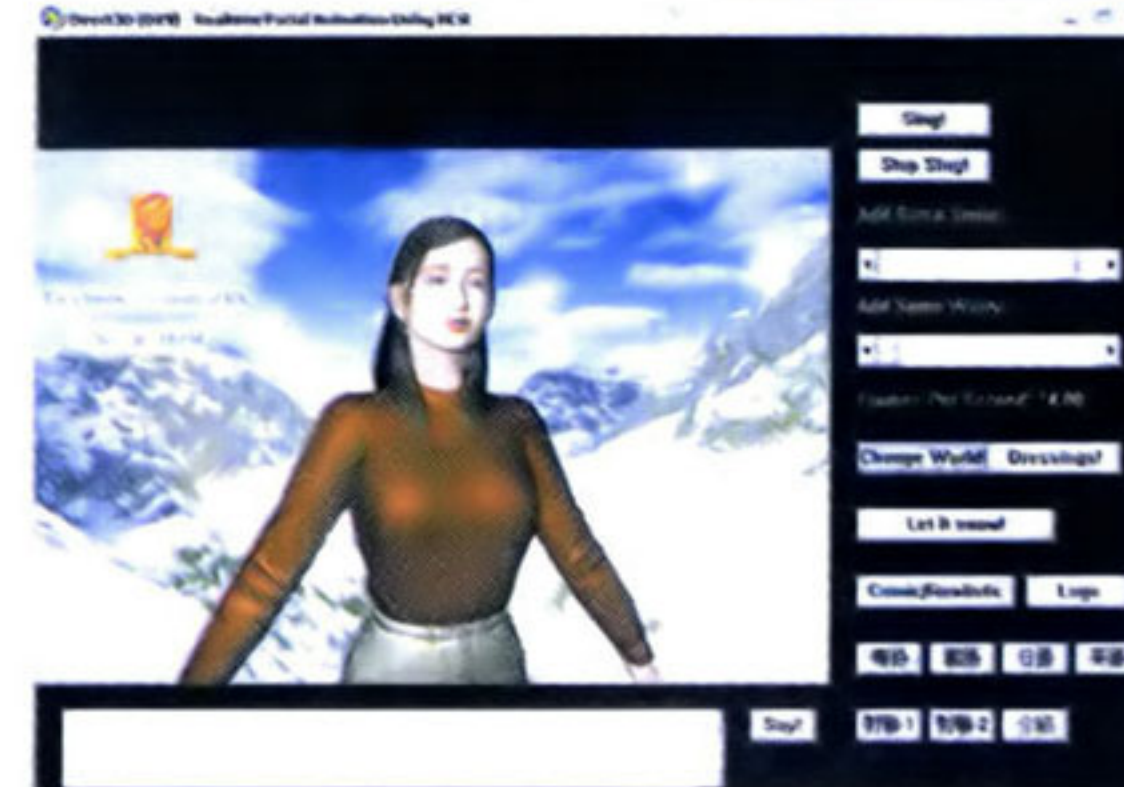
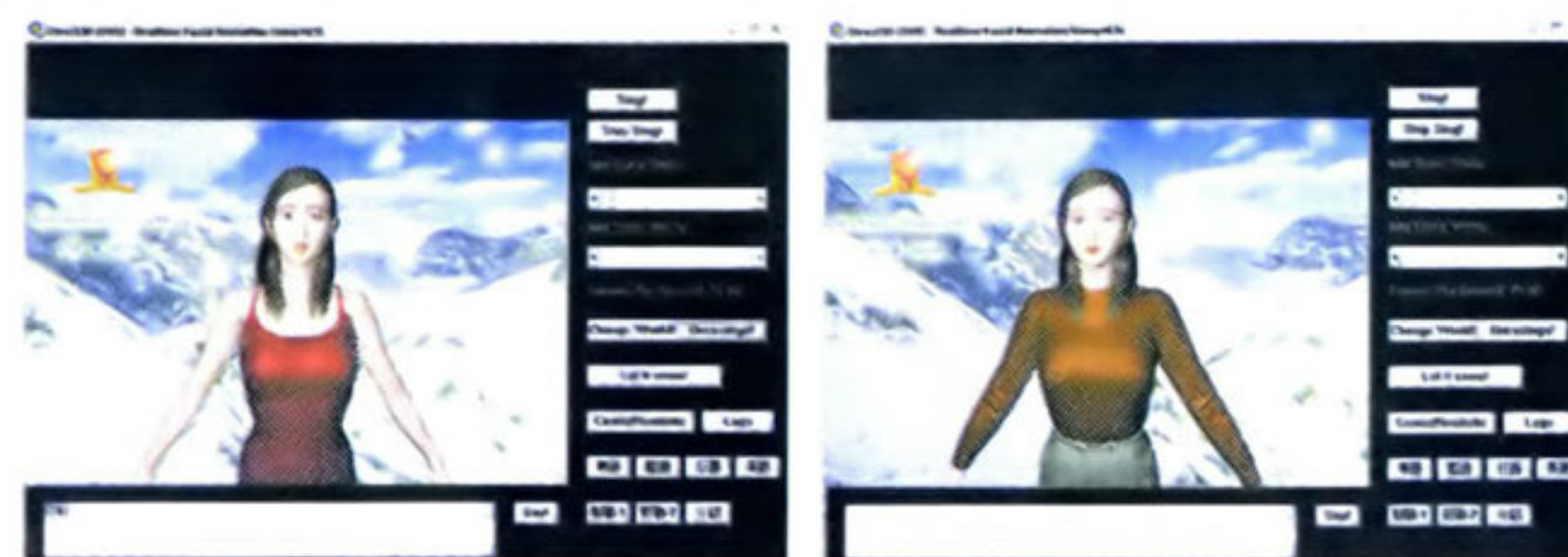


Technology

# 語音生成系統 令虛擬表演者製作成本大縮水

幾年前 DOTCOM 時代，互聯網曾掀起過一陣虛擬角色的熱潮，創造過不少維肖維妙的虛擬新聞主播或者偶像。中大計算機科學和工程學系最近只利用一張千多元圖像卡和已研發多年的自動文本語音生成系統，開發出亞洲首個有臉部表情，言詞舉止同真人相若的女虛擬表演者「琳琳」(Virtual Performer)，兼且訓練其擁有特殊專長，懂得 4 國語言與歌唱表演的能力。



琳琳懂得 4 種語言和唱歌，可手動加入自然面部表情變化，和唇形同步。

適逢中大 40 周年，創新科技成果盡出，虛擬表演者是中大計算機科學和工程學系跨部門大膽創新研究計劃之一。在剛舉行的資訊科技界的國慶酒會，「琳琳」更充當第三位司儀，與在場嘉賓進行實時對話，技驚四座。

## 未來向數碼娛樂發展



▲三位教授與兩位碩士生為虛擬角色創造生命，琳琳有臉部表情言詞舉止與真人相若，甚至懂得眉目傳情。

中大計算機科學與工程學系教授王平安指出，虛擬表演者是第一個能同時具備說普通話、粵語、日語與英語的能力，且實時進行唇形同步，互動臉部表情，有很高的實際應用價值，尤其在數碼娛樂方面的推廣前景極佳。「琳琳未來可以由動畫，演變為有智能的虛擬人，做到很細微的表情，最適合向導遊、數碼音樂和網上遊戲等應用層面發展。不過研究首先要得到政府資助，再考慮延伸到 PDA 或手機流動平台。」他說。現時中大在粵語語音合成技術領先，除虛擬表演者；他透露亦正研究可視人(visual human)

技術，看到人體細微到血管和內部器官。

## 協助視障人士看網頁

參與研究另一位中大系統工程與工程管理學系副教授蒙美玲談到應用前景表示，希望配合中大已開發出的語音合成和文語轉換系統，將網頁內容講出來，幫助視障人士看到網上資訊，掃除數碼隔膜。蒙教授表示，自動文本語音生成系統採用國際音標，除粵語亦可使用多種語言；暫時應用層面有二：

- 1)幫助視障人士聽取網上資訊內容，已答應免費授權給盲人協會，與瀏覽器結合；
- 2)做 e-learning 學習文字語言；應用到政府公共服務網站，閱讀新聞公告內容。



▲背後推動研究的三位教授，左起王平安、黃田津、蒙美玲，考慮應用到數碼導遊、數碼音樂和網上遊戲等層面。

## 節省大量動畫工序

蒙氏指計劃循文字轉換為語音，根據說話的口形輸送，與過往虛擬電腦動畫有很大分別，除

更加真實，亦省掉大量人力物力繪畫和製作動畫。研究中長遠目標是令「琳琳」說話更自然，發展出一套悠揚語音合成系統。現時「琳琳」說話是仍靠人類做文字輸入再讀出，要懂得思考對答比較複雜，需要加入視像、人工智能和語音識別，才能達到互動。

## 突破夾咀形難關

中大計算機科學與工程學系副教授黃田津則認為，以往虛擬人最難突破是自動唇形同步。令到口形和聲音配合，需要製作很多三角形，現時「琳琳」只做到約 2 萬多個；不過已可實時做到表情，頭髮有光澤；而且每秒鐘達到 70 幅圖，任何動作都可即時做到，一般普通 PC 亦可執行。硬件主要靠一張千多元的 ATI Radeon 9800 的圖像處理卡，而控制頭髮光澤唇形同步就靠程式。



▲研究動用的硬件只是這般普通 P4 電腦和一張千多元的 ATI Radeon 9800 圖像處理卡。

以往外國電視節目的虛擬人，因沒有自動唇形同步而製作緩慢，對每天廣播的節目會造成困難。新系統可加快了整個製作過程，亦可以很方便改換虛擬角色，更換不同的語言。

## 世界各地均可使用

參與研究還有兩位中大碩士生，包括計算機科學與工程碩士王劍清和系統工程與工程管理碩士黃嘉豪。他們透露虛擬表演者的設計採用最新並行圖形處理硬件編程技術，最大優點是將動畫師以往需要花上數小時甚至數十小時製的圖像，縮短至實時，大大縮短製作週期並且減低製作成本；系統可手動加入自然面部表情變化和唇形同步。只需經過簡單配置，此系統可被同時用於本地與海外市場，製作娛樂節目甚至虛擬偶像。