的眼神先記錄起來，在過世之後可以藉著上述互動情境下之眼神情感交流，將眼神合成至長輩生前之相片，藉此讓晚輩仍能與長輩互動，達到對逝者追思的目的。筆者另以此概念先於此論文揭露，同時申請專利。

（四）將指導者之視角投射到學習者之視角

在 COVID-19 疫情仍持續下，人與人也漸漸習慣無需實體接觸的遠端工作模式。在元宇宙的使用情境下，也可以成為具經驗的指導者傳遞經驗的學習過程。以維修複雜結構之機器為例，我們也可將指導者觀察問題之視角在學習者的屏幕上投射出同樣之視角，這樣藉由指導者觀察問題的角度，讓學習者快速掌握問題的癥結點，如此便可加速學習的效果。筆者另以此概念先於此論文揭露，同時進行專利申請。

眼睛是靈魂之窗，在目前逐漸形成的元宇宙的世界中，戴上VR或AR穿戴式裝置之後，只要注視畫面中的特定控制區域、滑動視線，亦即藉由本篇所介紹的眼球追蹤這項技術，不再靠著傳統控制板或手勢控制，即可操控眼鏡裝置裡的各項選項功能和外界交流。人機緊密的結合的結果，將會開創更多前所未有的應用領域。

陸、參考文獻

吳昭容（2019）。眼球追蹤技術在幾何教育的應用與限制。臺灣數學教育期刊，6（2），1-25。

陳學志、賴惠德、邱發忠（2010）。眼球追蹤技術在學習與教育上的應用。教育科學研究期刊第五十五卷第四期，55（4），39-68。

DisneyResearchHub (2021)。Realistic and Interactive Robot Gaze〔影片〕。YouTube。

https://www.youtube.com/watch?v=D8_VmWWRJgE

Android Castle (2017)。Collabration of intel lab with Dr. Stephan Hawking on .NET based Assistive Technology〔影片〕。YouTube。

<https://www.youtube.com/watch?v=w9wXZIILK8UA>

VR Expert (2022). What is foveated rendering and what does it mean for VR?, Retrieved August 22, 2022. from <https://reurl.cc/RXVO7Z>

Bonasio, Alice (2018). Looking Through Your User's Eyes With Virtual Reality Gaze Tracking, Retrieved Jun 21, 2018. from <https://reurl.cc/AOryej>

ABP News Bureau,(2022) “Mark Zuckerberg's Meta Wants You To Buy Designer Outfits For Your Digital Avatar, Launching Avatars Store” Retrieved 18 Jun 2022. from <https://reurl.cc/RXVOK9>

David Rosengrant et al. (2012) "Following student gaze patterns in physical science lectures", *2012 AIP Conference Proceedings* 1413, 323

X. Zhang, D. Wang, P. Chuang, S. Ma, D. Chen and Y. Li, (2021) "F-CAD: A Framework to Explore Hardware Accelerators for Codec Avatar Decoding," *2021 58th ACM/IEEE Design Automation Conference (DAC)*, 2021, pp. 763-768.

H. E. Sumbul et al. (2022) "System-Level Design and Integration of a Prototype AR/VR Hardware Featuring a Custom Low-Power DNN Accelerator Chip in 7nm Technology for Codec Avatars," *2022 IEEE Custom Integrated Circuits Conference (CICC)*, pp. 01-08.

M. K. X. J. Pan et al. (2020) "Realistic and Interactive Robot Gaze," *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 11072-11078.

Sharafi, Zohreh & Sharif, Bonita & Guéhéneuc, Yann-Gaël & Begel, Andrew & Bednarik, Roman & Crosby, Martha. (2020). “A practical guide on conducting eye tracking studies in software engineering” . *Empirical Software Engineering*. 25. 10.1007/s10664-020-09829-4.

Villanueva, A., Cabeza, R., Porta, S., Böhme, M., Droege, D., & Mulvey, F. (2009). Report on New Approaches to Eye Tracking. Summary of new algorithms. <https://reurl.cc/AOZ75d>

B. David-John, D. Hosfelt, K. Butler and E. Jain, (2021) "A privacy-preserving approach to streaming eye-tracking data," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 5, pp. 2555-2565.

S. L. Matthews, A. Uribe-Quevedo and A. Theodorou, (2020) "Rendering Optimizations for Virtual Reality Using Eye-Tracking," *2020 22nd Symposium on Virtual and Augmented Reality (SVR)*, 2020, pp. 398-405.

A. S, A. MB, S. D and A. C, (2021) "Drowsy Driver Detection using Eye-Tracking through Machine Learning," *2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 2021, pp. 1916-1923.

Tobii (2014), User's Manual : Tobii X2-30 Eye Tracker, July 22, 2022, <https://reurl.cc/O4NvgX>