

## SIGGRAPH 2024 參訪報告

國際策發部 王如蘭

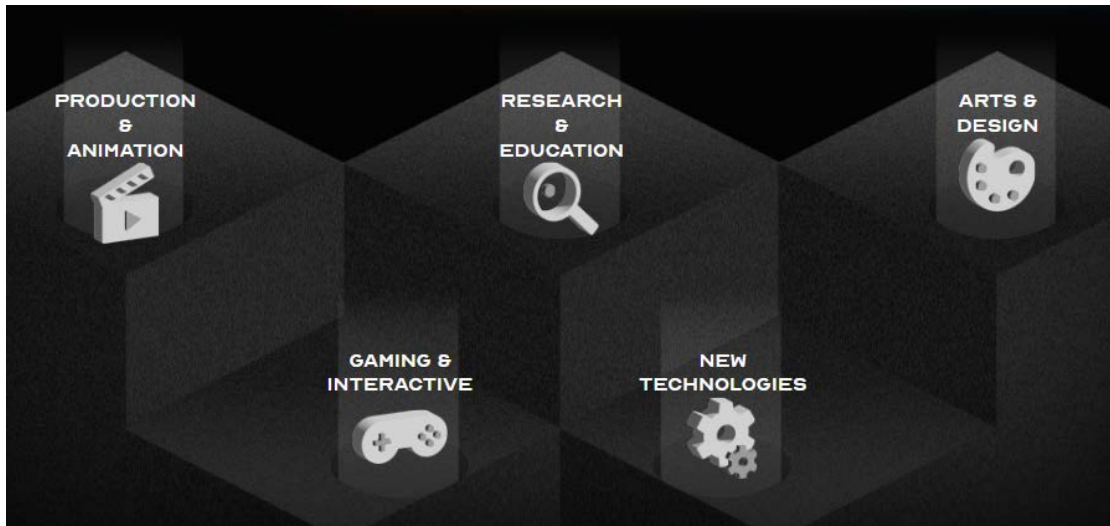
SIGGRAPH 全名為 Special Interest Group on Computer Graphics and Interactive Techniques，是美國電腦協會（Association for Computing Machinery, ACM）旗下的一個專門小組，ACM 是一個非營利性國際科學和教育組織，致力於電腦圖形學和互動技術領域的研究和發展。

SIGGRAPH 年度電腦繪圖大會是電腦圖形學界最重要的國際會議之一，已經舉辦了 50 屆。今年是迎向下一個 50 年的第 51 屆會議，於 7 月 28 日到 8 月 1 日回到了起源地美國 Colorado 州 Denver 舉行 SIGGRAPH 2024 大會。



SIGGRAPH 2024 會場位於 Colorado Convention Center；著名的大藍熊看著會場

由於 ACM 本身就是一個學術型研究組織，所以 SIGGRAPH 2024 會議內容十分充足，總共有製作與動畫、研究與教育、藝術與設計、遊戲與互動、新科技等五個主題。



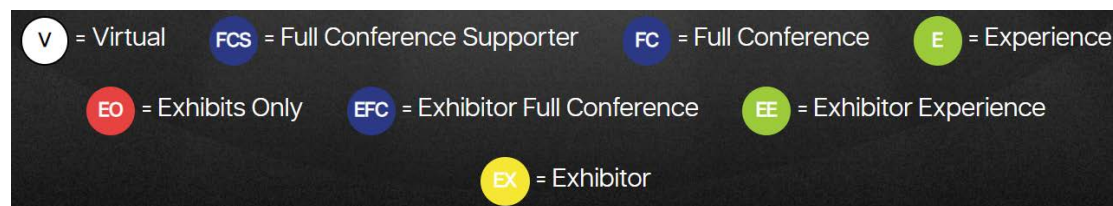
SIGGRAPH 2024 五主題；摘錄自 <https://s2024.siggraph.org/>

從主題下發展個子題分項，包括研究報告發表、紙本論文、教育論壇、沉浸式創新、電腦動畫、主題演講合計上百場次，內容等非常多元。

<b>Technical Papers</b> <b>PROGRAM</b>	<b>Art Papers</b> <b>PROGRAM</b>	<b>Talks</b> <b>PROGRAM</b>	<b>Production Sessions</b> <b>PROGRAM</b>	<b>Panels</b> <b>PROGRAM</b>	<b>Courses</b> <b>PROGRAM</b>
<b>Educators Forum</b> <b>PROGRAM</b>	<b>Frontiers</b> <b>PROGRAM</b>	<b>Awards</b> <b>PROGRAM</b>	<b>Electron Theater</b> Computer Animation Festival <b>PROGRAM</b>	<b>Real-Time Live!</b> <b>PROGRAM</b>	<b>VR Theater</b> <b>PROGRAM</b>
<b>Immersive Pavilion</b> <b>PROGRAM</b>	<b>Retrospective</b> <b>PROGRAM</b>	<b>Art Gallery</b> <b>PROGRAM</b>	<b>Posters</b> <b>PROGRAM</b>	<b>ACM SIGGRAPH Village</b> <b>PROGRAM</b>	<b>Labs Experience Hall</b> <b>PROGRAM</b>
<b>Emerging Technologies Experience Hall</b> <b>PROGRAM</b>	<b>Games</b> <b>PROGRAM</b>	<b>Diversity, Equity, and Inclusion</b> <b>PROGRAM</b>	<b>Birds of a Feather</b> <b>PROGRAM</b>	<b>Appy Hour</b> <b>EVENT</b>	<b>Student Volunteers</b> <b>PROGRAM</b>
<b>Student Research Competition</b> <b>EVENT</b>	<b>Lead</b> <b>PROGRAM</b>	<b>Keynote Presentations</b> <b>PROGRAM</b>	<b>SIGGRAPH for Peace</b> <b>PROGRAM</b>		

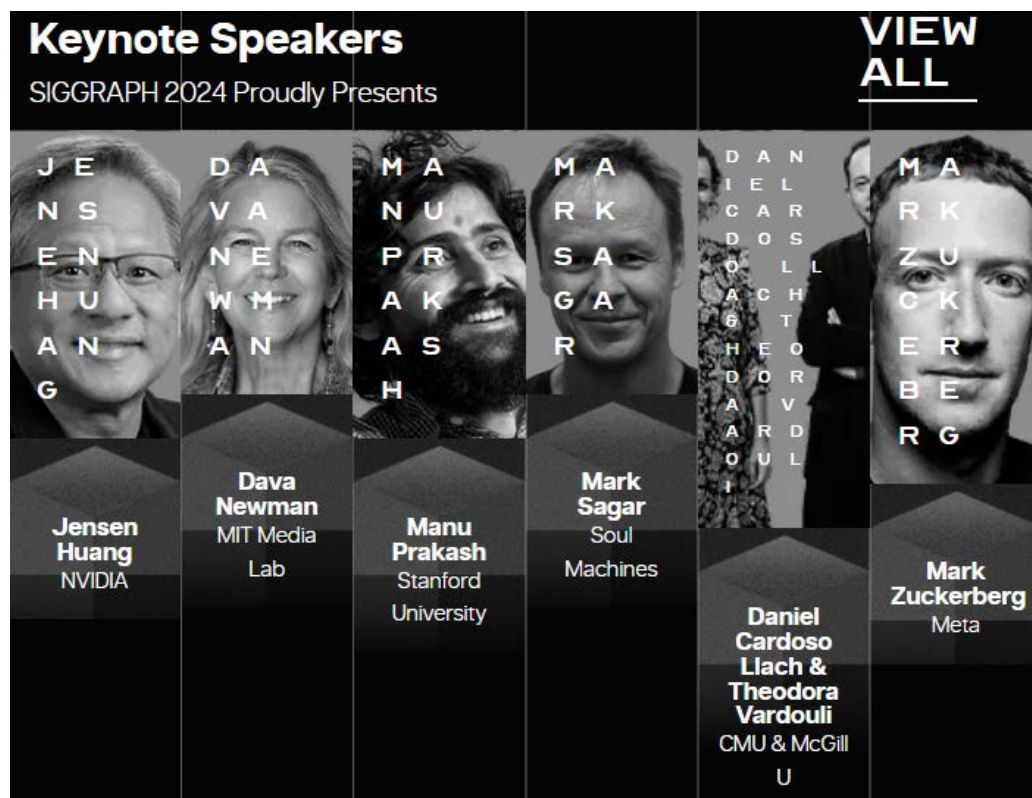
SIGGRAPH 2024 各個子題分項；摘錄自 <https://s2024.siggraph.org/>

但也由於內容多元，所以報名註冊也分為許多類別，各類別與會費用不同，能夠參與的項目也有很大的差異。由於與會費用高昂，本次是以早鳥 Experience 身分與會，所以能夠參與的場次有限，如 Talk、Panel、Course、Production Session 等就無法參與，但由於會議場次高達數百場，就算只參與部分也能稍能初步一窺 SIGGRAPH 的概貌。



SIGGRAPH 2024 報名註冊類別；摘錄自 <https://s2024.siggraph.org/>

當然，重頭戲還有主題演講。包括了 NVIDIA CEO 黃仁勳、Meta 創辦人暨執行長 Mark Zuckerberg 以及 MIT Media Lab Director Dava Newman 等人。以下將就最受矚目的 NVIDIA 與 Meta 主題演講內容進行重點摘要。



主題演講講者；摘錄自 <https://s2024.siggraph.org/>



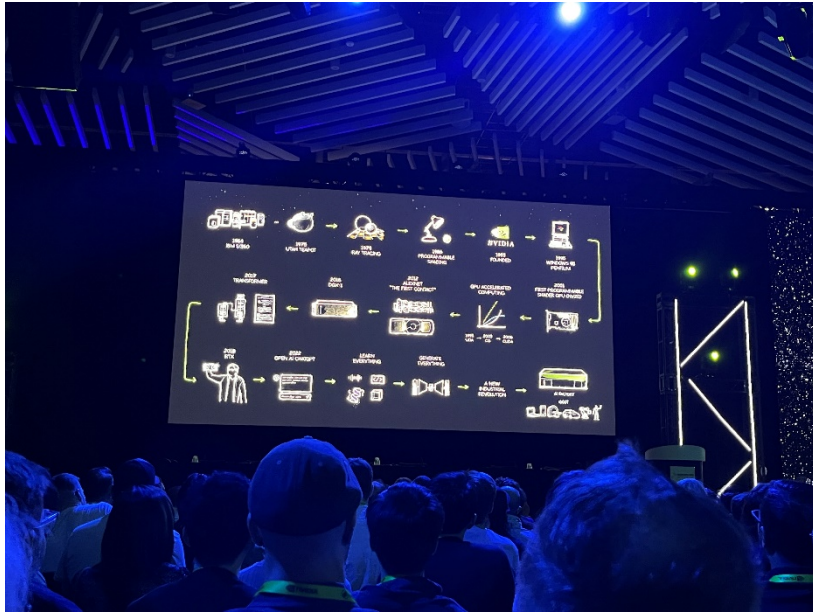
## NVIDIA+ Meta 兩大巨頭圍爐座談 帶領 AI 新想像

身為會議最大的贊助商 NVIDIA，NVIDIA Research 在會場發表超過 20 篇研究論文，每天安排 10 多場的演講或課程，11 個參展廠商現場展示是使用 NVIDIA 產品，SIGGRAPH 2024 真可謂是 NVIDIA 的秀場，也是 CEO 黃仁勳口中的 HUB。



NVIDIA 現場課程，每個人直接上線操作

NVIDIA CEO 黃仁勳在主題演講時以 AI 製作圖卡簡報，說明 NVIDIA 如何參與電腦動畫的歷史與 AI 進展。重大進成為 2001 年發明世界上第一個可編輯光影著色功能 GPU、接著發明了 CUDA 開啟了加速運算革命。2012 年推出了全球首款 NVIDIA RTX，由圖形和模擬共同推動了 AI 的進步，而 AI 的進步也反過來推動圖形和模擬的發展，AI 時代於此展開。



NVIDIA CEO 黃仁勳主題演講

在主題演講的第二階段為黃仁勳和 Meta 創辦人暨執行長 Mark Zuckerberg 進行圍爐座談，討論 AI 人工智慧的未來。Meta 日前推出迄今為止最大的開源 AI 語言模型，祖克柏表示 Meta AI 是走開源模式，除了透過 AI 模型優化社群系統運作，包括 Facebook 與 Instagram 上的動態消息與推薦系統，都將發生改變。未來 Meta 的推薦系統不再只是協助用戶與朋友建立聯繫，更可以推播更多用戶可能有興趣的內容。對此，黃仁勳補充其實世界上最大的運算系統之一，就是推薦系統。

黃仁勳與祖克柏一致認為接下來產業內將會出現大量的 AI 代理人(AI Agent)，為企業與個人拓展業務。以 Meta 來說，認為不論是企業或是社群上所有的創作者，未來都能建立 AI 代理人與客戶進行互動或銷售產品以及提供產品支援等，AI 代理人甚至能協助企業與社群上的其他人互動。現今 Meta 於美推出的 AI Studio 的新工具，讓美國用戶能在 Instagram 上製作 AI 分身、透過 AI 與粉絲互動交談，並回應原帳戶的留言等，就是個人專屬 AI 代理人的開始。黃仁勳對於未來 AI 代理人的想像則是認為未來 AI 運算會更類似人類思考的「決策樹」(decision tree) 模式，AI 將同時模擬多個決策可能

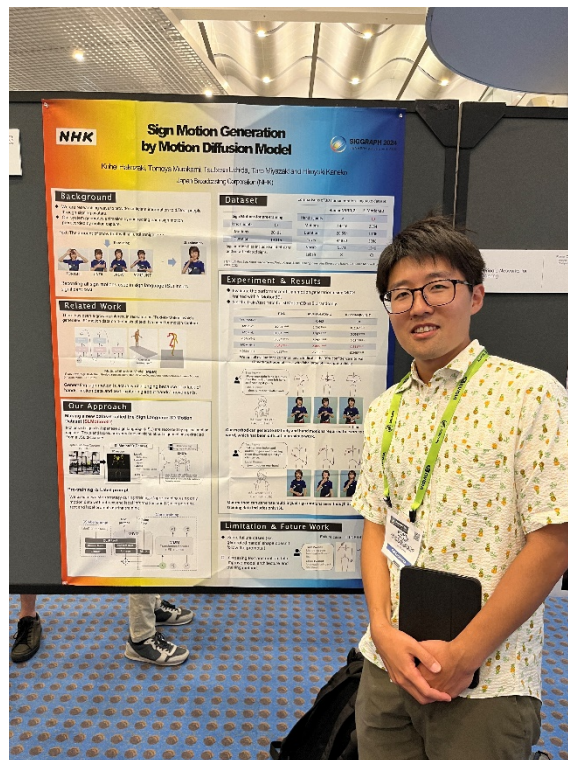
帶來的結果，最終選出最佳方案。

祖克柏於後又提到 META 與 Ray-Ban 捲土重來推出的第二代智慧眼鏡 Ray-Ban Meta Smart Glasses，與第一代一樣具備聽音樂、拍照、錄影等功能，但第二代眼鏡更可於 FB 與 IG 上進行直播，並具備回應(未來預計支援翻譯)的 AI 功能，希望可以開創可穿戴式智能設備新市場。

由於本次與會主要目的為了解 AI 手語生成可能技術，茲將於會場探訪之各式 AI 手語生成可能技術及廠商概述於下。

### 日本 NHK 技術研究所

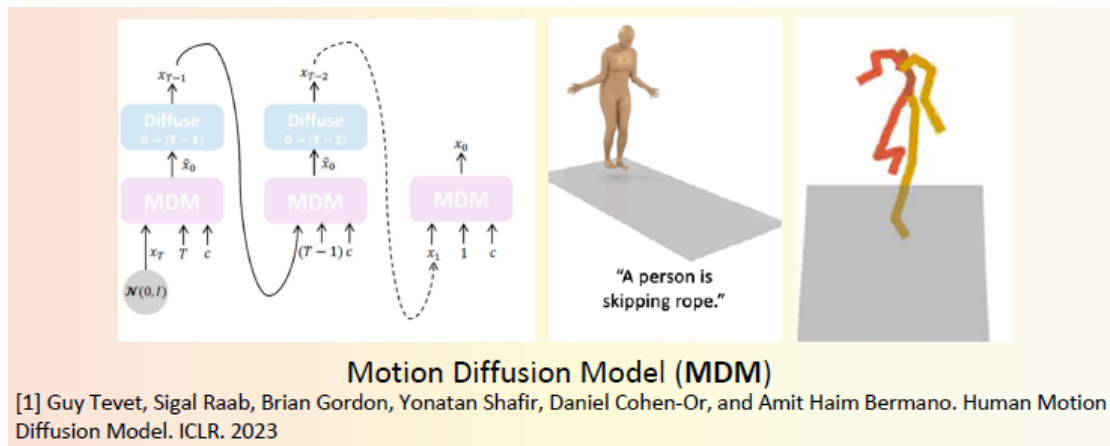
NHK 技術研究所在手語研究上已經投入了相當長的時間，從應用動作捕捉技術開始，到發展 CG 動畫以及至今的 AI 技術應用。在 SIGGRAPH 2024 大會中，NHK 技術研究所的箱(山奇)浩平 (Hakozaki Kohei) 研究員，發表了一篇《Sign Motion Generation by Motion Diffusion Model》的研究論文。



NHK 技術研究所 Hakozaki Kohei 研究員及其壁報論文



NHK 技術研究所現嘗試以 Guy Tevet 於 ICLR2023 會議中發表的 Motion Diffusion Model(MDM；運動擴散模型)為基礎，以既有的日本手語字典中的手語表述文字來自動生成 3D 手語動作(該動作為歷年由動作捕捉技術所累積)，如此一來預計可以減少後續再行利用動作捕捉系統剩餘字詞的高昂成本。



<https://guytevet.github.io/mdm-page/>

本次研究 NHK 技術研究所希望能夠蒐集一系列稱之為 the Sign Language-3D Motion Dataset (SLMotion3D) 的新資料集。技研以兩階段進行訓練，第一階段為 pre-training 階段只有動作數據沒有包含文字與標註資訊，第二階段為加入文字與標註。第一階段收錄了含手部關節的 20,127 個動作，時長總共 19.31 小時；第二階段則收錄了含手部關節的 2,354 個動作，時長 1.1 小時，以及 2,398 個對應文字以及 1,240 個語詞，並進行了標註。上述資料集雖然僅收錄日本手語(JSL)，但是系統可以自動轉換生成如美國手語(ASL)等多國手語動作。

研究最後對模型進行比較評估，結果顯示，在加入文字與標註資訊後，模型性能顯著提升。以輸入日本手語“LOVE”為例，輸入 LOVE 就會生成骨架結構，再套入 3D 模型生成 3D 的 LOVE 手語動作。另外也可自動轉換生成美國手語，打“CUTE”即可生成 3D 美國手語的 CUTE 手語動作。

Kohei 研究員說明由於目前是以文字生成動作，而 MDM 運動擴散模型中較缺乏對應於手語的細部手勢動作資料以及對動作的文字解釋，所以在兩手重疊以及旋轉的動作上會出現錯誤。比如說“ILLNESS”這種握拳輕敲額頭前後移動的動作，目前會出現錯誤。另外由於收錄的文字不夠多，後續必須再行增加收錄的規模。故目前此方法仍在研究測試階段。

由於目前 NHK 的 AI 手語服務套入的動作模型為由動作捕捉技術生成，接下來本文將分析 SIGGRAPH 2024 展場中各家動作捕捉技術的優缺點與適用性。

### **動作捕捉系統(Motion capture；mocap)**

動作捕捉系統是利用以下數種科技形式記錄人或物體的動作，動作可用主動式(Optical-Active)或被動式光學攝影機(Optical-Passive)或是慣性偵測器(inertial sensors)進行捕捉，或是不須標記直接把影片讓軟體或 AI 進行分析。動作捕捉系統如其名，只會捕捉動作形成骨架，後續可以將動作骨架套入任何想要生成的 2D 或 3D 模型。NHK 已經以動作捕捉系統收錄手語數年之久，近年因為 AI 以及科技進步，陸續也開始嘗試以其他方式收錄手語動作，但目前仍是以動作捕捉技術為主。動作捕捉系統有數種形式，茲將 SIGGRAPH 2024 動作捕捉設備相關參展廠商介紹如下

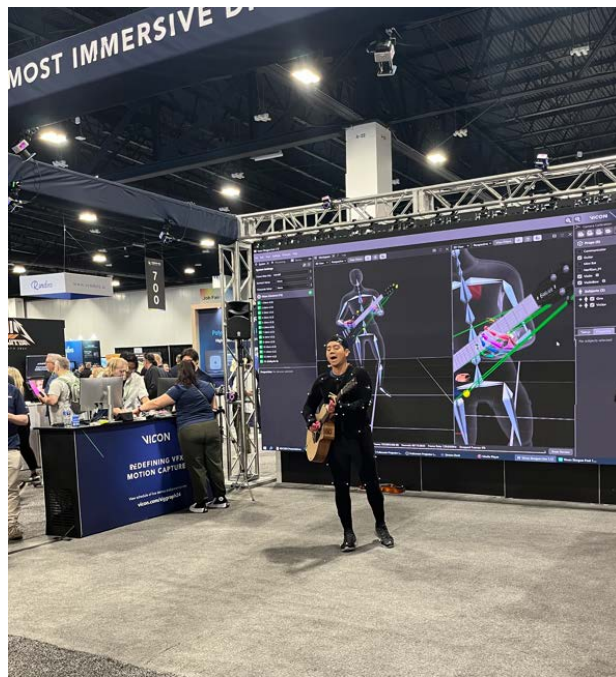
### **Vicon 動作捕捉系統**

Vicon 是目前 NHK 使用的動作捕捉攝影系統廠商，他是被動式光學技術(Optical-Passive)，透過架設多台紅外線高速動作捕捉攝影機建構起三維空間，利用配戴在人物或移動物品上反光球，即可將標記的空間位置資料轉換為數據資料。目前被動式光學動作捕捉是較準確靈活以及最常見的動作捕捉類型。

被動式光學技術要能捕捉到精細的動作，主要需要高解析度的攝影機，解析度越高從標記獲得的細節就越多。像手部動作的細節



以及臉部微妙的表情，都需要精確追蹤，這時候高解析度的攝影機就是關鍵。再來就是速度，對於快速移動的物體或運動，高速攝影才能有效追蹤到整個運動動作。此外影響整個動作捕捉結果好壞的還有動作捕捉軟體、系統校正、同步器以及電腦效能等。以往動作捕捉較難處理的遮擋(比如說趴下)以及手部細節動作，現在都因為軟體的進步，都已經獲得大幅度的改善，電腦運算系統可以直接根據可追蹤到的標記來計算看不到的標記位置。



Vicon 無線幕現場即時動捕生成骨架

Vicon 現已無需綠幕即可進行動作捕捉並即時套入 3D 模型，以往因場地受限的程度將大幅降低，大幅增加動作捕捉系統的可攜性。



Vicon 無線幕現場即時動捕套入 3D 模型

若考慮到需要同步捕捉臉部的表情，NHK 的解決方案是戴上 Headcam 頭部攝影機將影像帶入 Faceware 臉部表情系統，而 Vicon 在 SIGGRAPH 會場中的臉部辨識解決方式是與 Ozone Story Tech 合作，用左、右手操控 Henson puppet system 來表演出 3D 動畫人物的臉部表情與嘴型，相當吃操作人員的能力與熟悉度，但不用經過系統整合而可即時合成，十分令人驚豔。



操作 Henson puppet system 來演出 3D 動畫人物的臉部表情與嘴型

## QUALISYS 動作捕捉系統

QUALISYS 系統同時支援主動式和被動式光學捕捉系統，並且除了一般室內外活動外，是世界唯一可以捕捉水下動作的系統。強調可運用在大範圍、防水、防塵、耐熱，利用 AnyBody Modelling System 形成骨架資料，但由於強調大空間應用，故在手指動作的細緻度上較為不足。但是在操作介面上，可利用可攜性工具(比如說 Pad)操作，大幅提升了操控端的便利性。

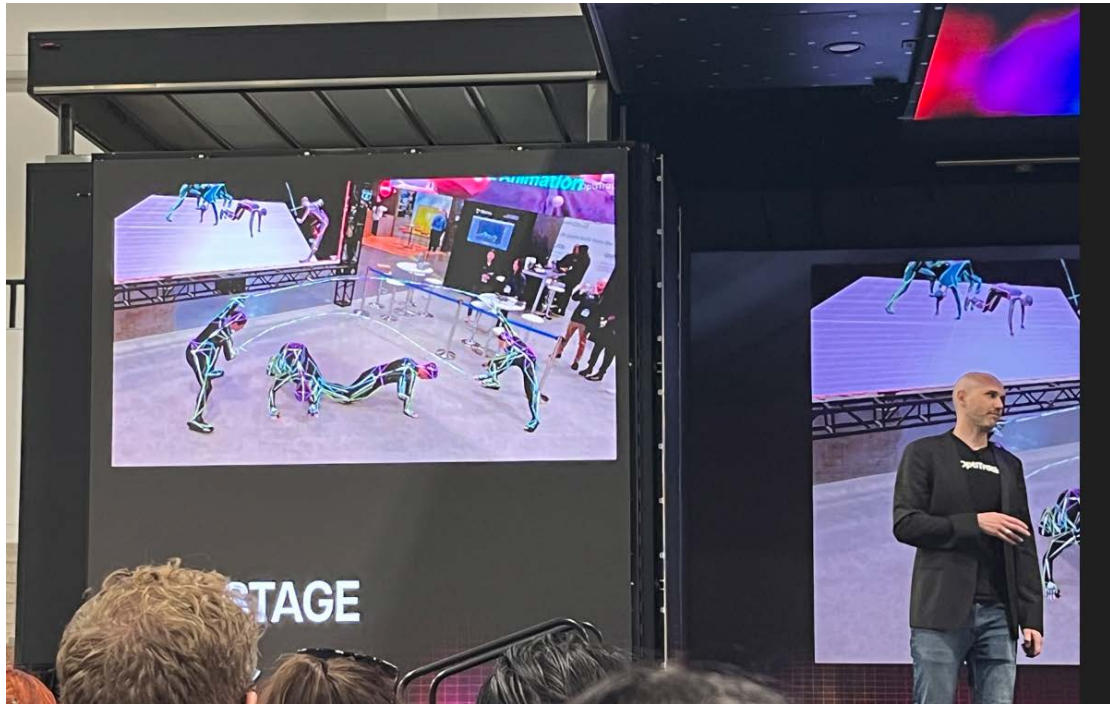




QUALISYS 動作捕捉系統控制端具移動性

## OptiTrack 動作捕捉系統

OptiTrack 動作捕捉系統，強調在多樣化的情境下皆有對應的攝影機種，並且具備快速追蹤功能，故此可以支援無人機進行較大區域的追蹤。另外還有桌面型的一體機，隨插即用，無需額外校正，只需一台機器就可以進行動作捕捉，可以對應如 YouTuber 獨立創作者或極小型空間的製作需求。



OptiTrack 設備可進行大範圍運用



OptiTrack 隨插即用一體機

## Markerless 系統

Markerless 動作捕捉系統不需要標記，而是依靠一組或多組不同視角之攝影機來追蹤人體的運動，再利用軟體或 AI 來合成人體三維關節與骨架空間座標資訊。優點是可以快速架設動作捕捉環境的便利性以及廣泛的場景適應性，但各廠商之不同的追蹤方法會產生



不同的結果，缺點是即時和最終數據誤差範圍往往比有標記的解決方案更大。

現場 Moverse 與 HP 均有展示相關方案，均可直接套入 3D 模型。但由於無法細部到手指追蹤，較不適合本次手語專案之需求。



Moverse Markerless 系統



HP AI Markerless 系統



## 結語

隨著 AI 技術的進步，各個領域及行業都受到了或多或少的影響，影視產業也是。這次是本會第一次參與 SIGGRAPH 此類電腦繪圖類的專業會議與展覽，深刻感受到電腦運算能力與 AI 的可能性與未來發展性。

公視現正與工研院合作，在預算有限的情況下嘗試以 AI 技術進行第一階段虛擬主播手語服務評估，此次在會場看到 NHK 發表的論文也是初步嘗試以 AI 生成手語，頗有英雄所見略同之感。

科技不斷進步推動各行各業的發展。尤其是虛擬製作、動作捕捉、電腦運算、AI 等的進步都將為電視台內容產製流程帶來突破性的改變。公視刻正進行數位轉型提供多平台服務，未來在內容產製、行政管理、數據分析、自動化服務等方面可進行更多的研討、評估與規劃，讓科技成為公視持續進步的動力與助力。

全文完