

# 第一章

## 導論

### 1.1 研究動機

虛擬實境(Virtual Reality)需要大量的計算資源，因此早期研究多只能在大型主機上進行。近年來，電腦硬體技術突飛猛進，使得這些技術被應用在個人及商用電腦上的可能性大為增加。隨著 3D 線上遊戲的盛行，由虛擬實境與分散式技術結合而成的多人虛擬環境系統(如圖 1.1)之相關應用也逐漸在我們日常生活中出現。



圖 1.1：典型的多人虛擬環境系統 – DIVE [15]，本圖取自

<http://www.sics.se/dive/demos/images/research.3.jpg>

近幾年語音技術漸漸進入成熟階段，因此語音技術是另一項近年來受到重視的新型態使用者界面。說話是人類最自然的溝通方式，若能使用語音與電腦互動，將能大幅降低資訊產品的使用門檻。傳統的虛擬實境應用程式中只能以滑鼠或鍵盤操作虛擬人物，在場景中移動或與其中的物體互動。將語音界面加入虛擬實境的應用程式中，將可以大幅地提昇其模擬的真實性。Jacob Nielsen 博士的研究調查指出，大多數 IT 領域的專家認為語音界面將為 21 世紀的人機界面帶來重大改變[36]。因此若能將以上二項技術加以結合，將使我們在虛擬環境系統中與其它人物互動時更為自然。本研究的主要目的在多人虛擬環境中實現互動式語音界面。

在這個研究領域中，我們發現有以下問題值得注意：

### **1.1.1 語音系統如何與人物動畫同步**

一般語音系統包含二大部份：語音合成器(TTS, Text to Speech)與語音辨識引擎(ASR, Automatic Speech Recognition)。欲將語音系統整合至虛擬環境中，首先會遇到的問題便是如何同步播放動畫與語音。舉例來說，我們希望人物邊揮手邊向使用者說 ” Hello” ，若不使用同步的機制協調動畫與語音，很可能發生揮完手才說 Hello，或還沒揮手就說 Hello 的情形。

### **1.1.2 須引入適當之對話管理(Dialog Management)機制**

在未加入語音界面的虛擬環境系統中，通常只須處理動畫播放，使用者可藉由 VRML[48]等高階腳本語言來描述場景，再由該腳本語言之動畫管理器播放。而語音界面主要著重與電腦的互動；在腳本播放途中，必須根據對話的內容(Dialog Context)停下來等待使用者的回應，再依據回應，動態地找出並執行下一段對話

腳本(Dialog Scripts)。另外，決定是否及如何啟始或結束一段對話，也是屬於對話管理的範圍。因此除整合語音合成器與語音辨識引擎外，尚須引入適當之對話管理(Dialog Management)機制，以解決對話管理之問題。

### 1.1.3 執行緒的控制與效能

語音界面的程式一般被認為開發難度較高。例如在每台電腦上，語音合成與辨識引擎只存在單一實體(Singleton)，通常會由所有該電腦上所執行的程式所共享。若架構未經適當的設計，很可能造成系統無法正常運作或效能低落。因此有一些學者針對如何寫作較好的語音應用程式，設計出可重用的語音應用程式框架(Framework) [43]。

虛擬環境與語音系統在執行時期均相當損耗資源，並且很有可能會獨佔某些 I/O 裝置使二者發生衝突。舉例來說，若將二者的程式碼在同一個 JVM 中執行，會發生動畫停格或語音播放斷斷續續等問題。如何對整個系統做適當的架構設計，透過合宜的執行緒的控制，使二者均能達成一定的效能亦是一個值得探討的問題。

### 1.1.4 多人虛擬環境中的對話及語音傳播處理

上述問題在多人虛擬環境中將更為複雜。以語音動畫同步問題為例，如何將二人間的對話以有效率的方式廣播到所有 Client 端，並且仍可做到有效的同步機制。另外，一段對話的語音內容只應該傳送到一定的範圍，如何決定虛擬環境中那些人才聽得到這些對話，並將語音訊息打包傳送到各端點及如何決定如何參與、啟始或結束一個多人的談話等。都是將語音界面加入多人虛擬環境時需探討與解決的問題。

## 1.2 研究目標

本研究的主要目標如下：

1. 虛擬環境與語音界面整合之研究與實作。
2. XAML-V (eXtensible Animation Markup Language - Voice Extension) 動畫與語音互動描述劇本語言之制定與實現。
3. 如何擴充 VoiceXML 對話管理(Dialog Management)機制，在多人環境下達成互動式語音界面之研究。

## 1.3 本論文的貢獻

我們將本論文的貢獻列舉如下：

1. 提出虛擬環境中語音與動畫的整合方式。
2. 利用 XAML 平台具擴充性的特色，結合 VoiceXML，設計出可在多人虛擬環境中提供語音與互動機制的 XAML-V 腳本語言。
3. 針對多人環境下語音互動對話相關問題提出解決方法。
4. 將 1~3 的結果加以整合並使用物件導向的設計樣式(Design Patterns)與框架(Framework)加以實作，以增加成果之可重用性。

## 1.4 研究限制

本論文基於主要的研究目標，在研究過程中做了部份假設與限制，列舉如下：

1. 本研究的實作部份是基於既有的 IMNet 虛擬環境平台與 XAML 腳本語言建置而成。
2. 研究過程中所設計的 XAML-V 腳本語言以達到功能性目標為主，並沒有特別

為執行上的效能做最佳化的動作。

3. 語音界面與 3D 虛擬人物動畫的同步機制並未包含臉部動畫。

## 1.5 論文章節架構

在第二章的相關研究中，我們介紹本研究相關的基礎知識與過去的研究成果。包括多人虛擬環境系統，腳本動畫語言，語音辨識和語音合成技術，語音與虛擬環境的整合與互動式的對話管理技術。將互動式語音界面整合至多人虛擬環境中，首要考量就是如何設計一個良好的整合架構，所以在第三章中我們討論互動式語音界面與 3D 虛擬環境整合時之重要考量，並列出本研究中所採用的整合方式。在第四章中我們主要針對多人虛擬環境中的角色互動與對話管理進行觀念上的討論，並以第五章中 XAML-V 對話腳本語言來實際解決這些問題，包含動畫的整合以及對話管理與協議機制。在第六章中我們會簡要說明目前的實作成果，並在第七章中說明未來可能的研究方向與進展。