

Utilizarea tehnologiei Motion-Capture in medii interactive in timp real

Rizea Victor Adrian, Dinca Florin Alexandru, Ilie Ciprian, Hramco Vera, Polceanu Mihai, Dorin-Mircea Popovici

Universitatea OVIDIUS Constanța

B-dul. Mamaia 124, 900527, Constanța

victor.rizea@gmail.com; cerva@univ-ovidius.ro; dmpopovici@univ-ovidius.ro

REZUMAT

Contribuția de față aduce în discuție utilizarea tehnologiei de captare a mișcării umane (*motion capture* - MoCap) într-un context virtual interactiv. Perspectivele abordate sunt complementare; prima tratând problema utilizării MoCap în obținerea comportamentelor la nivelul actorilor virtuali iar a doua problema captării, aplicării și redării în timp real a comportamentelor exprimate natural pe un personaj virtual, avatarul utilizatorului.

În vederea exemplificării problematicii discutate am făcut apel la mediul virtual, interactiv, al cetății Tomis. Aici, utilizatorii pot atât să experimenteze activități portuare prin intermediul interacțiunii cu personaje virtuale autonome ce populează mediul virtual, cât și să converseze cu un personaj virtual controlat de un expert în domeniul istoric, ghidul real.

După prezentarea soluției arhitecturale, articolul discută asupra aspectelor de aplicabilitate ale tehnologiei MoCap, demonstrând punctul de vedere pe baza aplicației susmenționate și încheie prin perspective și concluzii.

Cuvinte cheie

motion capture, mediu virtual, interacțiune

Clasificare ACM

H.5.2 User Interfaces (D.2.2, H.1.2, I.3.6)

INTRODUCERE

Captarea și redarea în timp real a mișcărilor umane într-un sistem virtual constituie o tehnologie remarcabilă care a devenit în ultimii ani accesibilă comunității de realități mixte (Mixed-Reality - MR). Indiferent dacă are la bază strict un sistem de procesare de imagini (Organic Motion[12]), utilizează un sistem bazat pe o configurație de camere infraroșu și un set de markeri (OptiTrack[11]), utilizează sisteme inerțiale (Xsens[16]), sau exoscheleți (eg., MetaMotion[10]), astfel un sistem de MoCap permite transpunerea mișcărilor unui actor real în mișcările unui personaj virtual.

Potențialul acestei tehnologii este dovedit în domeniul precum industria filmelor (Avatar[3]), cea a jocurilor pe calculator (FIFA[8]) și chiar în medii deschise mai puțin largului public precum cel al telemedicinii, sau de teleoperare a unor vehicule/dispozitive de precizie cum ar fi demonstrația din proiectul eMove[7].

Tehnologia MoCap a fost aplicată pentru simularea unei mulțimi de oameni [5] dar este o tehnologie ce presupune

un proces de lucru relativ anevoios. Dovadă stă proiectul EPOCH ce reconstituie virtual orașul Pompei recurgând animarea clasică prin adugarea manuală a cadrelor cheie [9].

În cele ce urmează, vom avea în vedere două tipuri de tratare a mișcărilor obținute prin tehnologia MoCap:

- înregistrarea mișcărilor și redarea lor ulterioară în mediul virtual
- detectarea mișcărilor și redarea acestora în timp real

Primul mod presupune înregistrarea unor mișcări generice ale actorului (e.g., mers, alergat, urcat/coborât scări). Acestea vor constitui intrări al unui sistem de control al animațiilor, pe baza cărora, devine posibilă sintetizarea unui comportament complex asociat personajului virtual.

Această soluție este aplicabilă în cazul în care utilizatorul unui mediu virtual dispune de un avatar pe care-l poate controla utilizând dispozitive de interacțiune uzuale (e.g., mouse, tastatură, joystick, wiimote, wand, comandă vocală, etc.). În acest fel, utilizatorul declanșează acțiuni pe care avatarul le duce la bun sfârșit (e.g., "Go Forward!" poate însemna "mers înainte" dar și "urcat scară").

În acest caz, soluția MoCap contribuie la verosimilitatea comportamentală a avatarului.

Cel de al doilea mod presupune transformarea actorului în utilizator i.e. mișcărilor acestuia sunt reprezentate în timp real în mediul virtual, prin intermediul avatarului asociat.

Interacțiunea cu mediul virtual devine astfel naturală deoarece limbajul corpului este utilizat atât în comunicarea dintre utilizator și mediul virtual, cât și în comunicarea dintre mediul virtual și utilizator, de preferință printr-un (*feedback*) multi-modal.

Acestea sunt cele două moduri de generare și transpunere a mișcării umanoizilor virtuali în cadrul proiectului Tomis[16].

PREZENTARE GENERALĂ

Notiunea de ghid se regăsește în [6] unde acesta este un agent autonom în mediu. În aceasta lucrare ne propunem să atribuim acestuia caracteristici dinamice umane.

Elementul în care ghidul își va desfășura activitatea este mediul din lumea reală. Această activitate este transpusă prin intermediul sistemului de MoCap în mediul virtual, ca

rezultat al proiectului Tomis. Acest proiect a avut ca scop - Utilizarea realității virtuale în reconstituirea Multimodala 3D a site-urilor Istorice - mai concret reconstituirea virtuală 3D a Edificiului Roman cu Mozaic. Ca rezultat direct, vizitatorul are posibilitatea unei experiențe interactive, creându-și o imagine completă, dar personală, asupra zonei portuare a anticului Tomis.

Interacțiunea cu mediul virtual s-a desfășurat pe 2 planuri :

- din punct de vedere al utilizatorilor, orice membru din public are posibilitatea de a explora mediul. Contribuția proiectului prezentat constă în introducerea în mediu a unui ghid care poate comunica cu publicul.
- din punct de vedere al ghidului, acesta interacționează cu mediul virtual folosind mișcări ale corpului vizualizate prin intermediul unui avatar existent în mediu.

Orice membru din public poate adresa întrebări ghidului iar răspunsul va fi oferit în timp real. Exprimarea membrilor poate fi naturală deși aceștia vor adresa întrebarea aparent unei mașini. Prin simplul fapt că răspunsul este furnizat de către persoană reală prin intermediul unui calculator, informația oferită este una dinamică. Membrul din public având ocazia să extindă întrebarea în cazul în care informația necesită clarificări ulterioare.

Ghidul se poate manifesta din punct de vedere al transmiterii informației către public folosind comunicarea orală completată de limbajul corpului.

Informația furnizată de ghid publicului vizitator, prin explicații verbale și gestuale supuse regulilor de comunicare naturală, explică sensul contextului dinamic al mediului virtual reconstituit.

Numitorul comun al celor două perspective îl constituie metafora utilizată în transmiterea informației, și anume interacțiunea prin limbajul corpului. Tehnologia MoCap a fost folosită pentru a genera animațiile umanoizilor prezenți în mediul virtual dar și pentru a reprezenta mișcările ghidului.

ARHITECTURA SISTEMULUI

Sistemul realizează legătura dintre două medii distincte situate în lumea reală, mediul ghidului turistic (MG) mediul publicului vizitator (MP), prin intermediul unei componente multimodale (Figura.1).

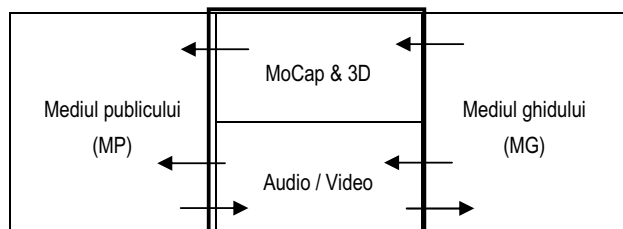


Figura 1. Conectarea multimodală a două medii

Cu alte cuvinte, între MG și MP există o legătură bidirecțională audio, ghidul aude publicul și viceversa. La nivelul legăturii video, legătura este unidirecțională orientată dinspre MP spre MG, ghidul vede publicul.

Feedback-ul vizual la nivelul publicului este obținut prin combinarea tehnologiei MoCap cu mediul virtual 3D.

Arhitectura întregului sistem este schematizată în (Figura.3). Mediul publicului presupune existența a două calculatoare, C1 (responsabil cu legătura audio/video) și C2 (responsabil cu redarea grafică a mediului virtual în care evoluează avatarul ghidului).

La nivelul lui C2 are loc sincronizarea mișcărilor avatarului cu mișcările în timp real ale ghidului, pe baza informațiilor generate de sistemul MoCap (C3).

Redarea mediului virtual se realizează prin componente periferice specifice (proiector, webcam, microfon, sistem audio).

Mediul ghidului este identic cu cel al publicului cu excepția sistemului de MoCap. Pentru a fluidiza traficul de date emise în timp real de sistemul MoCap, am dedicat un calculator (C3) acestui scop.

SOLUȚIA PROPUȘĂ

Arhitectura propusă a fost aplicată în contextul mediului reconstituit al cetății Tomis, prin aplicarea tehnologiei MoCap pentru animarea atât a personajelor virtuale autonome cât și a avatarului asociat ghidului.

Mediul virtual

Întrucât ghidului turistic îi este necesar un context în care să se manifeste, am ales să îl plasăm în zona portuară a reconstituirii cetății Tomis.



Figura 2. Exemplu de activitate portuara în Tomis

Modelarea mediului virtual s-a realizat atât din perspectiva geometrică, utilizând pachetele software Blender3D[4] și 3DSMax[1], cât și din cea comportamentală, utilizând platforma orientată spre agenți ARVi[2].

În ceea ce privește modelarea geometrică a trebuit să găsim o soluție de mijloc între fidelitatea geometrică (ce presupune modele 3d bine detaliate) și fidelitatea texturilor (ce presupune utilizarea unor texturi de rezoluție înaltă). În acest sens am aplicat tehnici de folosirea a hărților de normale și un sistem de control al nivelului de detaliu (LOD), un sistem asemător este regăsit și în proiectul EPOCH.

Prin modelarea comportamentală am obținut autonomia personajelor virtuale ce populează mediul virtual, urmând atingerea unor obiective proprii. În acest sens unul din elementele prezente în aplicație ce definește noțiunea de agent este reprezentarea 3D a tomitanilor. Acești agenți se prezintă cu proprietatea de a fi autonomi, având posibilitatea de a accesa zonele de importanță prezente în mediul virtual apelând la o rețea de drumuri. Cu alte

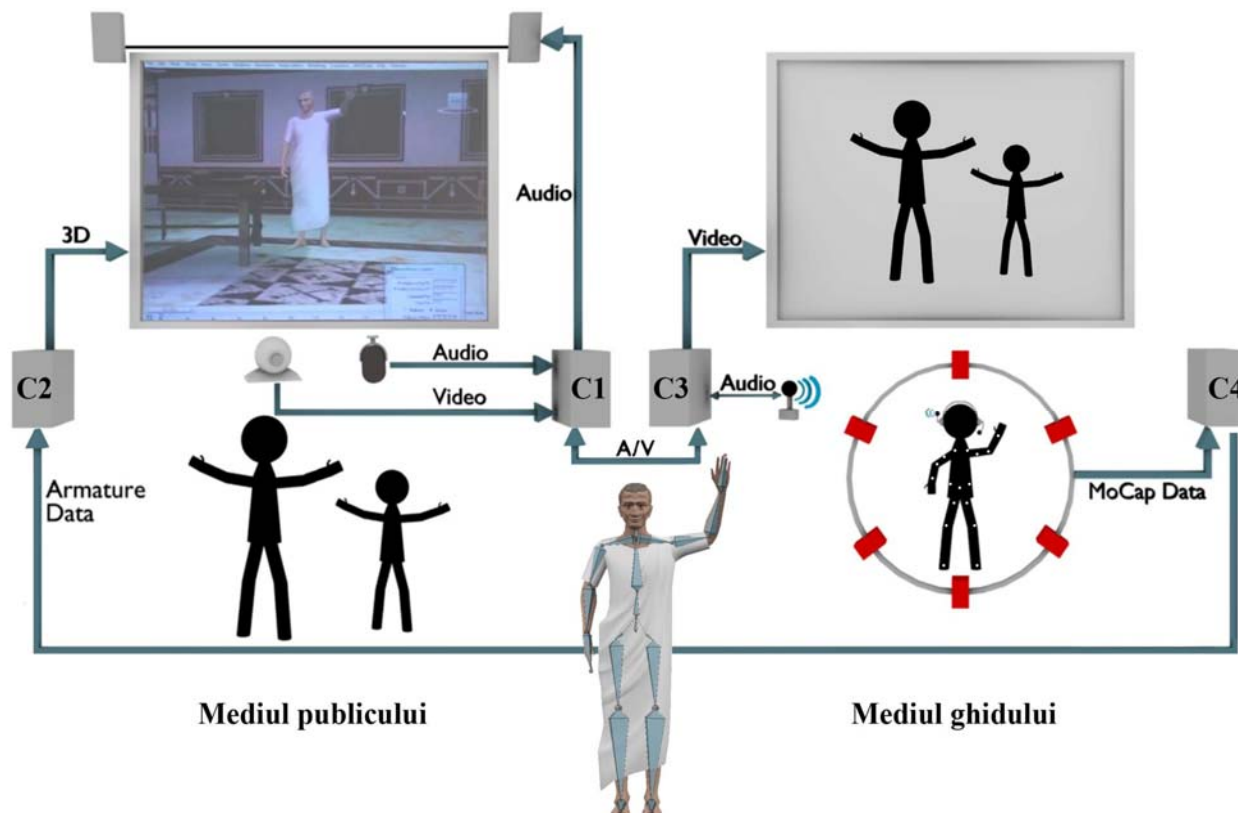


Figura 3. Arhitectura sistemului

cuvinte își pot planifica un traseu, iar în funcție de contextul destinației vor adopta un anumit comportament. Un comportament este definit de o serie de stări ce sunt vizualizate prin animații caracteristice. Mișcările ce definesc animațiile sunt mișcări generice ce pot fi îmbinate pentru a obține un comportament mai complex. Acest proces de modelare comportamentală este detaliat în [13].

Atunci când controlul unui personaj virtual încetează să aparțină strict personajului, putem vorbi de avatarul utilizatorului. Obținerea unui comportament credibil asociat acestuia dar și al agenților autonomi am realizat-o în două etape.

Captarea și înregistrarea mișcărilor corpului pentru animarea agenților

Astfel pentru a defini aceste mișcări generice (eg., mers, urcat scări, aplecare/ridicare, etc.) am folosit MoCap. Datele obținute prin sistemul MoCap sunt apoi analizate și prelucrate în programul de modelare 3D unde sunt utilizate în animarea personajelor. În acest sens, fiecare personaj dispune de un schelet (sau armătură) a cărui funcție este aceea de a acționa asupra punctelor (vertecșilor) ce definesc geometria umanoidului, într-o manieră ponderată (prin *weight paint*).

Așa cum sugerează primul cuvânt din această denumire, fiecare os poate fi privit ca un obiect ce are masă, astfel un os poate influența spațiul din jurul lui, spațiu în care se află vertecși ce definesc forma personajelor. Cel de al doilea cuvânt sugerează procesul prin care se definește câmpul gravitațional al unui os. Din punct de vedere vizual, procesul este concretizat sub forma unei unelte

specifice programului de modelare 3D prin care se desenează pe suprafața obiectului porțiunile ce vor fi deformate odată cu rotirea unui os.

Întrucât la momentul introducerii umanoizilor în mediul virtual posibilitatea folosirii unei armături în cadrul platformei grafice a fost inexistentă, animațiile umanoizilor au fost definite sub formă de fișiere individuale ce conțin cadrele cheie ale animației. Obținerea ulterioară a animației corespunzătoare unei acțiuni este obținută prin interpolarea cadrelor cheie.

Interacțiunea din perspectiva utilizatorului

Un utilizator are posibilitatea de a explora mediul virtual și de a observa evoluția agenților din mediu folosind diverse dispozitive (eg., *mouse*, *joystick*). Acesta își poate asuma un rol omniprezent (prin navigare liberă de orice constrângere) sau poate adopta perspectiva unui personaj prin selecția unui umanoid virtual și a asista sau controla acțiunile acestuia.

Un umanoid din mediu ce este controlat de un utilizator își pierde parțial autonomia deoarece comportamentul lui devine imprevizibil. Utilizatorul este cel care decide declanșarea acțiunilor umanoidului iar acesta din urmă execută mișcările asociate activității combinând animațiile.

Animarea avatarului controlat de ghid în timp real

Avatarul nostru este o reprezentare a faimosului poet roman Ovidius Publius Naso, exilat pe țărmul Tomitan al Pontului Euxin. Reprezentarea acestui personaj în mediul virtual a fost obținută pe baza informațiilor documentare existente referitoare la aspectul vizual al poetului. În ceea

ce privește tehnicile de modelare 3D a umanoidului, am recurs la modelarea poligonală prin subdiviziuni a corpului urmărind planuri ce definesc anatomia corpului uman. Ulterior s-a modelat vestimentația folosind aceeași tehnică, având ca referință tunica și toga cu care Ovidius apare îmbrăcat în reprezentările existente.

Spre deosebire de corp și vestimentație, pentru modelarea capului am optat pentru o tehnică diferită ce presupune sculptarea unui obiect pentru a ajunge la forma dorită. Această tehnică permite adăugarea detaliilor fine (eg., riduri) care printr-o modelare poligonală ar fi mult mai greu de obținut.

Interacțiunea din perspectiva ghidului

Dacă până acum doar un membru din public putea fi considerat utilizator al sistemului în cele ce urmează vom vedea că și ghidul face parte din aceeași categorie. Ghidului îi este atribuit avatarul care de data aceasta este o entitate în mediu virtual a cărui scop este doar să existe și să exprime în mediul virtual acțiunile ghidului din mediul real. În acest caz, avatarul ghidului este lipsit în totalitate de autonomia caracteristică agenților virtuali datorită faptului că execuția unei acțiuni la nivelul său presupune execuția acțiunii de însuși ghidul real, plasat fiind în sistemul MoCap.

Pentru a oferi utilizatorului control total asupra avatarului în timp real am implementat la nivelul modelului 3D al avatarului conceptele de armătură și weigh paint. Elementul "timp real" este esențial în această implementare deoarece procesarea datelor provenite de la sistemul MoCap în vederea aplicării lor pe sistemul osos al avatarului și calculul deformărilor corpului acestuia datorate armăturii trebuie să fie eficientizat la maxim, orice întârziere sau eroare de calcul producând efecte vizuale nedorite (tremurături ale avatarului, mișcări sacadate, blocaje, etc.).

Utilizatorul poate interacționa cu obiectele din mediul virtual prin intermediul avatarului său cu condiția ca avatarul să intersecteze, deci să atingă, obiectul în cauză. Dacă atingerea a fost realizată folosind mâna dreaptă utilizatorul va avea posibilitatea de a schimba poziția/orientarea obiectului.

Pentru activarea unui panou cu informații și opțiuni legate de obiectul deja selectat, utilizatorul poate afișa o postură ce implică utilizarea mâinii stângi, de exemplu.

Aspecte tehnologice

În implementarea sistemului nostru am ales să utilizăm soluția pentru body motion oferită de cei de la OptiTrack. Sistemul constă în 6 camere în infraroșu, un costum cu 36 markeri fotoreflexorizanți și pachetul software Arena de interpretare a datelor oferit de OptiTrack.

Redarea grafică 3D se realizează utilizând platforma AReVi. În aceeași platformă sunt implementați și agenții virtuali prin adoptarea unei abordări orientate multi-agent.

Redarea audio/video este realizată prin Skype[14], utilizând dispozitive specifice (webcam, sistem audio, proiector, căști wireless cu microfon, microfon, calculatoare asociate).

CONCLUZII

Sistemul propus oferă unui ghid turistic posibilitatea de a interacționa cu un mediu virtual cu scopul de a ajuta membri din public să înțeleagă mai ușor ceea ce le este prezentat. Publicul are posibilitatea de a interacționa cu mediul virtual în absența unui ghid, astfel interacțiunea cu mediul devine una personală.

Posibilitățile de extensie ale sistemului sunt multiple. În primul rând comunicarea dintre public și ghid, realizată momentan printr-o conexiune Skype, intenționăm să o integrăm direct în aplicație.

O altă direcție interesantă o constituie implementarea unui mecanism de recunoaștere a gesturilor la nivelul ghidului care va permite introducerea de noi metafore de navigare/interacțiune/prezentare a mediului virtual.

Nu în ultimul rând, posibilitatea de a deschide mediul mai multor grupuri de utilizatori, prin organizarea de tururi virtuale distribuite conduse de un același ghid, ne oferă o nouă perspectivă.

MULȚUMIRI

Dorim să mulțumim lui Matei-Ioan Popovici și Daniela Panait pentru contribuția lor la modelarea 3D a mediului virtual. Prezenta lucrare a fost parțial susținută din fondurile proiectului Tomis, PN II 11041/2007 finanțat de UEFISCDI.

REFERINȚE

1. 3DSMax - <http://usa.autodesk.com/3ds-max/>
2. AReVi - <http://svn.cerv.fr/trac/AReVi>
3. Avatar - http://www.hardocp.com/news/2009/12/19/how_mocap_wa_s_used_in_avatar
4. Blender3D - <http://www.blender.org/>
5. Ulicny, B., Thalmann, D., Crowd simulation for virtual heritage, Proc. First International Workshop on 3D Virtual Heritage, Geneva, pp.28-32, 2002.
6. D. Arnold et al, Tools for Populating Cultural Heritage Environments with Interactive Virtual Humans, EPOCH Conference on Open Digital Cultural Heritage Systems, 2008
7. eMove - <http://www.youtube.com/watch?v=YXa26PzpJ28>
8. Fifa 2009 - <http://www.youtube.com/watch?v=QoJfTzR6Vj4>
9. J. Maim, S. Haegler, B. Yersin, P. Miller, D. Thalmann, L. Van Gool, "Populating Ancient Pompeii with Crowds of Virtual Romans", In 8th International Symposium on Virtual Reality, Archeology and Cultural Heritage - VAST, 2007.
10. MetaMotion - <http://www.metamotion.com>
11. OptiTrack - <http://www.naturalpoint.com>
12. Organic Motion - <http://www.organicmotion.com>
13. Popovici D.M., Querrec R., Bogdan C.M., Popovici N., 2010. A Behavioral Perspective of Virtual Heritage Reconstruction, International Journal of Computers Communications & Control, ISSN 1841-9836, 5(5): 884-891.
14. Skype - <http://www.skype.com/intl/en/home>
15. TOMIS - <http://tomis.cerva.ro>
16. Xsens MVN - <http://www.xsens.com/en/general/mvn>