

Alberto Del Bimbo intervento a x media conference

Colloquio tra l'utente e il sistema

Questa mattina vi parlerò del colloquio tra uomo e sistema, e in particolare del colloquio tra uomo e sistema nel contesto delle applicazioni multimediali. Vi mostrerò inoltre, due esperimenti realizzati in collaborazione con il Master in Multimedia dell'Università di Firenze e con il Centro per la Comunicazione ed Interazione dei Media, due esperimenti d'interazione naturale come ha detto il Sig. Alagna, tra uomo ed applicazione multimediale.

Per cominciare iniziamo a riflettere sull'interfaccia posta tra l'uomo ed il sistema, un'interfaccia che oggi è sostanzialmente definita come se dessimo per scontato che il rapporto tra un uomo, il soggetto utente, ed un'applicazione multimediale sia un'interazione basata principalmente su informazione visuale. Come abbiamo potuto vedere durante questa conferenza, in tutto quello che è legato alla progettazione di esperienze d'interfacce per il web l'informazione visuale ha un ruolo dominante rispetto al resto.

Quando pensiamo ad un'interfaccia noi pensiamo alle metafore di interazione, alla sequenza di interazione, e a come progettare in modo che sia più attraente, ma sicuramente pensiamo alle immagini che usiamo per veicolare l'informazione e pensiamo anche a come progettare le caratteristiche visuali degli elementi che ci servono per realizzarla.

Questi aspetti della progettazione sono estremamente importanti, inoltre diamo per scontato che sia il mouse l'elemento tramite il quale si va ad interagire con un'applicazione multimediale sia questa mediata o meno dalla rete. Gli stessi studiosi che definiscono i principi di riferimento, quella che si chiama usability design, progetto di usabilità, stabiliscono delle misure secondo le quali un sito web risulta più usabile di un altro o più attraente di un altro facendo riferimento all'informazione visuale. Ci sono misure di apprendibilità di un sito, di efficacia (quanti errori si fanno per arrivare all'informazione), di efficienza (il tempo che ci mettiamo per concludere un certo compito su un sito), di soddisfazione (quanto ci riteniamo soddisfatti del rapporto con il sito ed in qualche modo gli rimaniamo fedeli). Ecco, tutte queste misure sono in qualche modo definite in funzione dell'informazione visuale, come sono le icone, quanto grandi, quanto chiare, quanti link mettiamo nelle pagine in modo da favorire la memorizzazione o la non memorizzazione. Ci preoccupiamo che il progetto del layout sia consistente, cioè che lo ripetiamo nelle diverse pagine di un sito in modo da dare omogeneità e quindi far sentire l'interfaccia il più leggera possibile, ci preoccupiamo del carattere dei font che usiamo, ci preoccupiamo dei colori, altro elemento importantissimo dal punto di vista della veicolazione dell'informazione, ci preoccupiamo di tutte quelle caratteristiche che sono eminentemente visuali.

Ora, se facciamo una riflessione sull'informazione multimediale, come informazione composta dei diversi media, potremmo concludere, che la possibilità di combinare in digitale diversi media, crea un contenuto informativo che tutto sommato si avvicina molto all'esperienza del soggetto nel contesto reale di tutti i giorni: vediamo, ma anche sentiamo, tocchiamo.

Quindi potenzialmente si apre la possibilità per un utente di interagire in modo naturale, cioè molto vicino al modo in cui farebbe la stessa cosa in un contesto reale. Ora per far questo è chiaro che il mouse sarebbe completamente da buttare, perché nel contesto reale difficilmente noi ci vediamo girare con un mouse in mano per interagire con gli oggetti, certamente sarebbe più diretto interagire sia attraverso la voce, oppure attraverso i gesti, cosa che tipicamente facciamo.

Da queste riflessioni nascono alcune ricerche che adesso vi presentiamo, in particolare ve ne presentiamo due che abbiamo fatto all'università di Firenze.

La prima è una ricerca in cui cerchiamo di catturare attraverso delle telecamere i gesti e i movimenti di un qualunque utente, li misuriamo, cioè comprendiamo di che movimenti si tratta, calcoliamo i gradi di libertà del sistema e le misure degli spostamenti in accordo a questi gradi di libertà; poi applichiamo queste misure a dei manichini in un contesto virtuale, di modo che, in tempo reale, cioè nello stesso momento che l'utente si muove ripreso dalle telecamere, il manichino si muova di concerto. Ora questa applicazione la sperimentaremo entro l'anno in un sistema che stiamo sviluppando insieme all'università dell'Illinois a Chicago, dove è stato provato per la prima volta, usato e poi diffuso in tutto il resto del mondo.

Il "Cave" è un ambiente di realtà virtuale, una stanza in cui l'utente entra equipaggiato di vari dispositivi: occhiali stereoscopici per visualizzare la tridimensionalità dell'esperienza, e puntatori tipo un mouse per spostarsi all'interno del Cave e quindi muoversi da un posto ad un'altro dell'ambiente virtuale presentato sulle pareti; in altre parole si dà all'utente l'opportunità di interagire con un ambiente virtuale, tramite gli stessi movimenti fatti nella realtà.

Con i colleghi dell'Università dell'Illinois cosa faremo? Loro hanno il Cave, noi no perché sfortunatamente costa un po' troppo, e l'Università di Firenze avendo dei problemi di bilancio non lo può comprare, però ci potremo collegare con loro attraverso la rete, allora: ci collegheremo al loro sistema, nel quale a Chicago ci saranno delle persone che entreranno nel Cave equipaggiate con tutti i loro dispositivi, noi invece non saremo equipaggiati, ma interagiranno in modo assolutamente naturale ripresi da delle telecamere qua a Firenze.

La nostra persona sarà ripresa, il sistema-software calcolerà i movimenti, queste misure saranno applicate ad un manichino virtuale che entrerà nell'ambiente del Cave al pari dei manichini delle altre persone, quindi potremo interagire con i colleghi americani. Loro con i dispositivi tradizionali che indossano, e noi invece senza indossare niente, ma interagiranno tutti e due, e ci vedremo in un contesto virtuale assieme. Questa è la prima cosa che vi farò vedere, poi ne discuteremo i contenuti.

La seconda cosa di cui parleremo, è un sistema che andremo ad installare già quest'anno qui a Firenze dentro un museo: è un sistema che abbiamo chiamato "Point at", cioè "punta a". Tramite esso faremo visitare ad un turista in un museo un affresco. In particolare l'affresco sarà quello posto nella Cappella di Benozzo Gozzoli a Palazzo Medici Riccardi, un'opera abbastanza importante perché è uno dei grandi capolavori del Rinascimento italiano. Vi sono rappresentati personaggi venuti a Firenze un po' da tutte le parti del mondo all'ora conosciuti, come il patriarca di Costantinopoli, banchieri famosi, alcuni componenti della famiglia Medici, ecc., in pratica risulta un'opera molto complessa da leggere e da raccontare.

I visitatori oggi si recano nella cappella, la guida racconta tutto, qualcuno guarda, chiede informazioni sui personaggi ritratti e la guida spiega... chi era per esempio Piero dei Medici e così via. Noi invece cosa vogliamo fare... creeremo una grande aula multimediale, proprio nella stanza dove dormiva Lorenzo dei Medici, ma questo è un caso era l'unica libera, ed in essa proietteremo la versione digitale dell'affresco, cioè l'affresco digitalizzato, reso striscia digitale, e consentiremo ad un utente di andare lì e puntare, ma senza niente in mano esattamente come farebbe nella Cappella, cioè tramite il dito, il personaggio a cui è interessato, il sistema basandosi su delle telecamere comprende dove l'utente punta, comprende il punto dello schermo su cui punta, e quindi comprende il personaggio puntato, capendo chi è il personaggio manderà in uscita il racconto relativo.

Vediamo il primo video:

-proiezione del video-

Quello che vedete qua è uno dei nostri ricercatori, vedete lui è in laboratorio, si sta muovendo, le sue mani, i suoi piedi e la sua testa hanno dei "blob", delle zone cromaticamente evidenziate a significare che le telecamere (due) che riprendono questo ricercatore acquisiscono due immagini, intanto un software di computer vision calcola quali sono la testa e le mani.

Mani e piedi sono state enucleate dal tronco, quelle palline che si muovono di concerto, e vedete le due immagini catturate dalle due telecamere, quella di destra e quella di sinistra, vedete come le palline di colore diverso seguono i movimenti che la persona sta facendo. Quindi il sistema di visione artificiale sta sostanzialmente seguendo in tempo reale quello che la persona sta facendo. Come fa a farlo? In questo caso è un sistema politicamente poco corretto, perché si basa sul colore della pelle, quindi se lì invece del nostro ricercatore che è europeo ci fosse una persona non europea, purtroppo non sarebbe seguito, o per lo meno bisognerebbe riprogrammare il sistema.

In pratica identifica le macchie cromatiche, le segue nel tempo e le isola calcolando 5 punti chiave, le mani i piedi e la testa. In realtà poiché il nostro obiettivo è riapplicare i movimenti che un uomo compie ad un manichino, questo con 5 punti si muove un po' maluccio, in realtà ne ha bisogno di più, semplificando abbastanza avremo bisogno di identificare 8 punti di articolazione, quindi gli altri punti che sono la vita per esempio o le spalle saranno calcolati in funzione dei 5 principali.

Quindi dati 5 punti se ne possono calcolare 8, e calcolandone 8 si possono ad esempio applicare al Dilbert, un tipo particolare di manichino, e farlo muovere esattamente come si muove la persona. Lui può compiere anche dei movimenti difficili... non so se apprezzate... adesso vedete che le mani si possono anche incrociare, e incrociandosi il software, la telecamera il sistema di visione, potrebbe anche perdere l'inseguimento, il sistema però è bravo abbastanza, perciò lo perde e lo recupera. Si può fare qualcosa di più però, ovvero applicarlo al vero manichino, allora vedete come questo manichino si piegherà... adesso... piega le gambe... piega il busto...

Se il primo non era un sistema politicamente corretto questo lo è, sia che uno sia basso, alto o di chissà quale colore non importa, il limite della testa ci dice la sua altezza e il sistema si adegua di conseguenza.

Il sistema, oltre ad essere corretto, è anche preciso perché qui vedete la persona (questo è un

laboratorio che abbiamo alla RAI) che sta puntando alla parete, sulla parete sono state messe delle targhette a distanza di 15 CM l'una dall'altra ed in alto di lato un pannello sul quale appare quale targhetta è puntata. Perché facciamo questo esperimento, per capire quanto è preciso il sistema, quindi l'utente punta una targhetta e guarda il sistema quale targhetta capisce essere puntata, quindi il quadratino nero dice qual'è.

Vedete allora, lui sposta la mano e la targhetta viene presa...ora da qui non potete capire quanto il sistema è preciso, ma in realtà lo è ad un altissimo livello, il sistema non sbaglia...fa un errore di circa 20 centimetri che può essere tranquillamente gestito progettando l'interfaccia.

Questa è l'interfaccia reale, questo è un prototipo che abbiamo presentato due settimane fa per la prima volta ad una manifestazione concorrente che c'era prima di questa che si chiamava WEB CITY, e ci fa piacere presentarlo anche oggi...lo installeremo a giugno probabilmente nel Palazzo Medici Riccardi. Quello che vedete è un grande schermo su cui è proiettato l'affresco...vedete un cerchio intorno ad alcuni personaggi che indica le zone dell'affresco che cliccabili, diciamo così, il puntamento di una persona fa esattamente quello che farebbe un mouse in un'applicazione convenzionale. Questa punta, il sistema calcola dove sta puntando, e come un mouse quando lo posizioniamo e lo premiamo, apre una finestra, in pratica capisce se si sta puntando in una zona precisa, capisce se quella zona contiene informazione indicandolo con il cerchio, e ci darà modo di visualizzarle se lo si vorrà. Quindi il sistema è molto semplice, la metafora di interazione è una metafora semplicissima, che è: punta, evidenzia, scorri, poiché l'affresco è una striscia.

Sono le tre azioni che se, pensate bene, l'utente farebbe se andasse nella cappella...punterebbe a qualcosa che lo interessa e chiederebbe di conseguenza alla guida informazioni, oppure si girerebbe nella sala ad osservare altre parti dell'affresco, risulta essere allora un'interazione assolutamente naturale.

Ora vediamo qui qualche esempio...vedete che la persona (ce ne sono di diverse, prima era una ragazza, ora è un uomo) sta puntando in alto a sinistra dove ci sono delle frecce, che sono quelle che ci consentono di scorrere, quindi lo scorrimento lungo l'affresco è possibile puntando ad esse...come l'utente punta si scorre, e quindi è come se l'utente girasse nella sala dove si trova l'affresco...punta...scorre...ancora...poi se ne va ed arriva un'altra persona. Quando non si vuole più scorrere, ma puntare a qualcosa, in questo caso si è puntato alla Natività, questa viene evidenziata, quindi viene fatto uno zoom di ciò che viene puntato e si dà l'opportunità di visionare le informazioni.

Questo è come funziona, la persona in tempo reale viene ripresa, vengono calcolati questi punti apparentemente semplici ma in realtà molto complessi, si calcola dove sta puntando e si dà la risposta.

Vi ho presentato questi esempi per mostrarvi una modalità naturale di interazione, è chiaro anche che comunque non tutte le applicazioni sono idonee a questo tipo di colloquio ...sostanzialmente sono idonee quelle applicazioni con le quali vogliamo riprodurre in un contesto digitale, un contesto reale, o comunque vogliamo rinforzare l'esperienza del reale. Per esempio in questo caso...perché in un museo realizzare tutto questo?

Tutti voi avrete avuto modo di vivere la devastante esperienza di usufruire dei dispositivi multimediali che affollano i musei italiani, sono pessimi, pessima la qualità, ma soprattutto pessimo il concetto per il quale si pensa che chi va al museo si debba sedere davanti ad uno schermo, ad un supporto multimediale e si debba sorbire ciò che qualcun altro ha pensato sia di interesse per lui, se lo fa e se la cosa è realmente di grande interesse probabilmente non resiste più di qualche minuto...non resiste più di qualche minuto perché la sua intenzione non era quella di andare al museo e di vedere l'applicazione multimediale, ma di andarci e vedere l'opera d'arte reale.

Il problema qual'è, ormai i musei italiani ci offrono dei tempi contingentati...non si può stare più di cinque dieci minuti in un'aula, dopodiché si deve passare a quella successiva, insomma è una visita a volo di uccello, non rimane quasi niente. Così anche per la cappella che abbiamo visto nell'esempio, questa poi ha anche l'aggravante di essere uno spazio molto ristretto e quindi visitabile da poche persone alla volta nonostante che ci sia la necessità di doverne fare entrare tante, questo è, infatti, il 3° museo per numero di visitatori di Firenze.

L'esperienza multimediale è un'esperienza che serve a rafforzare quella che l'utente farà poi nel reale, punta, chiede, guarda, interroga, come se avesse una guida personalizzata potendo inoltre selezionare la lingua desiderata. Per rinforzare l'esperienza reale, dobbiamo studiare dei meccanismi che replicano esattamente l'esperienza reale e se proponiamo dei meccanismi di interazione diversa, l'esperienza è un'esperienza diversa e non è detto che sia un'esperienza che rinforza quella conclusiva. Quindi questa è un po' la motivazione che sta dietro a questi tipi di ricerca.

MRK: Grazie dell'intervento, volevo fare un paio di domande al professore con qualche commento in realtà: E' interessante questa considerazione del rapporto con la realtà, se ho ben capito quello che si propone in questi progetti, non è più un simulacro del reale ma una realtà ampliata, vale a dire che probabilmente vado a vedere "oltre" a quello che potrei percepire e, nel caso di un affresco osservato da una certa distanza, arrivare ad avere una macropercezione dell'oggetto. Ma in questo senso mi chiedevo quale possa essere la fine di quella che qualcuno ha chiamato l'aura sacrale dell'opera d'arte?

Del Bimbo: Come dire, io penso che sarà bene sempre distinguere il reale dal virtuale, e personalmente appartengo a quella categoria di persone che pensano che l'esperienza virtuale sia un'esperienza poco soddisfacente.

MRK: anch'io...

Del Bimbo: però questo non significa che se è poco soddisfacente sia non utile, quindi l'esperienza reale francamente sarà bene mantenerla. In particolare se uno va ad un museo.... insomma una cosa è vederselo in rete e una cosa è entrarci veramente, ci sono tutto un contesto di emozioni non replicabili. Penso che la ricerca su queste cose sia una ricerca di rapporto uomo-macchina, che non deve essere la replica del reale per sostituirsi al reale, ma deve essere una cosa distinta. Penso che ci sia una familiarità nei prodotti che ci ha mostrato Matteo, come il riferimento ai caroselli, ed un'ironia che è fortemente appartenuta ad un nostro modo di promuovere i prodotti. Beh, questo mi sembra decisamente significativo ed interessante anche nell'ottica degli obbiettivi che si pone questo contesto. Grazie.