

行政院國家科學委員會專題研究計畫 成果報告

網路虛擬環境中自主式數位演員之實現(3/3)

計畫類別：個別型計畫

計畫編號：NSC93-2213-E-004-001-

執行期間：93年08月01日至94年07月31日

執行單位：國立政治大學資訊科學系

計畫主持人：李蔡彥

計畫參與人員：王智賢、黃培智、廖茂詠、陳培峰、廖峻鋒、周旭騏、陶百成、
許書璋、陳哲仁等

報告類型：完整報告

報告附件：出席國際會議研究心得報告及發表論文

處理方式：本計畫可公開查詢

中 華 民 國 94 年 10 月 31 日

行政院國家科學委員會補助專題研究計畫 成果報告
 期中進度報告

網路虛擬環境中自主式數位演員之實現

計畫類別： 個別型計畫 整合型計畫

計畫編號：NSC 93-2213-E-004-001-

執行期間：93年8月1日至94年7月31日

計畫主持人：李蔡彥

計畫參與人員：王智賢、黃培智、廖茂詠、陳培峰、廖峻鋒、周旭騏、陶百成、許書瑋、陳哲仁等

成果報告類型(依經費核定清單規定繳交)： 精簡報告 完整報告

本成果報告包括以下應繳交之附件：

赴國外出差或研習心得報告一份

赴大陸地區出差或研習心得報告一份

出席國際學術會議心得報告及發表之論文各一份

國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年 二年後可公開查詢

執行單位：國立政治大學資訊科學系

中華民國 94 年 10 月 31 日

行政院國家科學委員會
專題研究計畫結案報告
網路虛擬環境中自主式數位演員之實現(3/3)
**Realizing Autonomous Digital Actors
in a Networked Virtual Environment**

計畫編號：93-2213-E-004-001

報告期限：93年8月1日至94年7月31日

主持人：李蔡彥 Email: li@nccu.edu.tw

執行機構及單位名稱：國立政治大學資訊科學系

一、中英文摘要

(一)、中文摘要

電腦動畫是數位內容發展的一種主要項目，也是在技術上最需要突破的一個應用。自主式數位演員的概念是以人工智慧結合電腦動畫的技術，簡化動畫產生程序及所需時間，進而達到即時動態模擬的目的。本研究以三年的時間，建立數位演員及虛擬人群的運動模型，並設計開放式的3D虛擬環境，以利實驗的進行與成果的展示。在第一年的研究裡，我們已達到即時自動產生數位演員在高低不平環境中行走運動的目標。在第二年的研究裡，我們增加數位演員的數量，嘗試以不同的方法模擬虛擬人群的運動。在第三年的研究裡，我們設計了一個開放式的多人虛擬環境系統，並設計具延展性的動畫腳本語言，以利數位演員在即時動畫環境下的實驗。在這個虛擬環境實驗平台下，我們將多人語音對話管理的機制加入動畫腳本語言中，並將語音與動畫整合實現於同一虛擬環境中。此初步研究成果亦已整理成論文，於各國際期刊及研討會中發表。

(二)、英文摘要

In the development of digital content technologies, 3D animation and virtual reality are among those that urgently need technology breakthrough. The concept of Digital Actors (DA) is to apply Artificial Intelligence (AI) techniques in computer animation in order to make animation production more cost effective. We propose to use three years to establish a motion behavior model for a single digital actor as well as for virtual crowds and to design an open virtual environment

system to conduct experiments for simulation of digital actors. In the first year, we have achieved the goal of designing a humanoid motion planner to automatically generate collision-free walk motions for a digital actor in a 3D space. In the second year, we have also made several attempts to generate the motion of a large crowd of digital actors with various methods. In the third year, we have designed an open multi-user virtual environment as well as an extensible animation scripting language in order to conduct experiments for digital actors in an interactive animation environment. In such a platform, we have succeeded in extending the system to incorporate multi-user dialog management into our animation scripting language. The experimental results for the implemented system will be reported in this report and has been published in several international journals and conferences.

二、緣由與目的

電腦動畫與網路虛擬實境的技術是一直都是網路與多媒體領域裡極具挑戰性的一環。在電腦繪圖軟硬體技術日新月異的情況下，3D互動式電腦繪圖的普及日漸，儼然成為下一代人機界面的發展趨勢。目前動畫設計工具的學習曲線較高，人才訓練不易，且設計過程繁複，以致製作成本過高，難以達到經濟效益，因此其應用範圍受到相當的侷限。特別是即時性動畫的相關技術，要能跳脫目前線上遊戲2D動畫或3D罐裝動作的現況，勢必需要開發新的動畫製作技術，以提高動畫設計師的生產力，使其能以少量的高階輸入，即能產生多樣化的即時動畫。

虛擬人(Virtual Human)的研究是近年來電腦動畫研究的熱門課題[2]。在機器人學的領域

```

<AnimItem> := (<AnimItem> | <AnimHigh> | <AnimImport> |
  <AnimNode> | <AnimTransition> | <AnimPlugin>)+
<AnimHigh> := text
<AnimTransition> := (<AnimItem> | <AnimImport> | <Anim-
  Transition>)+
<AnimNode> := (<OrientationInterpolator> | <PositionInter-
  polator>)+
<AnimPlugin> := <xaml-v> | other-plugins

```

圖一、XAML 語言的文法

裡，人形機器人的研發與在娛樂界的應用，在近年來也備受矚目。然而，由於人體運動的複雜度高，而人們對人體運動的擬真度要求也較高，因此要能製作擬真的人體動畫，一直是一件困難的工作。近年來由於運動抓取技術的成熟，大多數人體角色動畫都透過此技術取得擬真的動作。由於是真人運動的重現，這個方式的好處在於其運動的可變度及自然度高；但缺點是無法產生事先未抓取的運動類型。

單一虛擬演員的運動模擬已十分複雜，如要同時模擬大型的人群運動，則計算複雜度更是龐大。然而一般人群運動時，在沒有中央控管的機制下，也都能表現出一些特殊的群體行為，因此如何找出好的局部運動規則，模擬出特定的群體行為，是一個具挑戰性的問題。

目前不論是單一虛擬角色或是虛擬人群運動的模擬，都缺乏較完整的平台可供實驗進行。市售的虛擬環境系統（如 ActiveWorld[12]，或 Blaxxun [14]），多只能提供部分獨家設計的 API，供應用系統作加值之用。而網路上開放原始碼的多人虛擬環境並不多，而且也缺乏較佳的延展功能。因此，自主式數位演員與虛擬人群的成果如要展現推廣，勢必需要仰賴一個開放且具有延展特性的多人虛擬環境系統。

本計畫的目標在以三年的時間，發展一套自主式數位演員技術及適合其運作的網路虛擬環境實驗平台。在第一年的研究裡，我們已經以運動計畫的技術[4]設計出一個在一般3D空間中具有自主性行動能力的數位演員，此演員能接受高階指令，並根據環境的設定，自動產生達到指定目標的運動[5][9]。在第二年的研究裡，我們對虛擬人群的模擬上，我們採用了兩種不同的方式（多人運動計畫演算法[6]及使用虛擬力場的基因演算法[12]），嘗試自動產生大群虛擬人群的模擬運動。在第三年的研究裡，我們將研究的重點放在研發一個開放式的多人網路虛擬環境系統，支援具延伸性的動畫腳本語言，並將系統延伸支援多人語音對話管理，以彰顯本系統的優

```

<AnimItem DEF="WaveWalk" cycle="2000">
  <AnimImport src="Walk">
    <AnimItem DEF="SimpleWave" cycle="1000">
      <AnimNode target="r_shoulder">
        <OrientationInterpolator
          key="..." keyValue="..." />
      </AnimNode>
    </AnimItem>
  </AnimItem>

```

圖二、XAML 動畫腳本的範例

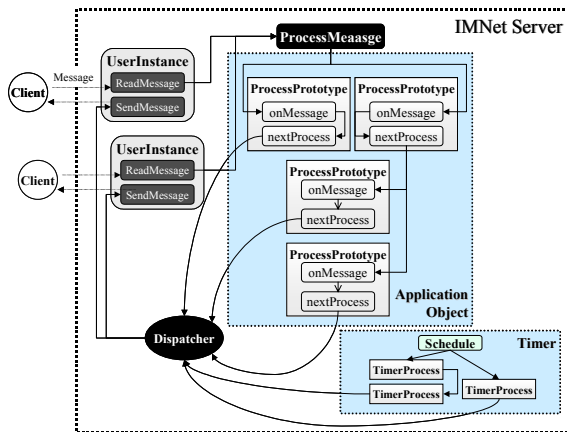
點。

三、 研究方法：

本研究的目標在實現數位演員自主運動的能力。如前一節所述，今年（第三年）的研究目標除了延展過去兩年的研究成果外，主要在發展一個開放式的多人虛擬環境平台。以下謹就動畫腳本語言的設計、虛擬環境系統伺服器端的設計、虛擬環境系統用戶端的設計、及多人語音對話管理的延伸應用作進一步的說明。

（一）、**動畫腳本語言的設計**：根據文獻上的研究[2]，一個理想的動畫腳本語言應具備以下幾點特色。高階（High-level）、完整（Complete）、可延伸（Extensible）、可支援參數化動作（Parameterized Action Support）、同步化（Synchronized）。本計畫所設計的動畫腳本語言稱為可延伸動畫塑模語言（eXtensible Animation Modeling Language, XAML）。此語言以目前最通用的資料註記語言XML為基礎，設計允許高中低階動畫指令能相容並存的結構。而此語言在動畫時間指定的完整性上，符合Allen所分析時間區間的13種可能關係[1]。另外，此語言的基本標記，提供各種屬性參數，以供不同動作的呈現；而使用者自行延伸的標記，也有充分的自由可設計動作的參數化屬性。至於動作的同步化，我們可以透過動畫時間指定與限制傳遞（Constraint Propagation）的方式，達到同步化的目標。整體而言，此動畫語言能符合上述五點動畫腳本語言的理想特質。

此語言的文法請參閱圖一。此語法的基本標記包含動畫高階（如 AnimHigh）、中階（如 AnimTransition）、及低階（如 AnimNode）的標記，並提供使用者自行延伸的插入標記（如 AnimPlugin），以做為此語言和其他以XML為基礎的語言進行整合的介面。圖二為一個以此腳本語言所設計的動畫範例。此語言以遞迴的結構提



圖三、IMServer伺服器的系統架構及其可插入式訊息處理機制

```

<serverConfig>
  <processors>
    <processor class="example.processorA">
      <processor class="example.processorC" />
      <processor class="example.processorD" />
    </processor>
    <processor class="example.processorB" />
    <timerprocessor class="exampleTestTimer"
      delay="5000">
      <processor class="example.processorE" />
    </timerprocessor>
    <timerprocessor class="exampleTestTimer"
      delay="2000" />
  </processors>
</serverConfig>

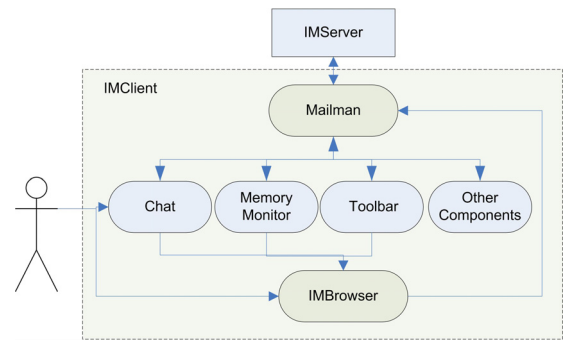
```

圖四、伺服器處理訊息結構之組態範例

供現有動畫腳本的重新組合，如此例中邊走邊揮手的動作，便是由一個外部的下半身行走罐裝動畫及上半身揮手動作的低階指定方式所構成。

(二)、IMNet伺服器端的設計：由於目前網路尚可資運用的虛擬環境系統，不是所費不貲，便是缺乏延展的特性，因此本研究自行設計開發了一個稱為智慧型媒體網路（Intelligent Media Network, 簡稱IMNet）的多人虛擬環境系統。此系統設計的主要目標，除了能滿足一般多人虛擬環境系統的功能外，最重要的在於提供系統的延展性。此系統目前所採用的架構為最常見的主從式架構，所有訊息都經由伺服器作交換。雖然此架構的瓶頸可能在伺服器端，但由於系統架構不是目前本研究的重點，而且此系統設計也可適用於其他架構，因此本系統目前仍以主從式架構進行實做。

IMNet的伺服器端的程式稱為IMServer。伺服器端的設計目標在於系統對訊息處理流程的彈性，以利各種改進措施及實驗的進行。目前



圖五、IMClient用戶端系統架構

IMServer的系統架構如圖三所示。每一個使用者（包含程式控制的虛擬使用者）的登入，都會產生一個UserInstance，接收及傳遞使用者端傳來的訊息。接收到的訊息傳遞至ProcessMessage的模組後，再根據訊息傳遞流程，將訊息以循序或平行的方式傳給相對應的處理模組（Process Prototype）進行處理。而伺服器也可以透過計時器的事件觸發，週期性的自動送出訊息至用戶端。

訊息處理流程是透過一個伺服器的組態檔在執行時期決定的。以圖四為例，一個使用者端來的訊息會平行傳給ProcessorA及ProcessorB，而ProcessorA在處理完後，會平行傳給ProcessorC及ProcessorD。而此組態包含兩個計時器的事件觸發處理器，其中一個又將所觸發的事件傳給另一個處理器ProcessorE作進一步的處理。由於處理器的組合方式及各處理器的程式碼（如example.processorA）都是在執行時動態決定的，因此如要進行不同的實驗，只需對此系統組態檔進行更改後重新啟動伺服器，即能產生預期的功效。

(三)、IMNet用戶端的設計：IMNet的伺服器端的程式稱為IMClient。此程式的設計目標在提供一個具客製化彈性的圖形化使用者介面。此用戶端程式的系統架構如圖五所示。此系統目前由Mailman、IMBrowser、及其他人機介面模組所構成。Mailman模組負責網路的連線、訊息的剖析、包裝與傳遞。其他模組以eventListener的機制向Mailman註冊登記後，便會得到伺服器端所傳來的訊息。IMBrowser則是虛擬環境中的主要3D瀏覽器。此模組的最大特色在於支援XAML語言的動畫腳本，並將使用者操作滑鼠的活動，轉換為適當的訊息，交由Mailman傳遞至伺服器端。

```

<imclient>
  <components>
    <component class="Mailman" name="mailman"/>
    <component class="SpChat" name="chat"/>
    <component class="SpActionButton"
name="actionButton"/>
    <component class="SpMemoryMonitor"
name="memoryMonitor"/>
    <component class="SpIMBro" name="browser"/>
  </components>
  <eventDispatcher name="mailman">
    <connect ip="127.0.0.1" port="62266"/>
    <eventlistener name="chat"/> ...
  </eventDispatcher>
  <toolbar>
    <button name="memoryMonitor"/> ...
  </toolbar>
  <toolbar>
    <buttonContainer name="actionButton">
      <xamlButton name="bow" text="Bow"
file="Behavior/Bow.xml"/>
      ...
    </buttonContainer>
  </toolbar>
  <panels>
    <panel name="browser"/>
    <panel name="chat"/>
  </panels>
</imclient>

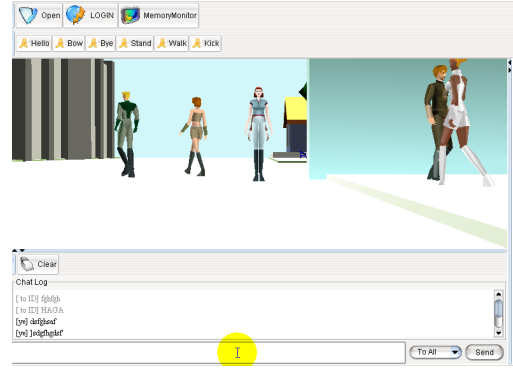
```

圖六、IMClient組態檔的範例

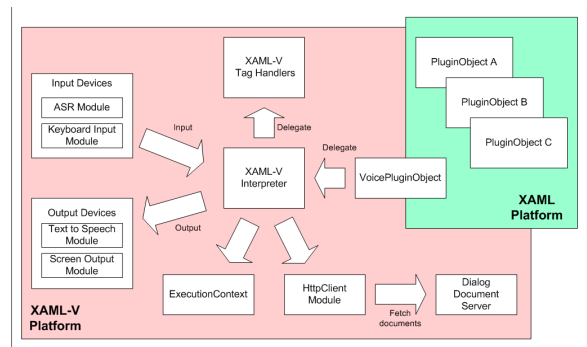
此系統中模組之間的溝通組態及人機介面，是由一個系統組態檔在執行時期所動態決定的。以圖六中的組態檔為例，我們描述此程式由哪些元件（Component）所組成、指定訊息事件的處理機制（Dispatch及Listener）、再指定人機介面中的各種工具列按鈕所對應的處理程式碼。其中工具列又分為兩類，一類需在程式啟始化的過程中即決定，另一類則允許使用者在載入不同的動作檔後再動態產生其動作按鈕。圖七為此組態檔所產生的用戶端圖形介面。

（四）、多人語音對話管理：上述多人虛擬環境系統除了所採用的架構具有開放的特性外，由於其所採用的動畫腳本語言為具有延展性的XML語言，因此可以將現有其他應用整合入動畫系統中。以下我們以多人虛擬環境系統中語音對話管理的機制為例，說明其整合的作法。在XAML中，對於語言中未明訂的應用均已AnimPlugin的標記與XAML系統進行整合。而我們的目標是將語音對話的機制整合入IMNet系統中。語音對話的研究，在過去多是應用在以電話應用為主的語音自動辨識與合成。其中，VoiceXML是近年來以XML語言為基礎的主要標準[15]。然而，VoiceXML設計時的主要應用多是以兩方對話為假設，因此不能直接使用在多人虛擬環境裡。

在此研究中，我們修改並採用部分VoiceXML語言，延伸發展成XAML-V（XAML



圖七、以圖六組態檔所產生的用戶端圖形介面範例



圖八、XAML-V 平台的系統架構圖

Voice Extension) 的對話管理語言。此系統根據對話的過程及對話者的角色（與談人或聽眾），即時產生適當的XAML-V腳本，分散至虛擬環境中的每個人。此系統設計架構如圖八所示。其中，我們在設計中允許XAML-V的腳本回頭呼叫XAML的動畫功能，以達到動畫與語音同步的效果。圖九即是在XAML中插入XAML-V腳本的範例。其中form, block, field等標記都是VoiceXML原有的標記，而Animation標記則是讓此plugin能加入動畫腳本的方式。進一步的例子請參考下一節中的動畫輸出成果。

四、研究結果：

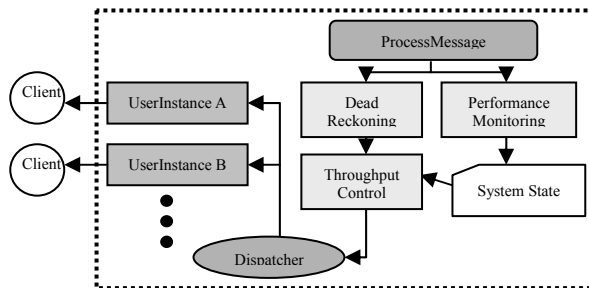
我們已經實做完成上述研究目標所需之軟體系統，並取得初步的成果。本研究在今年的主要成果是一個具有開放性特質的多人虛擬環境系統：IMNet。此系統的3D瀏覽器IMBrowser，是建立在Java3D程式庫的基礎之上，設計支援動畫腳本語言XAML。伺服器端及用戶端的程式都是由Java語言實做，並以XML為檔案資料描述及傳輸的標記語言。

```

1 <!-- initial script -->
2 <?xml version="1.0"?>
3 <AnimItem DEF="Service" playMode="seq">
4   <AnimHigh>Walk to user.location</AnimHigh>
5   <AnimPlugin>
6     <xaml-v version="1.0">
7       <form id="helloForm">
8         <block>Good morning, Sir.</block>
9         <field name="helpType">
10          <prompt>May I Help You? You can say : "I M Lab",
"Computer Center" or "no thanks".</prompt>
11          <animation><!--Field Level Anim-->
12            <AnimImport src="listen" />
13          </animation>
14        </field>
15      </block>
16      <submit next="helpFormResponse.jsp" namelist="helpType"/>
17    </form>
18  </xaml-v>
19 </AnimPlugin>
20 </AnimItem>

```

圖九、將多人對話管理機制 XAML-V 加入 XAML 的範例



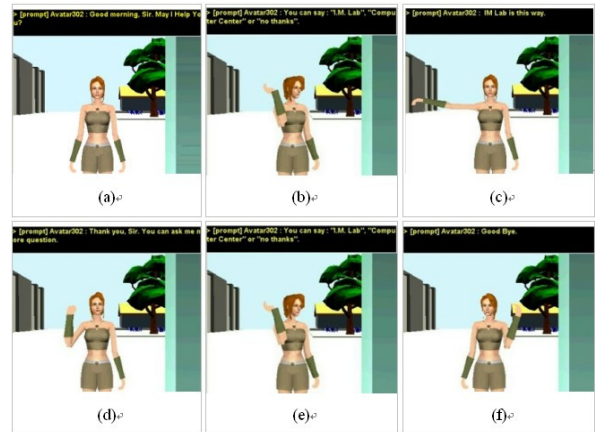
圖十、以一個可調整控制輸出之實驗為範例的伺服器組態

我們已根據系統發展的需要，設計了一些實驗以驗證此系統的延展性。例如，由於主從式架構的問題之一，便是伺服器容易在使用者人數較多的情況下成為系統的瓶頸所在，因此我們利用 IMServer 可延伸的特性，設計了一個資料過濾與監視的機制，以提升系統的效能。此實驗的設計架構如圖十所示。我們以 Dead Reckoning 的方式篩選用戶端的移動訊息 (Dead Reckoning 模組)，僅將變化超過一定門檻的移動送出，並根據用戶端的接收或顯示能力及伺服器的忙碌狀態 (Performance Monitoring 模組)，調整訊息的流量 (Throughput Control 模組)。任何新的訊息過濾機制，只需在組態檔中進行適當的設定，便能重組訊息過濾的流程。

另外，在多人虛擬環境中加入對話管理的機制，我們也進行了一些實驗。語音跟動畫整合的成果如圖十一所示。當系統代理人在等待使用者



圖十一：使用者(右)與代理人(左)進行交談



圖十二、在多人虛擬環境系統中進行對話之範例



圖十三：第三人觀察的效果

的回應時，代理人不但會提醒使用者輸入，還會作出一個聆聽的輔助動作。在使用者跟系統的代理人進行交談時，代理人同時使用動作和語音當作表達方式，使系統代理人更像真實世界的表現。系統會依照使用者的回應資訊的不同，產生不同的腳本。例如在圖九的對話腳本中，使用者回答「IM Lab」時，系統會播放一段動畫(圖十二(a)-(c))，告訴他IM Lab在那裏；使用者回答「No Thanks」時，系統會向他播一段揮手的動畫(圖十二(d)-(f))，並說「Good Bye」。在使用者與系統進行對話的同時，虛擬環境中的其它虛擬人物也能從旁觀者的角度來觀看剛才這段對話，如圖十三所示。其中畫面右方的男性虛擬人物代表使用者，正在與系統扮演的虛擬人物對話。

綜言之，本研究計畫已根據原提計畫目標，完成了以下工作：

1. 建立數位演員下半身的側走、跑步及跳躍的運動模型。
2. 設計適合數位演員多種運動及複雜地形的全域路徑計畫模組。
3. 設計以動態編組為基礎的人群運動計畫器，並以實驗方式證實其有效性。
4. 設計人群模擬之行為模型及環境模型，並以基因演算法求取最佳人群行為模擬之表現。
5. 設計具延展性的動畫腳本語言，並以語音管理系統為例進行延伸
6. 設計開放式且具延展性的多人虛擬環境系統，並在此平台上進行實驗。

本計畫所獲致的成果，已整理發表於知名國際期刊 [8][11] 及學術研討會 [5][6][7][9][10][12] 中；另外，計有八位同學根據本研究的成果，完成碩士論文，並已順利畢業投入職場。我們相信在此計畫中，我們已經建立了一個數位演員的實驗平台，未來我們將就數位演員的自主推理能力作進一步的研究，以達在多人虛擬環境中即時運動計畫的目標。

五、參考文獻

- [1] J.F. Allen, "Maintaining Knowledge about Temporal Intervals," *Communications of the ACM*, 26(11): 832-843, November 1983.
- [2] Y. Arafa, and E. H. Mamdani, "Scripting embodied agents behaviour with CML: character markup language," *Proceedings of Intelligent User Interfaces 2003*, pp. 313-316, 2003.
- [3] T.K. Capin, I.S. Pandzic, N. Magnenat-Thalmann, D. Thalmann, "Integration of Avatars and Autonomous Virtual Humans in: Networked Virtual Environments", *Proceedings of. ISCI 98, IOS Press*, pp. 326 - 333, Amsterdam, Netherlands, 1998.
- [4] J.C. Latombe, *Robot Motion Planning*, Kluwer, Boston, MA, 1991.
- [5] T.Y. Li, P.F. Chen, P.Z. Huang, "Motion Planning for Humanoid Walking in a Layered Environment," in *Proceedings of the 2003 International Conference on Robotics and Automation*, 2003.
- [6] T.Y. Li, H.C. Chou, "Motion Planning for a Crowd of Robots," in *Proceedings of the 2003 International Conference on Robotics and Automation*, 2003.
- [7] T.Y. Li, M.Y. Liao, J.F. Liao, "An Extensible Scripting Language for Interactive Animation in a Speech-Enabled Virtual Environment," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME2004)*, Taipei, 2004.
- [8] T.Y. Li, M.Y. Liao, and P.C. Tao, "IMNET: An Experimental Testbed for Extensible Multi-user Virtual Environment Systems," ICCSA 2005, LNCS 3480, O. Gervasi et al. (Eds.), Springer-Verlag Berlin Heidelberg, pp. 957-966, 2005.
- [9] T.Y. Li, and P.Z. Huang, "Planning Humanoid Motions with Striding Ability in a Virtual Environment," in *Proceedings of the 2004 International Conference on Robotics and Automation (ICRA2004)*, 2004.
- [10] J.F. Liao, and T.Y. Li, "Incorporating Voice Dialogs in a Multi-user Virtual Environment," in *Proceedings of 2004 International Computer Symposium*, Taiwan, 2004.
- [11] J.F. Liao, and T.Y. Li, "Realizing Voice Dialog Management in a Collaborative Virtual Environment," *International Journal of Computer Applications in Technology (IJCAT)*, 2006.
- [12] T.S. Wang, and T.Y. Li, "Simulating Crowd Motion Behavior with Evolutional Method," in *Proceedings of the Sixth Conference on Artificial Intelligence and Applications*, Taiwan, 2005.
- [13] ActiveWorlds, <http://www.activeworlds.com>
- [14] Blaxxun, <http://www.blaxxun.com/>
- [15] VoiceXML, <http://www.w3.org/TR/voicexml20/>