

全球資訊網多模式存取之研究

林俊卿 葉耀明

國立台灣師範大學資訊教育所

juinching@hs.ntnu.edu.tw ymyeh@ice.ntnu.edu.tw

摘要

近年來，人們利用全球資訊網來獲取資訊的需求與日俱增，然而傳統網頁以視覺呈現滑鼠操作的方式明顯不符合部分身心障礙者的使用需求，為此，全球資訊網協會開始推動網頁可及性(Web Accessibility)概念，目的是希望不論一般人或身心障礙者，都能夠非常方便地透過全球資訊網取得完整的資訊服務。另一方面，多模式互動機制(Multimodal interaction)強調由多種輸入輸出模式，包括鍵盤、滑鼠、文字、語音及多種無障礙設備與系統展開互動，在多種輸出輸入模式的互動下，使用者可針對自己的特殊需求，選擇最合適的方式存取網頁資訊。本研究根據 Web Accessibility 及 Multimodal interaction 的概念，提出具備雙模式互動的新網頁架構，其特色有以下三點：(1)滑鼠、鍵盤操作雙模式 (2)視覺、語音呈現雙模式(3)網頁符合行政院研考會無障礙機器檢測 AAA 級別。本研究並實作雙模式互動網頁之編輯器，以協助使用者快速開發具備雙模式互動機制的網站。

關鍵詞： Web Accessibility、Text to Speech、Multimodal Interaction。

1. 前言

近年來，全球資訊網以 HTML 語言的方式整合各種多媒體素材，其優異的多媒體呈現能力，生動活潑的互動方式，一躍成為網路上資訊傳遞的主流媒體。再加上電腦與網際網路的進步與普及，使得目前各國人民在日常生活中利用全球資訊網來獲取資訊的人數與日俱增。

然而，傳統的網頁以視覺呈現滑鼠操作的方式明顯無法符合部分身心障礙者的需求，為此，網頁可及性(Web Accessibility)的概念於是提出，其主要概念是透過一些網頁的設計規範和法規條例，來達成身心障礙者可以無障礙的在網際網路上瀏覽網頁資訊。此外，多模式互動機制(Multimodal interaction)也漸漸受到重視，在此機制下，系統可藉由多種的輸入輸出模式與使用者展開互動，包括鍵盤、滑鼠、文字、語音及多種無障礙設備等等，在多種輸出輸入模式的互動下，使用者可針對自己的特殊需求，選擇最合適的方式存取網頁資訊。

本研究根據多模式互動機制的概念，提出具備雙模式互動的新網頁架構，其特色有以下三點：(1)

滑鼠、鍵盤操作雙模式 (2)視覺、語音呈現雙模式 (3) 網頁符合行政院研考會無障礙機器檢測 AAA 級別。

本文包含下列五個章節，第一章簡介研究動機，第二章介紹相關研究，第三章介紹雙模式互動網頁，第四章介紹雙模式互動網頁產生器，第五章討論本研究的成果及未來的發展。

2. 相關研究

2.1 Web Accessibility & Multimodal interaction

根據國外一些相關研究報告中指出【3】【4】【5】【6】，傳統的網頁瀏覽器因其操作介面無法符合視障者的需求，而常常導致資訊取得不足或是資訊認知錯誤。為此，Web Accessibility 的概念【8】於是提出，目的是希望不論一般人或身心障礙者，都能夠非常方便地透過全球資訊網取得完整的資訊服務。Web Accessibility 的主要概念是透過一些網頁的設計規範和法規條例，來達成身心障礙者可以無障礙的在網際網路上瀏覽全球資訊網的網頁資訊。例如，網頁的文字可以使用點字設備顯示或用語音表達等等。

目前主導國際全球資訊網技術的標準機構「全球資訊網協會」(W3C)所形成的「資訊網可及性推動組織」(Web Accessibility Initiative, 簡稱 WAI)【9】已經在 1999 年五月五日訂定出一份「無障礙網頁內容可及性規範」(Web Content Accessibility Guidelines; 簡稱 WCAG)的國際標準。在國內，研考會於九十一年六月「無障礙網頁開發標準暨標準核發作業」委外案，參考 W3C 協會的 WAI 組織的無障礙網頁內容標準相關規範，並參照各國在制訂無障礙網頁相關政策和推廣策略，與國內近年來在身心障礙者保護政策等相關措施，訂定我國的「無障礙網頁開發規範」。

另外，有鑒於傳統網頁以視覺呈現滑鼠操作方式無法滿足部分身心障礙者的需求，多模式互動機制(Multimodal interaction)【1】【2】於是提出，即可藉由多種輸入模式，包括鍵盤、滑鼠、文字、語音及多種無障礙設備與系統展開互動。在支援多種輸出輸入模式的互動下，使用者可針對自己的特殊需求，選擇最合適的方式存取網頁資訊。

2.2 語音合成技術

語音合成 (Speech Synthesizer) 又名文句轉聲音(Text-To-Speech,TTS)，是指將輸入的文字或文件，轉換成清晰、流暢、自然的語音輸出技術。語音合成技術大致可分成文句分析、韻律產生器、合成單元產生器、語音合成等四個模組【7】。

目前的語音合成技術發展日趨成熟，許多語音合成器的產品也已經大量地商品化。因此，藉由運用已成熟的語音合成技術，可以大大減少研究時間的消耗，而更能專注於研究的本質。

3. 雙模式互動網頁

3.1 網站架構

有別於傳統的網頁架構以 HTML 網頁為基本單元，本研究以目錄為一個基本單元，在目錄下包括有：

- (1) XML 記錄檔案：記錄網頁標題、內容及段落等相關資料。
 - (2) HTML 網頁檔案：具備視覺、語音呈現雙模式，鍵盤、滑鼠操作雙模式的 HTML 網頁。
 - (3) MP3 語音檔案：網頁語音檔案，包括導覽語音、全文語音及段落語音。
 - (4) 網頁單元(選項)：表示子網頁單元。
- 其架構如圖 1 所示。

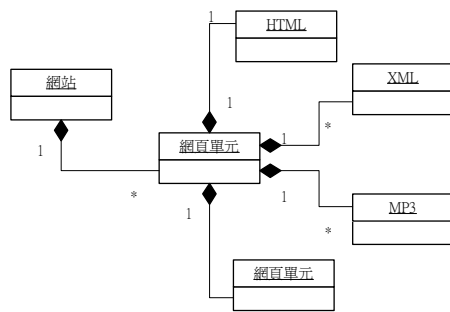


圖 1 網站架構圖

3.2 網頁結構

本研究所提出的網頁有以下三點特色：

- (1) 套用網頁樣版：網頁樣版觀念的引進，讓網頁風格趨於一致，且網站編輯者能輕易的改變網頁的佈景主題，讓網頁呈現更多樣化。
- (2) 網頁區塊劃分：網頁區塊的劃分讓網頁更具結構性，有助於瀏覽者對資訊的取得，本研究劃分區塊如下：
 - 網頁標幟：放置網頁圖示。
 - 目前位置：顯示目前網頁所在網站中之位置。
 - 子網頁：列出目前網頁有哪些子網頁。
 - 網頁內容：顯示網頁內容。

- 網頁語音圖示：每個圖示表示一個語音檔案，包含有網頁導覽語音、網頁全文語音及網頁段落語音。
- 頁尾文字：顯示版權資訊。

網頁結構如圖 2 所示。

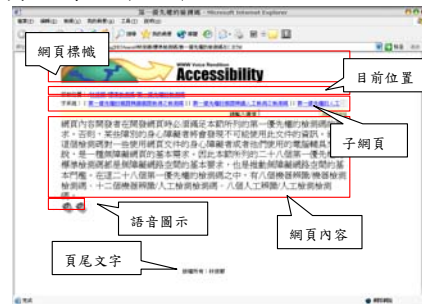


圖 2 網頁結構圖

- (3) 主題語音分段：視覺與聽覺在資訊取得的方式上有著根本的差異，相較於視覺存取方式可以隨機的瀏覽網頁上任何一段資訊，聽覺存取只能以循序方式逐步獲取資訊。為了加速聽覺存取網頁資訊的效率，本研究引進隨機存取的觀念，即除了全文語音閱讀外，編輯者可將全文以主題劃分成若干段落語音，讓使用者可以更有效率的取得需要的資訊。

3.3 呈現雙模式

呈現雙模式包括視覺、語音呈現兩種模式。在語音呈現方面，包含有下列三項：

- (1) 網頁導覽語音：呈現目前所在網頁位置、目前網頁有哪些子網頁及目前網頁有哪些文章段落。
- (2) 網頁全文語音：提供網頁全文語音。
- (3) 網頁段落語音：提供網頁段落語音，段落語音的呈現方式可讓使用者以更具彈性、更有效率的方式進行網頁資訊的讀取。

3.4 操作雙模式

操作雙模式包括鍵盤、滑鼠操作雙模式。在鍵盤操作模式方面，我們將使用者的鍵盤事件分為四種模式，包括有導盲模式、控制模式、連結模式、內容模式等，詳如表 1。

表 1 鍵盤操作模式說明表

模式	說明
導盲模式	雙模式互動網頁區分為標題區塊、子網頁區塊、內容區塊及頁尾文字區塊，導盲模式可以讓使用者用快速鍵於各區塊間切換。

控制模式	以快速鍵執行系統控制功能，包括回首頁、停止播放、繼續播放、暫停播放、開啟(關閉)語音、顯示輔助訊息及登出等。
連結模式	以快速鍵瀏覽頁面上的子網頁。
內容模式	以快速鍵瀏覽網頁全文或段落。

各種模式之操作快速鍵定義如下：

(1) 導盲模式：

Alt+B(Banner)：網頁標題，位於網頁上方。

Alt+N(liNk)：子網頁連結，位於標題列下方。

Alt+C(Content)：網頁主要內容，位於頁面的中間區域。

Alt+R(Right)：版權宣告、其他相關聯絡訊息，位於網頁最下方。

(2) 控制模式：

Alt+G(Guide information)：顯示輔助訊息。

Alt+I(Index page)：回首頁。

Alt+S(Stop)：停止播放。

Alt+P(Play)：繼續播放。

Alt+U(paUse)：暫停播放。

Alt+O(vOice)：開啟/關閉語音。

Alt+L(Logout)：登出。

(3) 連結模式：

0+enter：回上一頁

1+enter：現行頁面下的子網頁 1

2+enter：現行頁面下的子網頁 2

n+enter：現行頁面下的子網頁 n

(4) 內容模式：

Alt+0：閱讀全文

Alt+1：閱讀段落 1

Alt+2：閱讀段落 2

Alt+n：閱讀段落 n

3.5 符合機器檢測三星級

本網頁的設計基於行政院研考會頒發的「無障礙網頁開發規範」以及 W3C 的 Web Content Accessibility Guideline (WCAG 1.0)，並採用研考會之無障礙網路空間服務網之網頁檢測功能，如圖 3 所示，其網址為<http://enable.nat.gov.tw/check.jsp>。本研究將以符合機器檢測三星級為目標。



圖 3 無障礙網路空間服務網

4. 網頁產生器

4.1 系統架構

本系統是以 Apache Tomcat 為網頁伺服器，並採用 JSP(Java Server Pages)伺服器端網頁技術開發而成的網頁應用程式，使用者區分為編輯者及一般使用者，使用案例圖如圖 4 所示。網站編輯及匯出僅限編輯者使用，網站建構共分為三個步驟，分述如下：

- (1) 登入：進行身份確認後方能使用編輯系統。
- (2) 網站編輯：網站管理以檔案總管方式呈現，網站編輯以所見即所得方式呈現，如圖 5 所示，編輯者應可輕易上手。
- (3) 網站匯出：將編輯完成之網站匯出後成為 ZIP 壓縮檔，解壓縮後置於任何網頁伺服器上即可順利執行，不需進行任何額外設定。

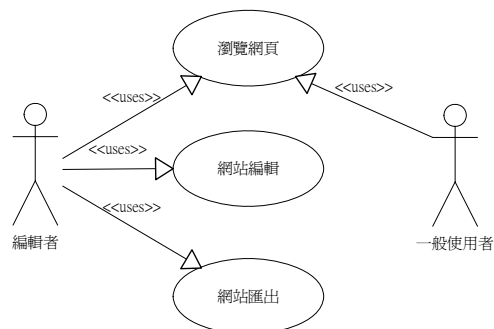


圖 4 使用者案例圖



圖 5 網站編輯介面圖

而系統的組成包括以下四個模組：

- (1) 網站編輯模組：產生網頁單元的 XML 記錄檔案及 HTML 檔案。
- (2) 網頁語音處理模組：分析 XML 記錄檔案並呼叫語音製作模組產生語音檔案，並嵌入 HTML 檔案中。
- (3) 語音製作模組：接收來自網頁語音處理模組的文字檔案並製作語音檔案。
- (4) 快速鍵處理模組：分析 XML 記錄檔案產生快速鍵處理程式碼並嵌入 HTML 檔案中。

模組協同工作圖如圖 6 所示，各模組將於以下各節分別介紹之。

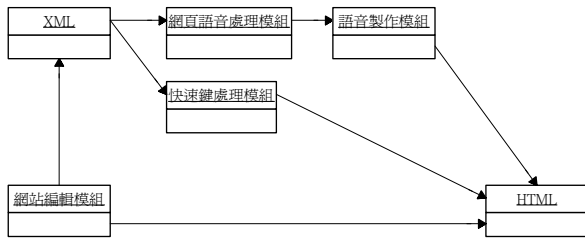


圖 6 模組相關圖

4.2 網站編輯模組

網站編輯模組包括以下三個主要功能：

- (1) 網站編輯：提供檔案總管的操作方式，進行網站結構的維護。
- (2) 網頁編輯：提供所見即所得的操作方式，進行網頁內容的設計。
- (3) 網站匯出：將設計好的網站打包成 ZIP 壓縮檔案。

在網頁編輯方面，編輯者所編輯的資料網頁，將會被分類記錄到多個 XML 檔案，而後套用網頁樣版，產生以視覺呈現方式的 HTML 網頁，其概念圖如圖 7 所示。

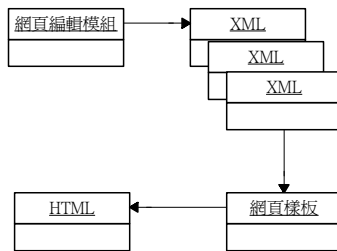


圖 7 網頁編輯模組循序圖

4.3 語音製作模組

語音製作模組可以將文字檔案轉換成 mp3 語音檔案，其動作包含以下兩個步驟：

- (1) 語音轉換：利用 Text to Speech 文字轉語音引擎將文字檔案(txt)轉換成語音檔案(wav)。
- (2) 語音壓縮：利用 LAME encoder 壓縮程式將轉換後的語音檔案(wav)壓縮成 mp3 語音檔案。

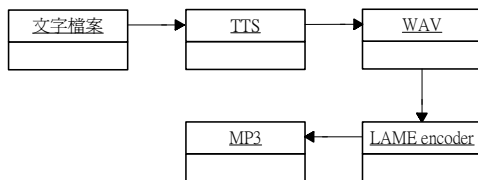


圖 8 語音製作模組循序圖

4.4 網頁語音處理模組

網頁語音處理模組目的在產生網頁瀏覽所需要

的各種語音檔案，並將語音檔案嵌入網頁中，以達到呈現雙模式的目標。其步驟如下：

- (1) 分析記錄檔案：分析網頁編輯模組產生的各個 XML 記錄檔案。
- (2) 產生文字檔案：產生網頁瀏覽各種模式所需要的文字檔案，包括導覽模式、全文模式及段落模式文字檔案。
- (3) 呼叫語音製作模組：呼叫語音製作模組將各種文字檔案轉換成相對應的語音檔案，包括導覽語音、全文語音及段落語音檔案。
- (4) 嵌入語音檔案：將語音檔案嵌入由網頁編輯模組所產生的 HTML 檔案，嵌入語音後的網頁將具備視覺、語音呈現雙模式。

其概念如圖 9 所示。

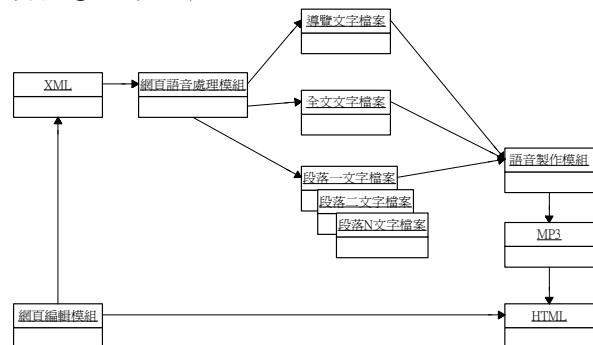


圖 9 網頁語音處理模組循序圖

4.5 快速鍵處理模組

快速鍵處理模組目的在產生網頁所需的快速鍵處理程式碼，並將程式碼片段嵌入網頁中，以達到操作雙模式。其步驟如下：

- (1) 分析記錄檔案：分析網頁編輯模組所產生的各項 XML 記錄檔案。
- (2) 產生各種鍵值處理程式碼：產生各種鍵盤操作模式的處理程式碼，包括導盲模式、控制模式、連結模式及內容模式程式碼。
- (3) 嵌入程式碼：將四種鍵值程式碼嵌入 HTML 檔案中，嵌入程式碼後的網頁將具備滑鼠、鍵盤操作雙模式。

其概念如圖 10 所示。

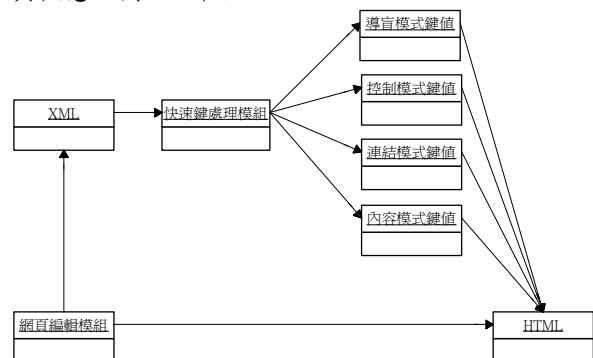


圖 10 快速鍵處理模組循序圖

快速鍵處理模組利用 JavaScript 的鍵盤處理機制捕捉使用者的鍵盤活動，其演算法如下：

```

//四種鍵盤模式之處理
function handleKeys(evt) {
  evt = (evt) ? evt : ((window.event) ? event : null);
  if (evt) {
    //導盲模式鍵值處理
    //if (keyCode=...) ...
    //控制模式鍵值處理
    //if (keyCode=...) ...
    //連結模式鍵值處理
    //if (keyCode=...) ...
    //內容模式鍵值處理
    //if (keyCode=...) ...
  }
}
//設定由 handleKeys 處理鍵盤事件
document.onkeyup = handleKeys;

```

5. 結論

有鑒於身心障礙者在存取全球資訊網時的不便，本研究在符合無障礙網頁規範下，以多模式互動機制為概念，提出具備視覺、聽覺雙模式，滑鼠、鍵盤操作雙模式且通過行政院研考會無障礙機器檢測 AAA 級的雙模式互動網頁，以協助身心障礙者對網路資訊的取得。雖然雙模式互動網頁有如此的優點，但以傳統方式製作雙模式互動網頁有其難度，網站編輯者須具備程式語言基礎及語音檔案製作的背景知識。本研究提出雙模式互動網頁編輯器，以檔案總管的方式管理網站架構，以所見即所得的方式編輯雙模式互動網頁，大大簡化雙模式互動網頁的製作難度。

展望未來，除了雙模式互動網頁的推廣外，本研究更將朝以下方向繼續進行：

- (1) 支援多國語言：目前本研究僅支援中文，依照本系統模組化的設計，其他語言的支援僅需更換 TTS 引擎即可。
- (2) 結合自然語言辨識系統：目前本系統以類似電話語音選單系統方式進行網頁的操作，未來希望可結合自然語言辨識系統，讓網頁的操作更具人性化。
- (3) 結合 VoiceXML：配合政府 M 台灣計劃，擬透過 VoiceXML 結合現有電話系統，使用者只需透過電話系統即可存取網站資訊，讓資訊更隨手可得。

參考文獻

- [1] Consortium, W.W.W.2000a, Multimodal Interaction Use Cases. Available at <http://www.w3.org/TR/mmi-use-cases/>.
- [2] Consortium, W.W.W.2000a. Multimodal Interaction Framework: Available at <http://www.w3.org/TR/2003/NOTE-mmi-framework20030506/>.

- [3] James, F. "AHA: Audio HTML Access", The Six International World Wide Web Conference. Ed, by Michael R. Genesereth and Anna Patterson, Santa Clara, CA, 7-11 April 1997. IW3C2, pp. 129-139.
- [4] James, F. "Presenting HTML Structure in Audio: User Satisfaction with Audio Hypertext", ICAD 96 Proceedings, Xerox PARC, 4-6 Nov. 1996, pp. 97-103
- [5] James, F. "Lessons from Developing Audio HTML Interfaces", ASSETS 98, April 1998, pp. 15-17.
- [6] Rollins, S. and N. Sundaresan, "AVoN calling: AXL for voice-enabled Web navigation", Elsevier Science, Computer Networks, Vol: 33, Issue: 1-6, pp.533-551, June 2000..
- [7] Speech synthesize note, <http://www2.cyut.edu.tw/~s9154610/speech.html>
- [8] Waters, C. "Universal Web Design", New Riders Co., 1997.
- [9] WWW Consortium, "Web Accessibility Initiative". <http://www.w3.org/WAI/>