

運用生成式人工智慧技術導入互動式擴增實境室內設計系統

張耀中，胡修銘，李振偉，賴盈勳*

摘要

隨著資訊科技的快速發展，科技導入在室內設計領域的應用也變得日益普及，為設計流程帶來了前所未有的便利與創意發揮空間。然而，傳統的設計輔助工具通常缺乏即時交互性，大多僅能提供靜態或單向性的設計建議，導致使用者在設計過程中難以直接參與或即時調整設計內容，限制了設計師的自主性與創意表現。本研究提出一個結合生成式人工智慧與擴增實境技術建立一互動式室內設計系統，目的在提供更直覺且沉浸式室內設計體驗。本研究系統架構包含生成式物件模型，其能根據使用者輸入的關鍵字自動生成各式室內設計所需的家具或裝飾物件等，並搭配擴增實境技術將生成之物件即時呈現於現實空間中。使用者可透過手勢控制或觸控操作，在擴增實境環境中自由地調整、旋轉與擺放生成物件，實現具互動性的設計體驗。此外，本系統具備即時預覽與修改功能，使用者可即時檢視空間中物件的擺放效果，並依需求快速調整設計布局，大幅減少傳統設計過程中繁瑣的重複嘗試與修正。本研究並進一步探討根據不同生成物件參數下對於整體物件生成影響與分析，降低整體物件生成幻覺與錯誤分析。本研究透過擴增實境沉浸式特性，使用者能從多角度觀察設計成果，有助於提前發現潛在的設計問題與空間不協調之處，提升整體設計品質與效率。

關鍵字：生成式人工智慧、擴增實境、互動設計、室內設計系統

張耀中，國立臺東大學資訊工程學系 教授、E-mail: ycc@nttu.edu.tw

胡修銘，國立臺東大學資訊工程學系、E-mail: 11201903@gm.nttu.edu.tw

李振偉，國立臺東大學資訊工程學系、E-mail: 11301904@gm.nttu.edu.tw

賴盈勳(通訊作者)，國立臺東大學資訊工程系所 教授、E-mail: yhlai@nttu.edu.tw

Applying Generative Artificial Intelligence for Interactive Augmented Reality Interior Design System

Yao-Chung Chang, Hu- Hsiu Ming, Li- Chen Wei, Ying-Hsun Lai*

Abstract

With the rapid advancement of information technology, the integration of technological innovations into the field of interior design has become increasingly prevalent, offering unprecedented convenience and creative flexibility to the design process. However, traditional design assistance tools often lack real-time interactivity, typically providing only static or unidirectional design suggestions. To address these challenges, this study proposes an interactive interior design system that integrates generative artificial intelligence (AI) with augmented reality (AR) technology, aiming to provide a more intuitive and immersive design experience. The system architecture features a generative object modeling component that automatically generates various furniture and decorative items based on user-inputted keywords. Users can intuitively manipulate the generated objects within the AR environment using gesture controls or touchscreen interactions, enabling free adjustment, rotation, and placement of objects to achieve an interactive design experience. Moreover, the system offers real-time preview and modification functionalities, allowing users to immediately visualize the spatial arrangement of objects and rapidly adjust the design layout as needed, thereby significantly reducing the iterative trial-and-error processes commonly associated with traditional design workflows. This study further investigates the impact of different generation parameters on the quality and accuracy of the produced objects, aiming to minimize generative hallucinations and reduce object generation errors. Consequently, the proposed system enhances both the overall quality and efficiency of interior design projects.

Keywords: Generative AI、Augmented Reality、Interactive Design、Interior Design System

Yao-Chung Chang, Professor, Department of Computer Science and Information Engineering, National Taitung University, Email: ycc@nttu.edu.tw

Hu- Hsiu-Ming, Department of Computer Science and Information Engineering, National Taitung University, Email: 11201903@gm.nttu.edu.tw

Li- Chen-Wei, Department of Computer Science and Information Engineering, National Taitung University, Email: 11301904@gm.nttu.edu.tw

Ying-Hsun Lai (Corresponding Author), Professor, Department of Computer Science and Information Engineering, National Taitung University, Email: yhlai@nttu.edu.tw

壹、研究背景

近年來，資訊與通訊技術的迅速發展對各產業領域帶來了深遠的變革，其中尤以室內設計領域為甚。先進數位工具的導入為設計流程提供了提升效率、精確度與創意探索的新契機。然而，儘管技術有所進步，傳統的電腦輔助設計系統仍普遍存在互動性不足的問題，多僅能提供靜態視覺化與單向性建議，致使設計師與最終使用者在設計環境中無法動態互動，限制了即時回饋、反覆試驗與創意自主的空間。在室內設計領域，AI 技術的引入使得設計流程得以更加高效、智慧與個性化。特別是生成式AI的崛起，為設計者提供了多樣且富有創意的設計選項，使得室內設計的可能性大幅擴展。Liu 和 Wang (2024) 提出「Mental-Gen」設計方法，透過腦機介面 (Brain-Computer Interface, BCI) 結合生成式 AI 模型，來協助使用者更有效地表達其室內空間設計意圖。此外，生成式AI也透過自然語言和視覺引導，讓設計方案的生成更加直觀且易於理解 (Le 等, 2023)。然而，目前多數 AI 設計輔助工具仍存在著明顯的局限性，特別是在互動性與即時性方面，往往無法滿足設計師或一般消費者對於設計自由度與即時回饋的需求，因此Jang 和Lee(2023) 提出一套名為 GAIA (Generative AI-enabled Interactive Architectural design) 的方法，結合大型語言模於建築資訊建模系統，以實現更互動式的設計過程。其研究結果顯示，該方法不僅能加強 AI 與設計師間的互動效率，也突顯了生成式語言模型在建築設計協作中的潛力。另一方面，擴增實境技術的興起，也為室內設計領域帶來不同之契機。透過擴增實境技術，使用者能夠將虛擬物件直接投影至現實空間中，提供沉浸式的視覺與操作體驗，進一步提升設計過程的直觀性與參與感 (Shum 等, 2023)。然而，目前市場上的設計系統大多僅將 AR 作為展示工具，缺乏與 AI 系統的深度整合，導致設計流程仍然存在互動性不足與修改不便的問題。

基於上述挑戰與需求，本研究提出一套結合生成式人工智慧與擴增實境技術之互動式室內設計系統，其透過即時、直覺且沉浸式的互動方式，提升使用者參與度、創意表現與設計效率。系統核心包含一套基於關鍵字驅動的生成式人工智慧模型，能依據使用者輸入的關鍵字，自動生成各式家具與裝飾物件，並動態對應個別設計圖。此系統不僅具自動生成設計物件的能力，亦能透過擴增實境環境提供使用者即時擺放、調整與修改物件的功能。生成的物件將藉由AR視覺化技術即時呈現在現實空間中，使用者可透過自然互動方式，如手勢控制與觸控操作，自由操控、旋轉與調整物件位置，實現高度互動的設計體驗。本研究亦探討生成參數對於物件品質的影響，並致力於降低生成過程中可能出現的幻覺與錯誤，確保設計成果在功能性與美學上的可靠性。本研究期望透過此系統的建置，解決傳統 AI 輔助設計工具所面臨的互動性不足問題，並提供設計師與消費者一個易於操作、靈活且具備即時回饋的設計平台。最終，本研究將有助於推動室內設計邁向智慧化、個性化與高度參與性的發展方向，為未來設計產業注入新的創新動能。

貳、文獻探討

2.1. 生成式人工智慧

生成式人工智慧 (Generative Artificial Intelligence, GAI) 已成為能自動生成數位內容 (如影像、音樂與自然語言文本) 的關鍵技術。GAI 的主要目標在於提升內容創作的效率與可及性, 使高品質內容得以更快速地產出。隨著資料集規模與模型架構的持續擴展, GAI 系統能夠從使用者輸入中提取並理解意圖資訊, 進而學習更全面且貼近真實世界的資料分佈, 產生更具真實感與精緻度的內容 (Cao 等, 2023)。與此同時, 「人工智慧生成內容」(AI-Generated Content, AIGC) 概念亦逐漸受到重視。AIGC 指的是藉由先進 GAI 技術自動產生、非由人類直接創作的內容, 能顯著簡化內容生產流程, 在短時間內大量生成各類型的內容。例如, 由 OpenAI 開發的語言模型 ChatGPT, 能有效理解並有意義地回應人類語言輸入; 而 DALL-E 則能根據簡單的文字描述, 在數分鐘內生成獨特且高品質的圖像 (Cao 等, 2025; Liu 等, 2024)。

近年來, AI 技術迅速發展, 並廣泛應用於藝術、廣告與教育等多個領域 (Anantrasirichai & Bull, 2022; Kietzmann 等, 2018)。其中, AI 繪圖工具如 Stable Diffusion 與 ComfyUI 已被廣泛應用於快速生成高品質的圖像資源。這些工具不僅豐富了教學內容, 亦有效提升了學習效率與互動性。透過即時的視覺回饋與創意探索功能, 生成式人工智慧工具有助於營造更生動、有趣且具參與感的學習環境, 促進學習者的積極投入與學習成果的提升 (羅禎偉, 2024; 徐志溢, 2023)。而在藝術設計方面, 生成式人工智慧可透過大規模資料學習不同的設計風格與元素, 快速生成多樣化的設計方案, 顯著提升設計效率與創造力 (Zhou & Wang, 2024)。此外, 生成式人工智慧結合自然語言處理技術, 使用者僅需輸入簡單的文字描述, 即可自動生成具體的設計圖像與方案, 此舉大幅降低了設計門檻, 並促進了使用者參與設計流程的意願與可能性, 進一步擴展了創意產出的範疇。

2.2 擴增實境

擴增實境 (Augmented Reality, AR) 作為一種新興體驗技術, 透過將電腦生成的虛擬內容與特定地點或活動結合, 強化了使用者對現實世界的感知與互動體驗。近年來, 隨著行動裝置的普及與運算能力提升, AR 應用已變得更加便攜並廣泛部署於行動平台上。初期的 AR 內容觀看方式多透過網路攝影機或行動裝置掃描 QR 碼, 進行網頁或行動應用程式的互動, 而後續技術演進則引入了頭戴式顯示器 (Head-Mounted Display, HMD) 作為觀看介面。使用者配戴 HMD (例如 AR 眼鏡) 時, 能夠在顯示螢幕上同時觀看虛擬內容與真實環境, 或透過連接相機即時映射真實場景 (Yuen et al., 2011)。目前, AR 技術已廣泛應用於醫療保健、餐旅、教育、旅遊、文化、軍事、建築、設計、工程、遊戲與娛樂等多個領域, 展現出其跨領域整合與推動產業創新的潛力 (Hsieh, 2018)。

在設計領域中，AR 技術可讓使用者將虛擬設計元素即時融合至現實環境中，創造沉浸式的互動體驗。使用者得以即時觀察、調整虛擬物件的位置、尺寸與其他屬性，直觀且高效地實現設計構想。此特性不僅顯著提升了專業設計師的工作效率，同時亦降低了一般使用者參與設計過程的門檻，使其能夠以更具體、真實的方式參與空間規劃與設計決策。近年來，業界與學界亦開始探索將 GAI 與 AR 技術結合，期望開發出更具互動性與直覺性的設計輔助工具。透過將 GAI 自動生成的設計元素即時導入 AR 環境，使用者能於真實空間中迅速完成設計布局，並即時進行調整與最佳化 (Zhang et al., 2024)。然而，目前此領域之研究與應用仍處於初步階段，大多數現有系統在互動性、即時反饋與自適應調整等方面仍存在不足，尚無法完全滿足使用者對於即時參與與自由設計操作的需求。

基於上述現況，如何有效整合生成式人工智慧與擴增實境技術，建立一套真正具備即時互動性與沉浸感的室內設計系統，成為值得深入探討的重要課題。因此，本研究將在系統性分析 GAI 與 AR 技術的整合模式，提出更完善的系統架構與技術方法，以提升室內設計過程中的互動性、即時性與創造力表現，並為未來智慧設計應用提供具體可行的技術參考與發展方向。

參、研究方法

一、系統架構

本研究將主要探討 GAI 與 AR 技術整合於室內設計應用服務，在 GAI 部分主要採用一開源生成式圖像生成工具 Stable Diffusion 作為整體圖像產生應用。Stable Diffusion 是一個深度學習生成式模型，可以根據文字描述於潛在空間中進行噪聲預測與邊解碼建立圖像或物件。本研究搭配採用 ComfyUI 使用者圖像介面來管理整體 Stable Diffusion 自定義的生成流程。ComfyUI 支援多種 Stable Diffusion AI 模型並可透過流程管理方式建立自己想要風格與透過提示詞來產生想要圖像。而在 AR 部分則主要採用 Unity 套件在行動裝置部分建立使用者相關互動介面與功能服務，將運用 Unity 行動裝置應用程式串接 ComfyUI 提供之 API，供使用者從語音或文字輸入框輸入預想生成物件的提示詞，提送至 ComfyUI 生成 3D 模型並儲存至指定資料夾。系統運用 System watcher 來監聽該資料夾有無檔案新增，若有則匯入至 Unity 並加載至場景中，使用者再將 AR 眼鏡連結到行動裝置上，AR 眼鏡即可呈現生成的 3D 模型，整體系統流程圖如圖 1 所示。

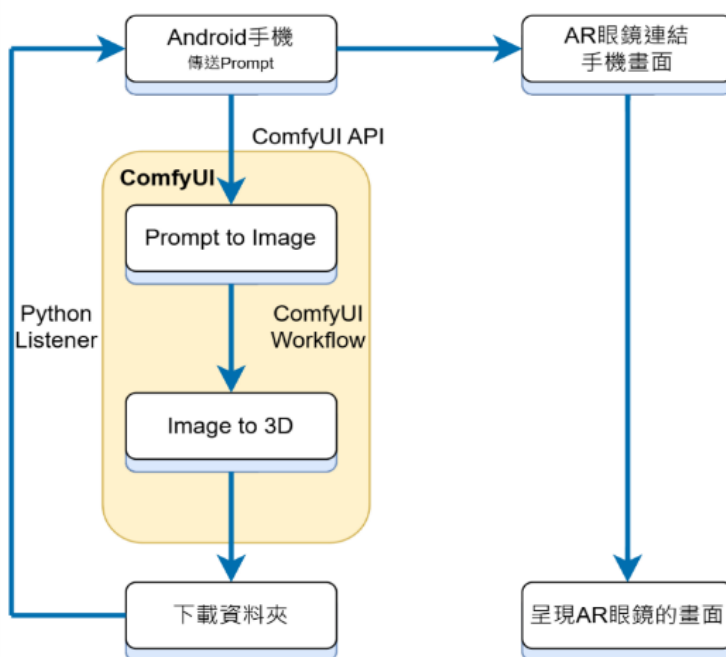


圖 1. 系統流程圖

二、生成式 2D 物件模組

首先，本研究利用 ComfyUI 建立一個基於生成式 AI 的 2D 物件生成流程。該流程以自然語言提示 (Prompt) 為起點，並透過文生圖 (Text-to-Image) 模型將描述性文字轉換為對應的圖像。整體生成模型流程中將主要包含以下功能區塊

- (一) Checkpoint 加載器：主要用於載入 Stable Diffusion 的主模型提供後續圖片生成風格參考，而 Checkpoint 加載器其可串接包含模型、CLIP 模組與變分自編碼器 (Variational Autoencoder, VAE)。
- (二) CLIP 模組：該模組將主要用於理解提示詞語意並將提示詞語意轉換成為 Stable Diffusion 所需採用語意向量，以便讓模型理解所需生成圖像。
- (三) 空 Latent 圖像模組 (Empty Latent Image Node)：該模組主要用於建立一個潛在向量空間 (Latent Space)，讓圖像生成於該設定潛在向量空間降低整體運算效能。
- (四) K 採樣器：K 採樣器負責生成潛在圖像 (Latent Image) 顯示，控制生成過程中的隨機種子、cfg 比例與取樣演算法等相關參數。其根據初始潛在圖像、文字提示語意向量與採樣演算法生成最後潛在圖像。
- (五) VAE 解碼：該模組則負責將最後潛在圖像解碼還原至圖像空間，產出對應的視覺內容。VAE 同時可透過編碼將圖像空間圖片編碼至潛在圖像空間，讓其生成模型可進行參考。

本研究流程以 Checkpoint 加載器載入經訓練完成的模型檔案（本案例中為 realvisxl），並連接至 CLIP 模組以解析語意。主要提示詞經由 CLIP 文本編碼模組進行語意嵌入，並同時接入負面提示詞以過濾掉不必要的影像特徵。同時，本研究透過空 Latent 圖像模組初始化圖像的介面，設定最後圖像解析度為 512 x 512，並將其與語意條件一起傳入 K 採樣器。K 採樣器則根據設定條件進行生成潛在空間圖片，圖片完成後再透過 VAE 解碼將潛在表示還原為最終圖片並顯示使用者可見之圖像檢視。整題 ComfyUI 生成流程如圖 2 所示。

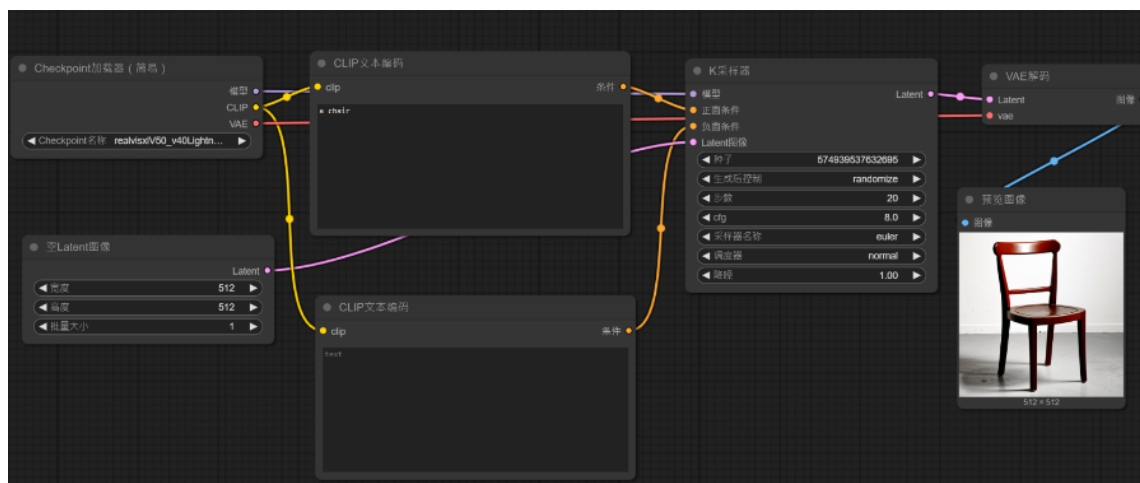


圖 2. ComfyUI Text to Image 流程

三、2D 圖像轉換 3D 物件

在完成 2D 圖像生成之後，利用 ComfyUI 進一步將靜態圖像轉換為可用於擴增實境的 3D 模型，流程如圖 3 所示。此流程整合了 Stable-Fast-3D 節點，將生成圖像透過 Stable-Fast-3D 建構對應的三維網格結構，以利擴增實境應用。因為生成式 3D 模型容易將 2D 圖像的背景和主體物件做混和再生成，因此首先透過 Image Remove Background (rembg) 模組，去除 2D 圖像背景，只保留乾淨的主體，以利後續生成乾淨的 3D 模型。背景去除後的圖像，會被傳入圖像轉換為遮罩的相關 模組，圖 3 範例是使用 alpha mask，作為 3D 建模時的遮罩參考。

接著，該遮罩與圖像一併輸入至 Stable Fast 3D Sampler 模組進行 3D 模型構建。在此步驟中，系統會根據輸入參數如 foreground_ratio、texture_resolution、remesh 及 vertex_count 來控制生成結果的細緻程度與模型大小。以圖 3 為例，設定 foreground_ratio 為 0.90、texture_resolution 為 2048、remesh 為 none，並將 vertex_count 設定為 20000，以確保模型細節完整呈現。生成後的 3D 網格模型可即時預覽，透過 Stable Fast 3D Preview 節點檢視生成結果之視角與貼圖效果，並可透過 Stable Fast 3D Save 節點將結果匯出成標準 3D 格式檔案(glb)，進行後續 AR 系統整合、3D 列印或動畫應用。

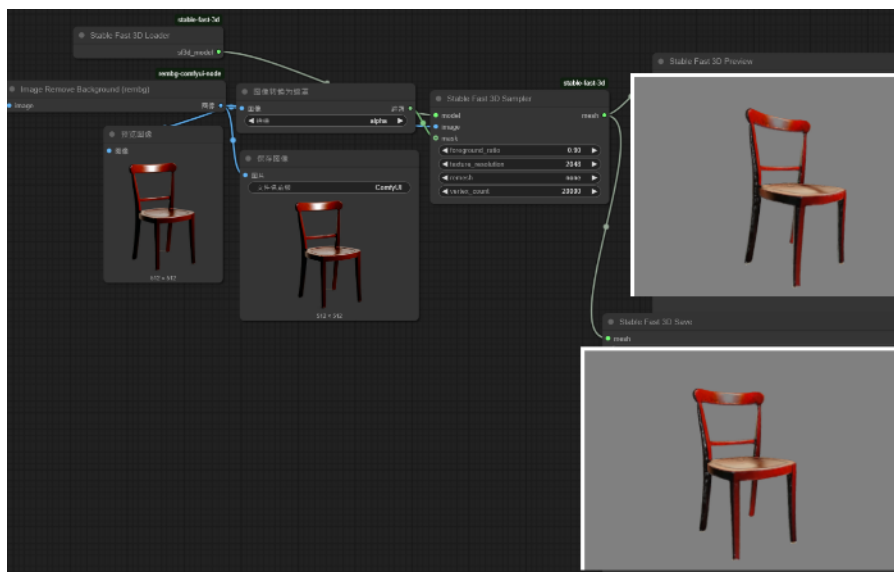


圖 3. ComfyUI Image to 3D 流程

四、API 串接與擴增實境應用

將利用 ComfyUI 製作的 workflow 打包成 API，Unity 使用此 API 即可做到同樣效果，接著透過指定 3D 模型輸出路徑，將生成之 3D 模型自動儲存至輸出路徑的資料夾。此輸出機制為 Unity 前端應用程式抓取 3D 模型呈現給使用者的資料來源。

在 Unity 端，系統利用 C# 語言中的 FileSystemWatcher 元件建立一套事件驅動的即時資源監控架構。該元件可持續偵測指定資料夾中新增的模型檔案，一旦偵測到新資源生成，便會自動觸發事件，將 3D 模型載入與做渲染，並將模型呈現出來，如圖 4 所示。最終，此流程可打包成 APK 應用程式，供使用者下載至 Android 行動裝置之中，並配合 AR 裝置進行顯示。使用者透過行動裝置即可產製 3D 模型，並將其利用 AR 裝置即時映射至真實場域中，呈現沉浸式的擴增實境體驗。

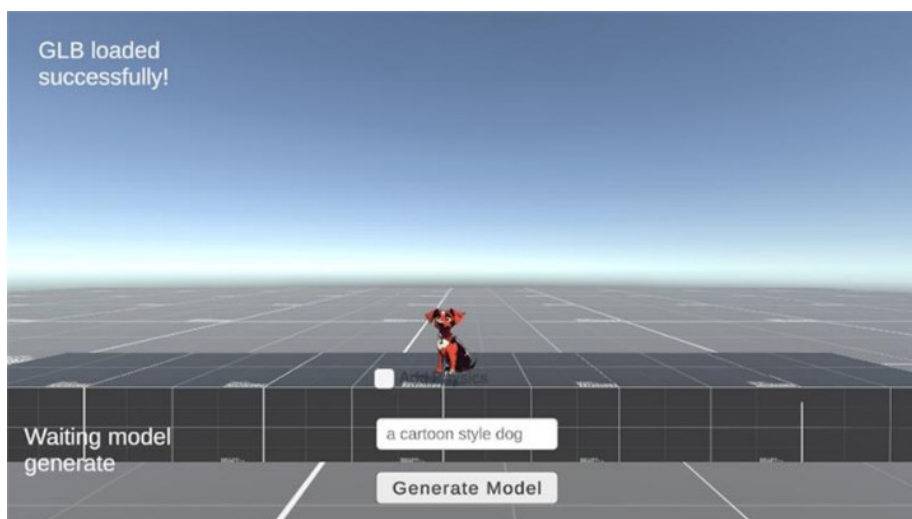


圖 4. 偵測是否有輸出新的 3D 模型，並將 3D 模型呈現出來

肆、實驗成果

一、生成式家具功能實現

本研究以室內設計為核心，開發了一套創新的 AI 驅動三維空間佈局生成系統。該系統透過生成式 AI 技術，能夠根據使用者的自然語言描述，快速生成客製化的室內設計所需物件，並且使用者能配戴 AR 眼鏡將 3D 模型呈現出來，供使用者做設計使用。本系統採用 AI 生成技術，將使用者的自然語言需求轉化為具體的三維空間佈局。例如，當使用者輸入「舒適的書房空間，包含木質桌子、木質椅子和木質書櫃」時，系統能夠即時生成符合描述的室內空間模型，如圖 5 所示。



圖 5. 使用者可以配戴 AR 眼鏡將 3D 模型呈現在現實生活中







二、生成式參數實驗影響

本研究進一步探討對於生成式模型中相關參數對於整體物件生成影響，本研究將主要採用以下參數進行分析設定討論。

- **Foreground_ratio**: 該參數控制圖片中哪個區域是要進行建模的區塊，參數依據是根據遮罩 (mask) 來判定，數值介於 0~1 之間，如果參數太低會導致只生成部分區域，例如只生成椅腳，太高則會因為太多雜訊導致生成的模型出現異常。
- **Texture_resolution**: **Texture_resolution** 為控制生成模型貼圖的細節度。太低會導致圖片模糊，太高則會大幅增加生成時間。
- **Remesh**: 該參數指是控制是否調整、重新構建潛在圖像 (latent image) 或實際圖像的結構資訊，以利後續的重繪 (inpainting)、合成、解析度轉換或風格融合。
- **Vertex_count**: 該參數則為控制模型的複雜度與細緻度，即用多少個頂點構成整個 3D 網格，在本研究採用 Stable-Fast-3D 模型目前版本最高只能設定 20,000。

整體相關實驗數據與生成結果如表 1 所示，經實驗各參數的影響結果，影響最大為 vertex_count，如果參數過低會導致模型複雜度和細緻度大幅下滑，因此盡量保持設定最高 20,000，而第二影響最大為 Foreground_ratio，因為這會影響訓練內容，參數過低只會生成圖片部分區塊的模型，參數過高則會把圖片雜訊全部都一起拿去生成模型。Texture_resolution 和 Remesh 影響會小一點，主要是影響最終模型的精緻度如何。

表 1. 生成式參數實驗結果

| | | | | | | |
|---------------------------|---|---|---|--|---|---|
| Foreground_ratio | 0.90 | 1.00 | 0.20 | 0.90 | 0.90 | 0.90 |
| Texture_resolution | 2,048 | 2,048 | 2,048 | 512 | 2,048 | 2,048 |
| Remesh | none | none | none | none | triangle | none |
| Vertex_count | 20,000 | 20,000 | 20,000 | 20,000 | 20,000 | 100 |
| 3D 模型 |  |  |  |  |  |  |

結論

本研究成功整合 ComfyUI 生成式 AI 技術與 Unity 環境，建立一套從自然語言輸入開始直至 3D 模型即時生成完整流程，並將流程打包成 API 供 Unity 做使用，將生成出來了 3D 模型應用在擴增實境。透過 Stable Fast 3D 等模型的串接與資源監控機制的建構，使用者僅需輸入描述語句，即可一鍵完成模型生成並於行動裝置中配合擴增實境進行視覺化呈現，實現高效率 3D 模型的製作。

本系統有找出相對穩定的參數與模型配置，但仍不是最好的結果，如模型生成精度、背面結構還原、貼圖細節還原與載入效能等問題。針對這些限制，後續將持續優化生成參數的設定邏輯（如 cfg、step、seed）、進行提示詞模板化開發，以及更多模型加入訓練過程。目前參數與模型設定皆是設定好不可更改，因此未來會進一步發展動態參數調整與自動優化機制，以提升整體生成品質與使用體驗。

本研究未來將持續朝向三個主要方向進行發展：一為提升語意準確性與視覺保真度，二為優化互動操作與即時體驗，三為拓展應用情境至教育、數位展演等多元領域，而不是僅使用在室內設計領域。藉由技術不斷迭代與創新，本系統未來期望可以持續降低 3D 創作門檻，為大眾提供更智慧且高效的創作工具，推動生成式 AI 技術不僅在空間與內容創作領域中的落地應用，也能運用在更多的領域上。

參考文獻

- 羅禎佺 (2024)。AI 工具 ComfyUI 輔助 2D 人型角色動畫應用於各景別之研究。〔碩士論文。嶺東科技大學〕臺灣博碩士論文知識加值系統。
<https://hdl.handle.net/11296/uhj3zm>。
- 徐志溢 (2023)。基於 Stable Diffusion 模型和 LoRA 優化訓練生成人臉表情。〔碩士論文。國立高雄科技大學〕臺灣博碩士論文知識加值系統。
<https://hdl.handle.net/11296/j597b4>。
- Anantrasirichai, N., & Bull, D. (2022). Artificial intelligence in the creative industries: a review. *Artificial intelligence review*, 55(1), 589-656.
- Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt. *arXiv preprint arXiv:2303.04226*.
- Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P., & Sun, L. (2025). A survey of ai-generated content (aigc). *ACM Computing Surveys*, 57(5), 1-38.
- Hsieh, M. C., & Lee, J. J. (2018). Preliminary study of VR and AR applications in medical and healthcare education. *J Nurs Health Stud*, 3(1), 1.
- Liu, J., Wang, Z., Xie, J., & Pei, L. (2024). From ChatGPT, DALL-E 3 to Sora: How has Generative AI Changed Digital Humanities Research and Services?. *arXiv preprint arXiv:2404.18518*.
- Kietzmann, J., Paschen, J., & Treen, E. (2018). Artificial intelligence in advertising: How marketers can leverage artificial intelligence along the consumer journey. *Journal of Advertising Research*, 58(3), 263-267.
- Yuen, S. C. Y., Yaoyuneyong, G., & Johnson, E. (2011). Augmented reality: An overview and five directions for AR in education. *Journal of Educational Technology Development and Exchange (JETDE)*, 4(1), 11.