

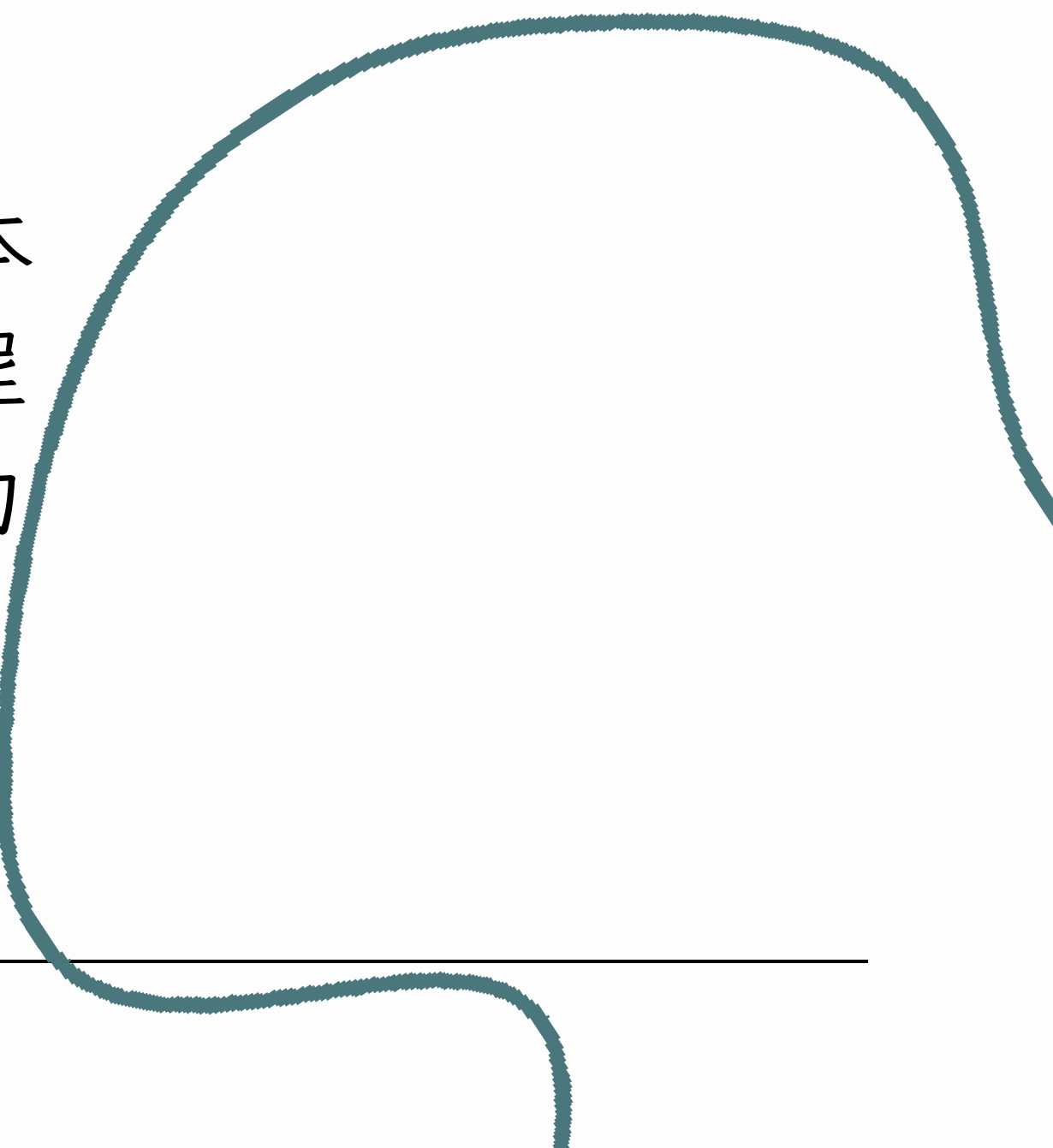


AI虛擬主播技術應用於 永續文化行銷

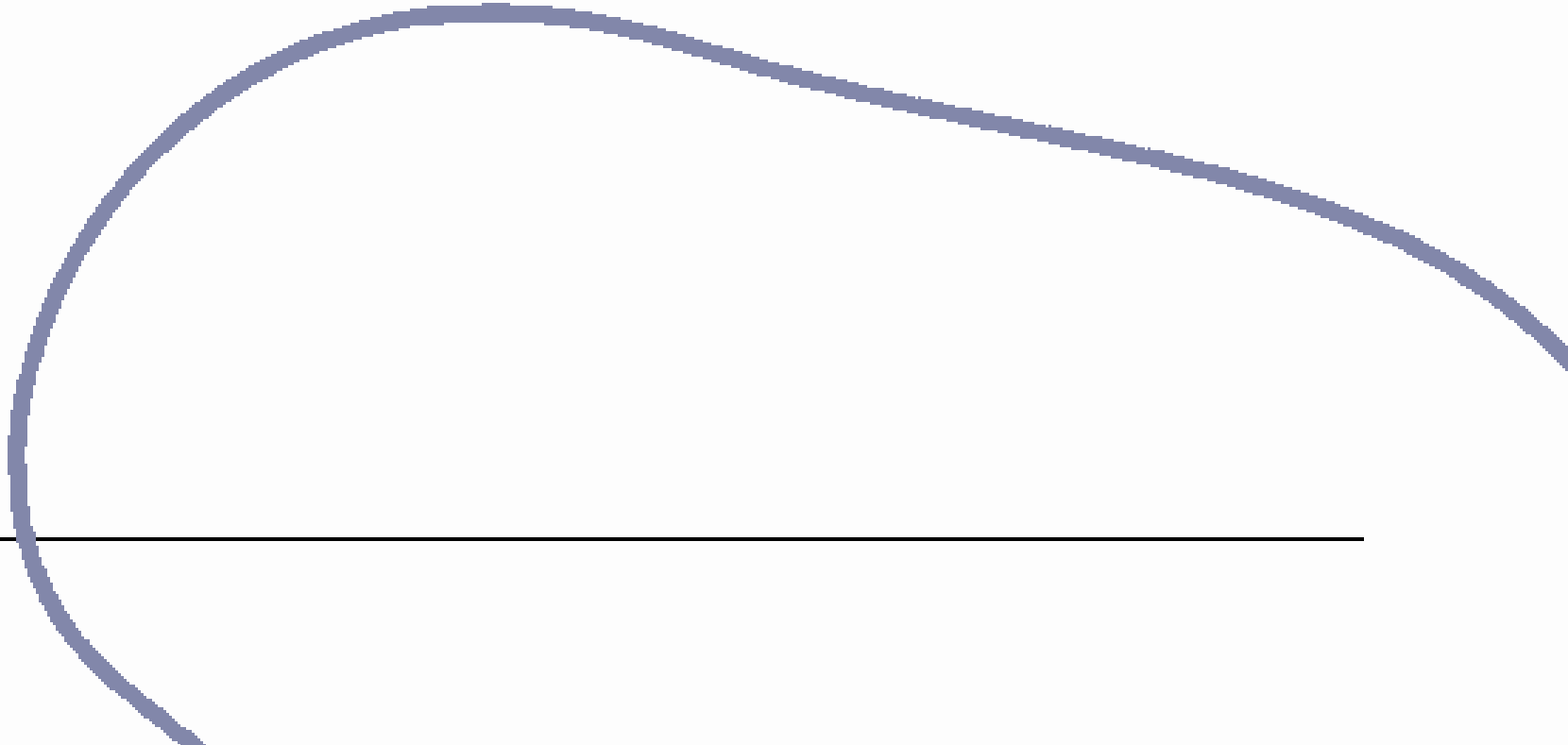
目錄

01.	研究背景與動機	07.	SDGs與永續文化	13.	案例比較：Neuro-sama
02.	研究目的	08.	系統設計模組化	14.	案例比較：神霄股份有限公司
03.	研究流程	09.	角色設計流程	15.	成果評估與反思
04.	技術介紹：Live2D	10.	聲音模型建構	16.	研究貢獻與限制
05.	技術介紹：GPT-SoVits	11.	整合與輸出	17.	未來發展建議
06.	技術介紹：VTube Studio	12.	實作案例展示	18.	結語與Q&A

研究背景與動機

- AI技術讓影音創作變得更快速與低成本
 - 華語系老師錄製影片困難，需簡化流程
 - 虛擬主播技術成趨勢，具文化傳播潛力
- 

研究目的

- 建立一套文化導向AI虛擬主播製作流程
 - 降低教師與創作者的時間與人力負擔
 - 推動永續文化
-
- 

研究流程

1. 虛擬角色設計 (Live2D)
 2. 聲音模擬訓練 (GPT-SoVits)
 3. 整合與輸出 (VTube Studio)
 4. 實作案例應用與評估
-

技術介紹：Live2D

- 將2D圖層進行拆件與動態建模
 - 實現眨眼、嘴型、表情、肢體動作
 - 高效呈現2D動畫風格虛擬角色
-

技術介紹：GPT-SoVits

- 基於少量語音樣本進行AI仿聲

- 可輸入文字自動合成擬真人聲音

- 調整語速與語調，應用彈性高

技術介紹：VTube Studio

虛 擬 角 色 嘴
型、頭 部、
表 情 即 時 控
制

與聲音輸出同步

支 援 錄 影 與
直 播 場 景

SDGs與永續文化

- SDG 3：健康與福祉（減輕錄音負擔）
- SDG 4：優質教育（提高學習平等）
- SDG 11：永續城市（文化多樣性）
- SDG 12：負責任消費（數位可複用系統）



研究生對於永續文化之理解

- 讓人類的發展走的更長遠的精神和行為文化
 - 永續發展的文化
 - 精神與行為性：文化有他一定的精神及主要行為方向。
 - 轉譯與再生性：文化必須被當代理解、轉化與創造，維持再生能力。
 - 跨語境與共融性：內容應具有多語言、多文化接觸與理解能力。
- 文化行銷的實踐方式
- 低參與門檻：製作流程簡化，非專業者亦能使用。
- 模組化與可複製性：建構出可快速套用至他類主題的模板。
- 教育應用友善：支援課堂教學、文化導覽與公開推廣活動。



系統設計模組化

- Live2D角色建模模組
 - 聲音模擬與合成模組
 - 整合控制與輸出模組
 - 對應永續與文化需求調整彈性高
-

角色設計流程

- 建構「中式端莊」的虛擬角色形象
- 拆件：髮絲、五官、服裝、裝飾
- 設定表情動作參數與動畫連動



角色設計流程

參考資料收集、草稿設計



最終定案

角色設計流程

繪製Live2D專用分層插圖

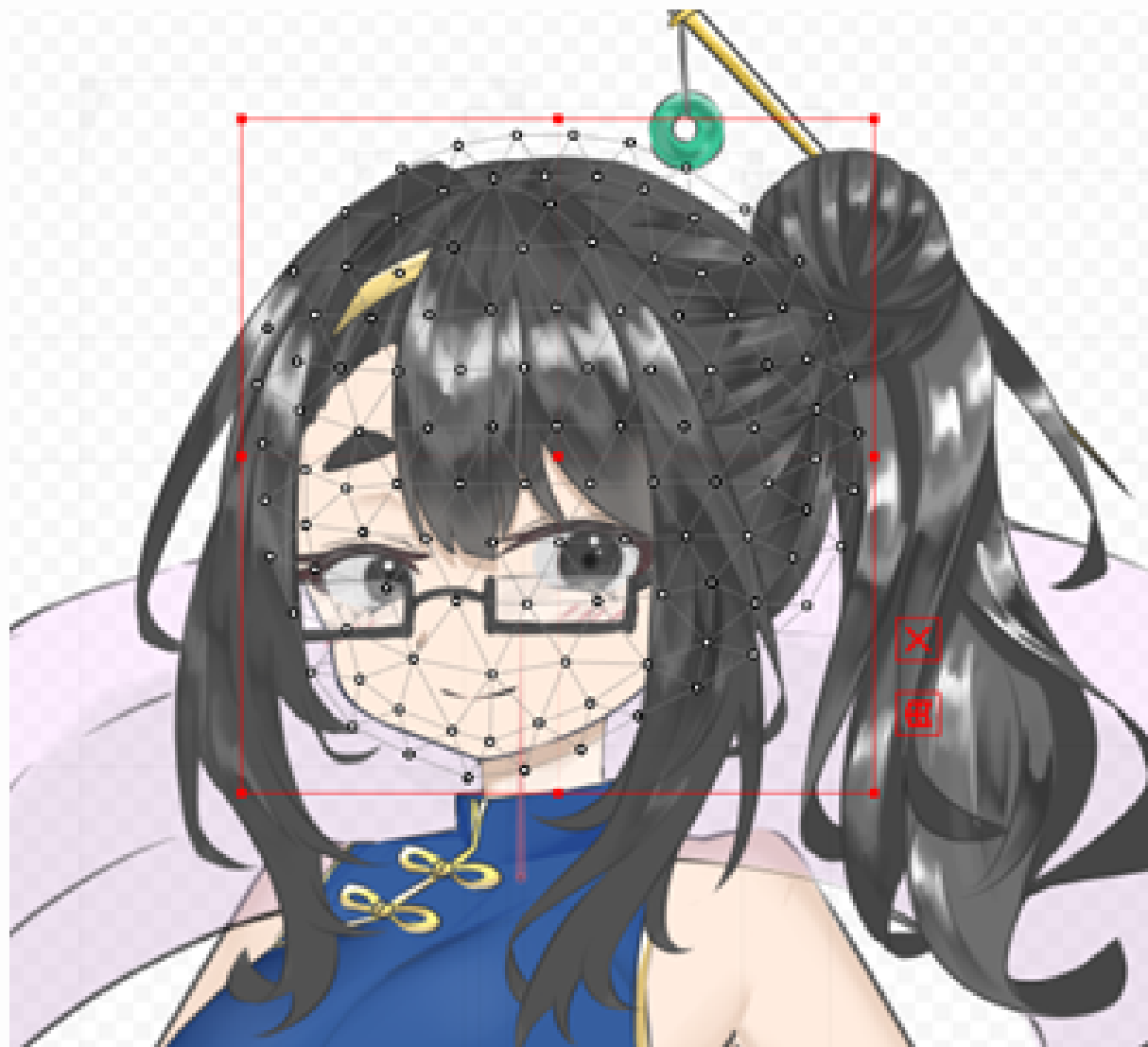
身體、頭髮、眼睛、嘴巴等部位分開繪製



Live2D 建模

將 PSD 檔案匯入 Live2D Cubism Editor

調整各部件的錨點與變形組 設定動作與測試



角度 X:	<input type="range"/>	0.0
角度 Y:	<input type="range"/>	0.0
角度 Z:	<input type="range"/>	0.0
左眼 ...:	<input type="range"/>	1.0
左眼 ...:	<input type="range"/>	0.0
右眼 ...:	<input type="range"/>	1.0
右眼 ...:	<input type="range"/>	0.0
右眼 ...:	<input type="range"/>	0.0
眼珠 X:	<input type="range"/>	0.0
眼珠 Y:	<input type="range"/>	0.0
左眉 ...:	<input type="range"/>	0.0



測試模型

聲音模型建構

- 錄製指定人聲樣本（至少30分鐘）

- 降躁、切片與標記文字對應

- 使用GPT-SoVits進行預訓練與正式訓練

聲音模型建構

錄製指定人聲樣本，音檔前處理(括去除噪音、裁切無效片段



降噪



使用AI先進行辨識再手動修正辨識錯誤的底方



聲音模型建構

選擇預訓練模型，開始預訓練

開啟SoVITS訓練

SoVITS訓練配置輸入

SoVITS訓練配置

100s-GPT訓練，用于分段的模型文件输出在GPT_weights下。

每批显卡的
batch_size

4

總訓練輪數
total_epochs

10

☐ 是否开启edge训练选项
(实验性)

☒ 是否保存每轮的logit
值用以数据清洗空间

☒ 是否使用大模型初始化
您最新的小模型并设置
weights输入

保存频率
save_every_epoch

1

GPU卡号以分割，每个卡号
一个逗号

0

開啟GPT訓練

GPT訓練配置輸入

GPT訓練配置

預訓練完成後進行正式訓練

1A-训练基础设施工具1A-资源列表1C-训练

选择训练所需放在SoVITS_weights和GPT_weights下的模型，默认的一个是预训练，您也可以选择TTS用。

GPT模型列表

GPT_SoVITS/pretrained_models/v1bert25hz

SoVITS模型列表

GPT_SoVITS/pretrained_models/sovits48hz.pt

GPU卡号,只能填1个整数

0

刷新模型列表

☐ 是否开启TTS模型训练

TTS模型训练的输入信息

聲音模型建構

訓練完成後，進入推理，再上傳一參考音檔，就可以輸入想生成的語音

模型切換

GPT模型列表

SoVITS模型列表

參考音檔列表

刷新模型列表

參考音檔中...

請上傳并填寫參考音檔

請填寫需要合成的目標文本和語種模式

請上傳5-10秒的參考音檔，最好會發聲！

拖放音訊至此
或
點擊上傳

需要合成的文本

需要合成的語種

中文

開啟無參考文本模式，不過參考文本亦視乎語種。

使用無參考文本模式時建议使用自训练的GPT，若不用参考音檔的话(不曉得寫啥)可以开，开后就无视填写的参考文本。

聲音模型建構

錄製指定人聲樣本，音檔前處理(括去除噪音、裁切無效片段)



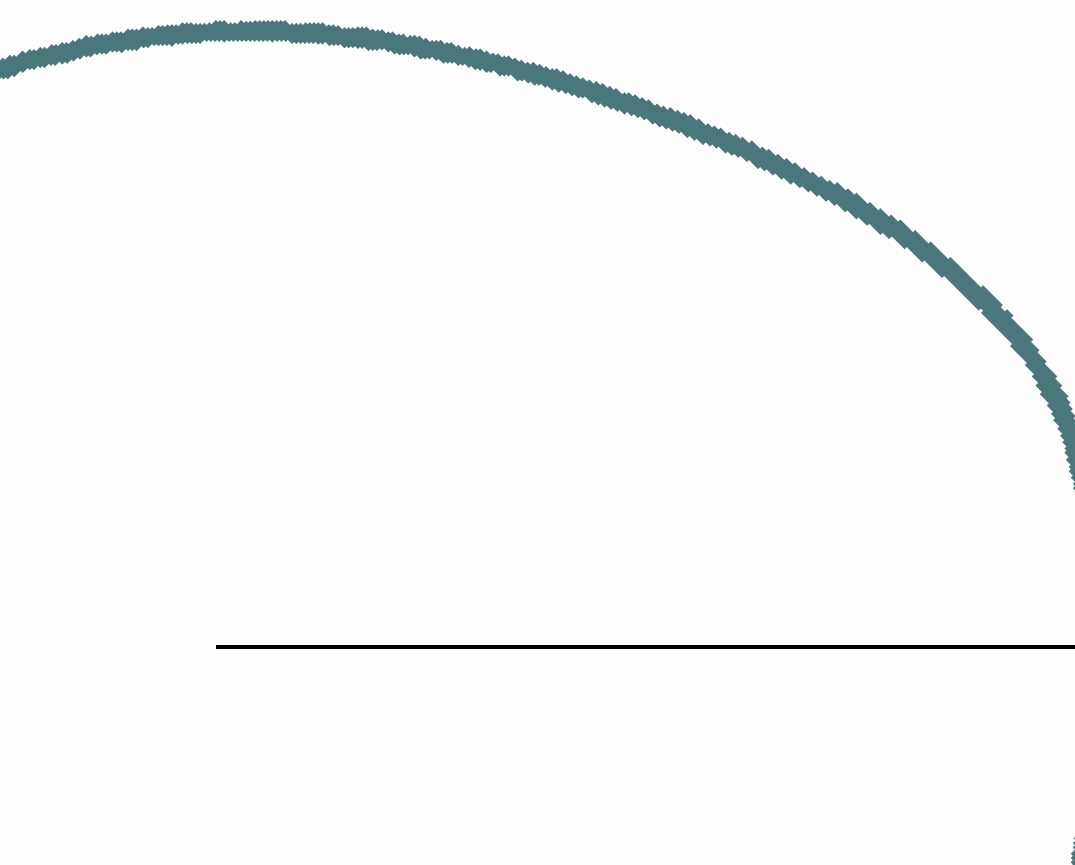
降噪



使用AI先進行辨識再手動修正辨識錯誤的底方



整合與輸出

- 輸入文字 → AI聲音 → 音訊觸發嘴型
 - VTube Studio完成同步與動畫輸出
 - 輕鬆製作文化教育影片
-
- 
- A decorative teal-colored curved line starts from the left edge of the slide and curves downwards towards the bottom center.

整合與輸出

参考音频审阅

*请上传并填写参考信息

🔊 请上传3~10秒内参考音频，超过会报错！

0:05

1x

⏮️ ⏪ ⏩ ⏭

🔄 ✂️

📶 🎤

参考音频的文本

他走了幾步 回頭看見我 說進去吧!

参考音频的语种

中文

☐ 开启无参考文本模式 · 不填参考文本亦相当于开启 ·

使用无参考文本模式时建议使用微调的GPT，听不清参考音频说的啥(不晓得写啥)可以开，开启后无视填写的参考文本。

*请填写需要合成的目标文本和语种模式

需要合成的文本

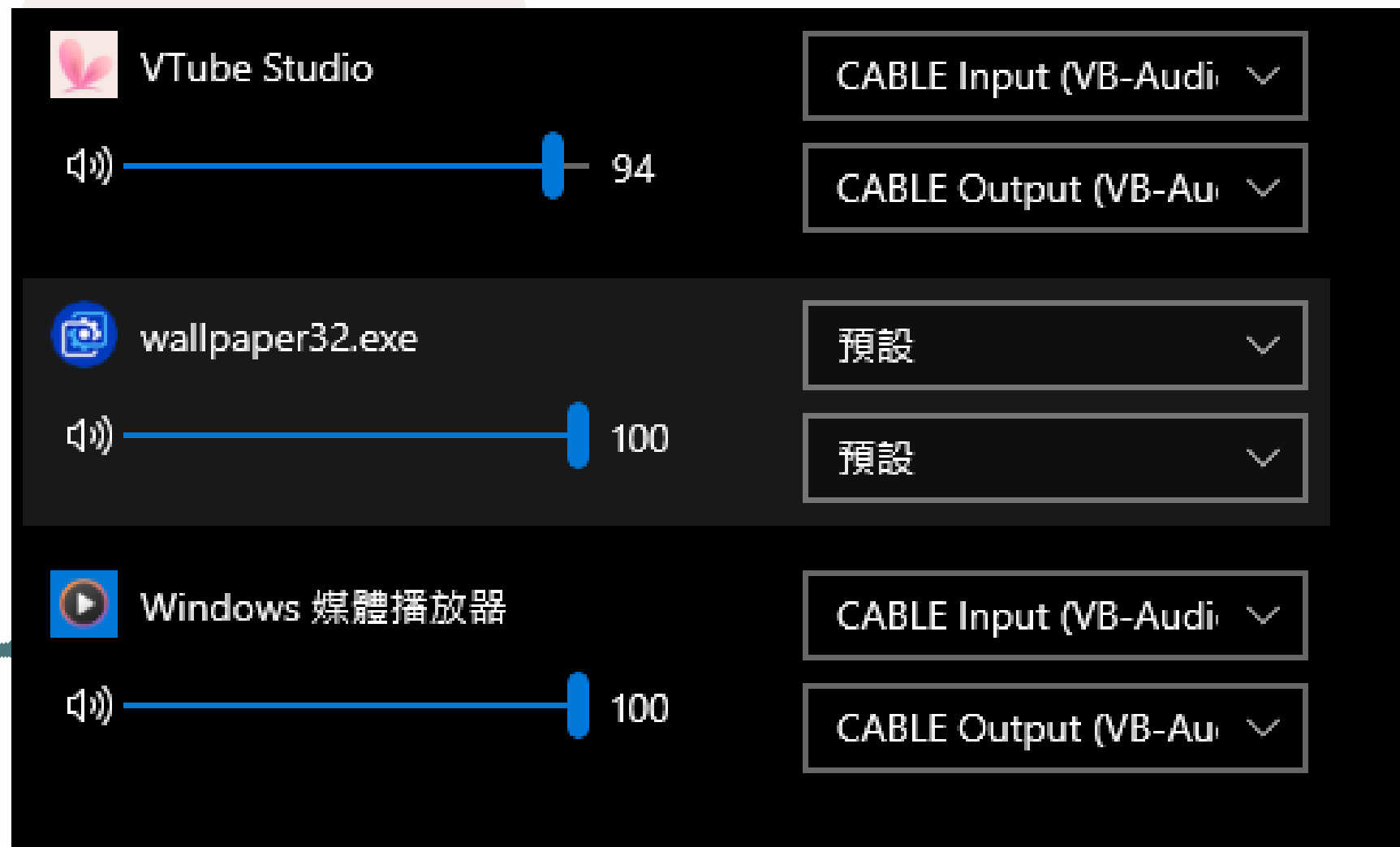
敦煌は東西文化が交差する重要な通過点です
ここには、二つの世界文化のひとつが色濃く映し出されています
それは、敦煌石窟群が世界文化遺産に登録されたことを意味します
そして、ここは東西文化の交差点でもあります
この地には、二千年以上にわたる古代文明が息づいています
では、なぜ「敦煌学」は重要なのでしょうか
国学の大家、季羨林先生の言葉を借りれば
世界には長い歴史を持ち、
広大な地域をカバーし、独自の体系を持ち、
そして深い影響力を持つ文化体系は、
中国、インド、ギリシャ、
そしてイスラム、この四つしかない
第五の文化は存在しない
そしてこの四つの文化体系が一堂に会する場所は、
中国の敦煌しかありません
そこは、古代文明の文化体系が融合する唯一の地なのです

需要合成的语种

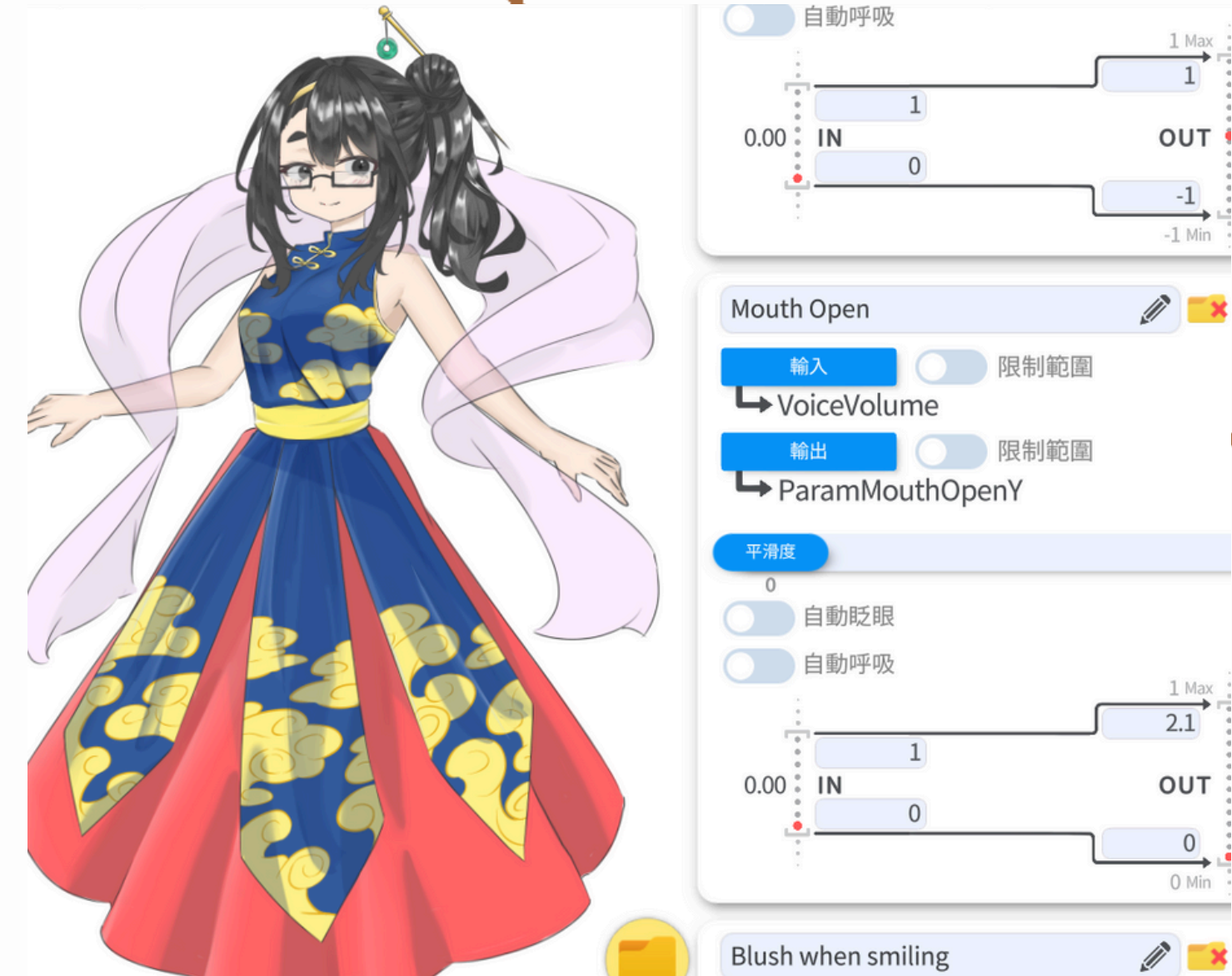
日文

合成文本

整合與輸出

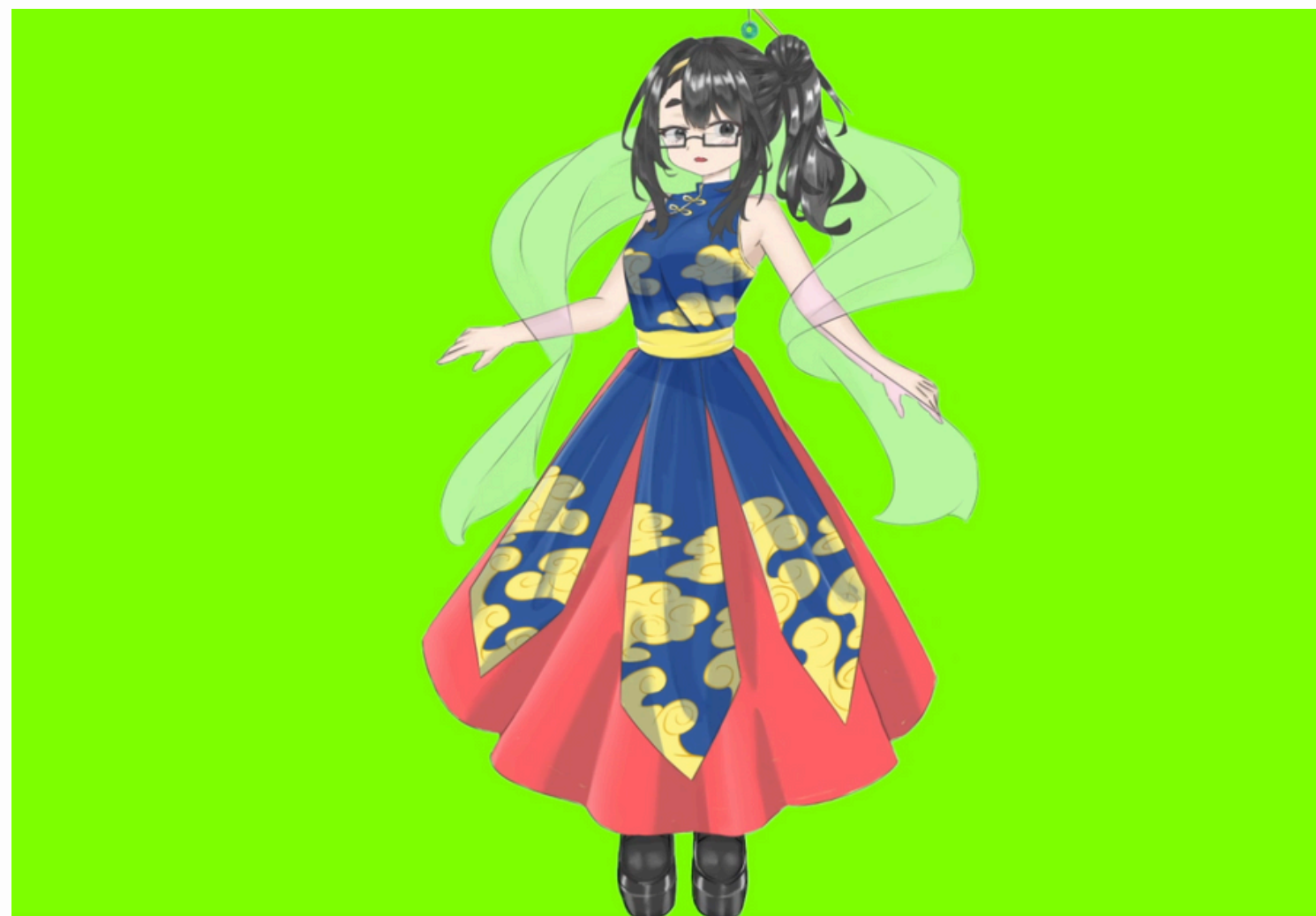


設定虛擬音源



透過虛擬音源來偵測

整合與輸出



錄製音訊觸發嘴型

整合與輸出



製作文化教育影片

實作案例展示

- 再剪輯敦煌介紹
 - 使用AI虛擬主播進行完整講解
 - 表情與嘴型配合語音自動變化
-

成果評估與反思

- 聲音自然度：3.0
 - 嘴型同步度：2.5
 - 整體流暢度：4.0
 - 使用者滿意度：3.8
 - 評估指出：尚有進步空間
-

研究貢獻與限制

貢獻

- 模組化建構AI虛擬講解系統
- 實踐SDGs理念的教育工具

限制

- 嘴型與語音同步尚不精準
- 聲音情緒尚無法自然表現

未來發展建議

- 引入 viseme 分析強化嘴型控制

- 增加即時互動與語調情感變化

- 發展跨語系版本與ESG導覽應用

結語與Q&A

- 我們對於永續文化的理解及行銷相較於我們的實作，還是相對薄弱，但AI虛擬主播是一項具教育與文化潛力的工具可以讓每個人輕鬆的宣傳自己想宣傳的東西。
-



Thanks!
