MIND REALITY

詳細解說影片



簡介

隨著數位內容快速成長，3D 內容創作在各領域的重要逐年提升，但傳統製作流程非常複雜，專業門檻成為大眾難以逾越的壁壘。

為此，我們提出一套結合 LLM 與混合實境（MR）的 3D 生成系統，使用者僅需透過語音描述，即可即時生成 3D 物件與對應的互動行為。

與此同時，本系統具備有場景理解能力，能依據環境主動提供生成建議，將 3D 創作轉化為直覺、輕鬆且很好上手的互動體驗。



方法與設計

Multi-Agent 架構

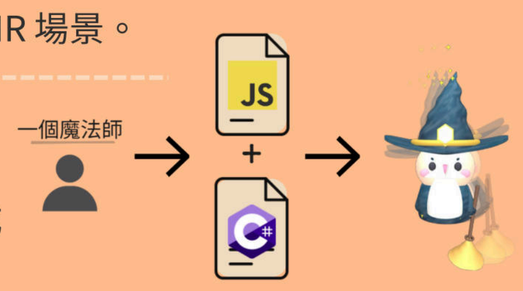
採用 Supervisor 模式，以 LangGraph 作為實作框架，由 Orchestrator 作為中央協調者，透過即時 AI 與使用者互動，理解意圖後將任務委派給子代理執行並整合回覆。各子代理具備獨立上下文，僅接收自身任務相關資訊，避免角色混淆與內容污染，提升系統穩定性。

物件生成

LLM 先依據使用者描述產生 Three.js 程式碼作為中介表示，兼具可讀性與可修改性；模型在算繪後，由 AI 透過多視角預覽圖評估缺陷以確保品質。通過評估後，最終匯入至 MR 場景。

動畫編排

提出 Planning-based Architecture
根據模型以及敘述生成動畫流程
在此架構下，動畫流程會被動作組成 Runtime 下編譯並加到物件上。



上下文工程

我們將系統提示、對話歷史、場景截圖、現有程式碼及工具描述整合為統一上下文，以提升 LLM 輸出之穩定性與可控性。透過系統提示規範 Real-time AI 對話行為；各 Agent 則依任務性質設計專屬提示詞，說明可用函式庫用法，並標示常見錯誤以降低幻覺發生。

視覺場景理解

Real-time AI 可以調用工具透過 HoloLens 2 的前置攝影機擷取場景影像，主動觀察環境的能力。在分析場景特點後，提出適合加入的 3D 物件建議。



成果展示



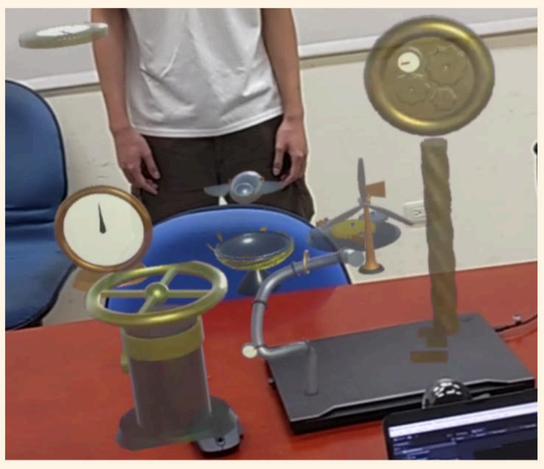
系統根據生成要求：機械鳥、齒輪組，理解語意並生成相關的機械主題金屬類物件，有機械懷錶、閥門、風車等。

圖中生成物皆掛載相應動畫

▶ 機械鳥依據語意生成動畫，翅膀可上下擺動，並呈現飛舞效果。
▶ 齒輪組依照物理產生連動動畫，不同齒輪可依尺寸呈現相反方向與不同轉速。



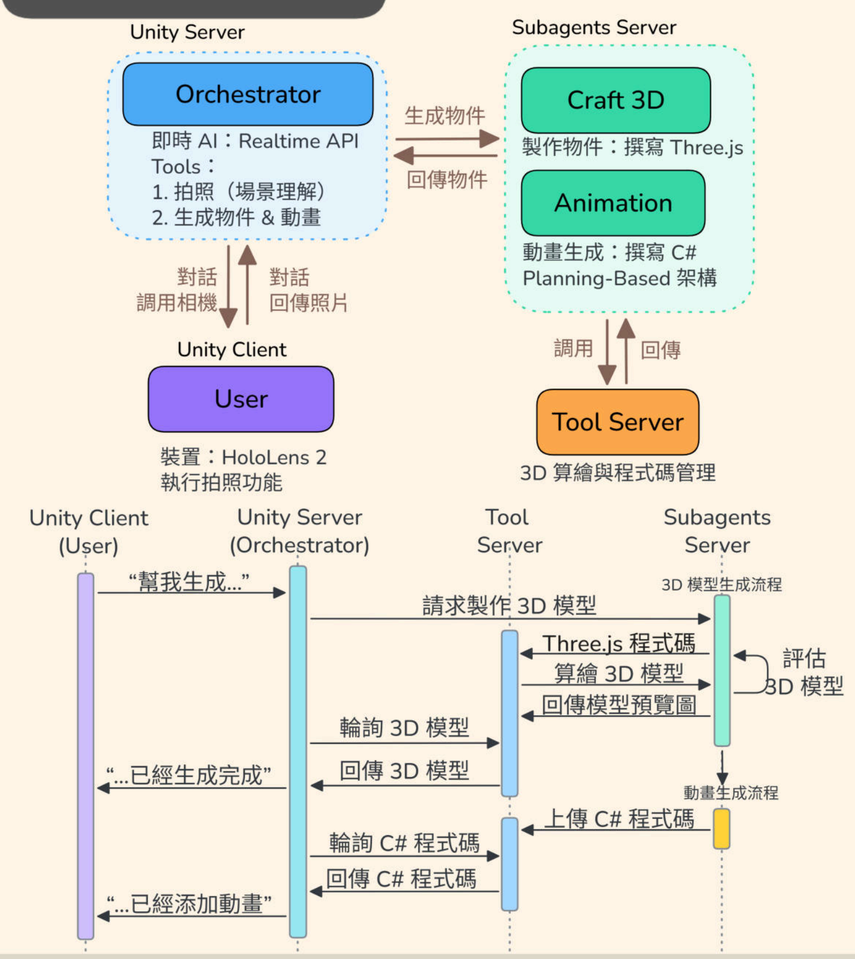
使用者1視角



使用者2視角

運用 LARR 平台透過計算物件與使用者視角的相對關係，實現極低延遲的多裝置同步顯示，支援多視角一致渲染，不同使用者可從各自角度觀察同一虛擬物件，並可對其進行拖拽等操作。

系統說明



結論

▶ 自然語言驅動 3D 生成 — 大幅降低創作門檻。
▶ MR 多裝置即時互動 — 提升沉浸式體驗。
▶ 提出語意驅動 + Planning-based 方法 — 保證生成品質。
▶ 場景理解能力 — 提供情境感知的生成建議。
▶ LARR 架構支援 — 確保低延遲與同步一致性。





REALTIME AI · MIXED REALITY · MULTI-AGENT · CRAFT 3D