國立台東大學

資訊工程學系碩士班

碩士論文

智慧眼鏡結合多模態生成式AI之效能與應用特性研究：以室內設計輔助為例

A Study on Performance and Application Characteristics of Smart Glasses Combined with Multimodal Generative AI: Using Interior Design Assistance as a Case Study

研 究 生 : 胡修銘

指導教授 : 賴盈勳 博士

狄旻

目錄

[第一章 緒論 3](#_Toc181787650)

[1-1 研究背景與動機 3](#_Toc181787651)

[1-2 研究目的 5](#_Toc181787652)

[1-3 研究流程 5](#_Toc181787653)

[1-4 研究限制與範圍 5](#_Toc181787654)

[第二章 文獻探討 5](#_Toc181787655)

[2-1 AR眼鏡 5](#_Toc181787656)

[2-2 ComfyUI 5](#_Toc181787657)

[2-3 生成式AI 5](#_Toc181787658)

[2-4 選用的ComfyUI模型 5](#_Toc181787659)

[2-4-1 TripoSR 5](#_Toc181787660)

[2-4-2 Flux 5](#_Toc181787661)

[2-4-3 SF3D 5](#_Toc181787662)

[2-4-5 SDXL 5](#_Toc181787663)

[第三章 流程設計與架構 6](#_Toc181787664)

[3-1 AR眼鏡前端與使用者介面 6](#_Toc181787665)

[3-2 ComfyUI 與WorkFlow 設置 6](#_Toc181787666)

[3-3 3D物件生成模型 6](#_Toc181787667)

[3-4 Unity API串接方式 6](#_Toc181787668)

[第四章 研究結果 7](#_Toc181787669)

[4-1 AR眼鏡選用 7](#_Toc181787670)

[4-2 3D生成模型選擇 7](#_Toc181787671)

[4-3 節點參數設置 7](#_Toc181787672)

[4-4 模型生成速度 7](#_Toc181787673)

[第五章 結論與建議 8](#_Toc181787674)

[5-1 合適的AR眼鏡款式 8](#_Toc181787675)

[5-2 選用模型的差異及生成時間突破 8](#_Toc181787676)

[5-3 參數或額外功能性的節點設置 8](#_Toc181787677)

第一章 緒論

1-1 研究背景與動機

隨著資訊科技的快速發展，智慧眼鏡作為新一代的人機互動介面，正逐漸改變人們與空間互動的方式，並在遊戲、醫療保健、教育、軍事國防等領域展現出廣泛的應用前景。根據市場研究機構 Orion Market Research 的預測，全球 AR/VR 智慧眼鏡市場預計在 2023 年到 2030 年間，將以 8.1% 的複合年增長率持續成長，由於「隨看即得」的特性，讓智慧眼鏡可以在各個行業帶來革新應用。

生成式 AI 是人工智慧領域中一個快速發展的分支，其核心概念在於\*\*讓 AI 不僅僅是分析和預測現有資料，更能創造新的資料，\*\*是一種技術，它更像是一種賦予 AI 創造力的新方法，讓 AI 能夠像人類一樣進行創作。 生成式 AI 模型通常具備多模態能力，可以處理文字、圖像、聲音、程式碼等多種類型的資料，在跨媒體內容的生成和理解上擁有驚人潛力。

這個技術的發展建基於三股主要力量：

資料工程的進步: 現代化的資料工程發展出大數據治理技術，能夠收集和清理網路上大量的數據，為訓練 AI 模型提供了可靠的資源。

電腦算力的提高: 以 NVIDIA 為首的科技大廠投入加速器硬體技術研發，提供超高速的叢集運算能力，為生成式 AI 模型的訓練和推理提供必要的運算資源。

AI 演算法的突破:深度學習技術取得重大突破，例如生成對抗網路 (GAN)、擴散模型 (Diffusion Model) 和採用注意力機制的變換器 (Transformer) 網路，這些技術的發展讓 AI 能生成更逼真和充滿創意的內容，並在自然語言處理領域取得巨大成功。

生成式 AI 的興起，為各行各業帶來了新的突破和應用場景，例如：

**智慧 AOI 檢測**: 生成式 AI 可以生成各種可能的缺陷圖像，擴增訓練數據集，提高模型的泛化能力和檢測準確率，特別是在需要大量影像資料的智慧檢測辨識上，能有效縮減模型產生的時程。

**藝術與設計**: 生成式 AI 能快速產出圖像、轉換藝術風格或提供配色建議，提升藝術家或設計師的工作效率，甚至改變傳統的設計流程。

**影視娛樂**: 生成式 AI 被應用於影視特效、遊戲開發等領域，透過深偽 (Deepfake) 或虛擬人技術，創造出逼真的視覺效果和互動體驗。

**教育**: 生成式 AI 可輔助教學，例如批改作業、製作學習歷程檔案或多媒體教材等。

**醫療保健**: 生成式 AI 被用於醫學影像處理、疾病診斷、新藥研發等任務。

在人工智慧發展的潮流中，ComfyUI 作為一個創新的開源框架，ComfyUI 是一款開源的人工智慧工作流程工具，最大特色就是採用節點式的視覺化介面。讓使用者可以用拖拉的方式，把不同的 AI 功能組合在一起，特別適合用來生成或修改圖片。

功能支援方面，ComfyUI 展現出優異的擴充性與相容性。系統支援多種主流擴散模型，包含 Stable Diffusion 1.x、2.x 系列及 SDXL 等，同時擴展至影片生成（Stable Video Diffusion）和音訊處理（Stable Audio）等多媒體領域，在工作流程管理層面，ComfyUI 提供了完整的工作流程儲存與讀取機制，並具備模型和檢查點管理功能。系統整合的即時效能監控與除錯工具，有效協助研究者追蹤與優化實驗流程。這些功能的整合不僅提升了研究效率，更確保了實驗過程的可追溯性與重現性

本研究旨在探討如何在智慧眼鏡等資源受限設備上有效運行生成式 AI 模型，ComfyUI 展現出在學術研究與實務應用上的諸多優勢。其高度客製化的使用者介面允許研究者根據個別需求進行調整，完善的除錯機制則有助於實驗過程中的問題排除。基於上述特性，選用ConfyUI做為生成式AI開發工具，並探索其在領域的應用。相信隨著技術的進步和應用生態的完善，生成式 AI 與智慧眼鏡的結合將為人們帶來更智能、更便捷、更富創造力的未來生活體驗。

1-2 研究目的

1-3 研究流程

1-4 研究限制與範圍

第二章 文獻探討

2-1 AR眼鏡

2-2 ComfyUI

2-3 生成式AI

2-4 選用的ComfyUI模型

2-4-1 TripoSR

2-4-2 Flux

2-4-3 SF3D

2-4-5 SDXL

第三章 流程設計與架構

3-1 AR眼鏡前端與使用者介面

3-2 ComfyUI 與WorkFlow 設置

3-3 3D物件生成模型

3-4 Unity API串接方式

第四章 研究結果

4-1 AR眼鏡選用

4-2 3D生成模型選擇

4-3 節點參數設置

4-4 模型生成速度

第五章 結論與建議

5-1 合適的AR眼鏡款式

5-2 選用模型的差異及生成時間突破

5-3 參數或額外功能性的節點設置