

William Buxton
Silicon Graphics

Els pensaments intercanviats entre dues persones
no són els mateixos en una habitació que en una altra habitació.

Louis I. KAHN

Resum

El 1991, Mark Weiser va publicar un article en el qual esbossava el projecte de la Xerox PARC sobre la propera generació informàtica (Weiser, 1991). Es va referir a aquest model amb el nom de *Ubiquitous Computing* (Informàtica Omnipresent) o *UbiComp*. En l'article següent introduïm un component complementari al pla de Weiser: l'anomenem *Ubiquitous Video* (Vídeo Omni-present) o *UbiVid*.

El treball preliminar a UbiVid va partir de la recerca en els «espais mediàtics» (Gaver *et al.*, 1992; Mantei *et al.*, 1991; Stults, 1986; Bly, Harrison i Irwin, 1993). Els conceptes que analitzem es basen en aquests treballs. El nostre punt de vista és que UbiComp i UbiVid són els dos vessants del mateix projecte. Junts, conformen el que hauríem d'anomenar, per ser més precisos, *Ubiquitous Media* (espais mediàtics omnipresents). Creiem que la noció d'espais mediàtics omnipresents esdevé un model útil per dissenyar sistemes futurs i els models d'ús.

Aquest article es basa en la recerca duta a terme pel Projecte de Telepresència d'Ontario (Ontario Telepresence Project) i de Xerox PARC entre el 1990 i el 1994. La majoria de les idees que exposarem han estat implementades o se n'han fet prototipus. Tanmateix, el propòsit no és informar sobre la recerca *per se*, sinó que la voluntat és l'estímul i l'explicació de l'armadura arquitectònica i espacial del nostre model.

Com a resum del projecte, el nostre argument és que no hauríem d'estar tan interessats en l'arquitectura del ciberespai com en la ciberització de l'espai arquitectònic.

Informàtica omnipresent: una breu ressenya

Tal com ho va descriure Weiser, l'Ubiquitous Computing (UbiComp o Informàtica Omnipresent) es pot caracteritzar a partir de dos trets principals:

- Omnipresència: les interaccions no es canalitzen per mitjà d'una única estació de treball. L'accés informàtic és «pertot». Per exemple, en el despatx hi hauria d'haver desenes d'ordinadors, pantalles, etc. Les mesures haurien d'anar de tabuladors de la mida d'un rellotge, a *blocs de notes* de la mida d'una llibreta i fins a plafons de la mida d'una pissarra. I tot plegat connectat per xarxa. S'hauria d'accedir a les xarxes sense fils per donar suport als accessos remots i mòbils.
- Transparència: aquesta tecnologia no és intrusiva i és tan invisible i integrada en l'ecologia general de la llar o del lloc de treball com, per exemple, l'escriptori, la cadira o un llibre.

Aquests dos trets fan palesa una paradoxa evident: com pot ser que alguna cosa sigui pertot arreu i al mateix temps sigui invisible? Resoldre aquesta paradoxa ens condueix a l'essència del projecte subjacent. La solució no és que no puguem veure (sentir o tocar) la tecnologia, sinó que la seva presència no causi cap intromissió en l'entorn del lloc de treball (tant en relació amb l'espai físic com amb les activitats que s'hi realitzen). Com la tecnologia convencional del lloc de treball (l'arquitectura i el mobiliari, per exemple), la utilització és evident i l'aplicació física s'ha dissenyat de manera específica per a l'espai i la funció que havia d'acomplir. El projecte central d'UbiComp és poder trencar amb el model informàtic de Henry Ford i que es pot parafrasejar de la manera següent:

Pot adquirir la forma que vulguis, sempre que tingui un ratolí, un teclat i una pantalla.

No té cabuda en aquest projecte de tecnologies futures voler fer encaixar el motlle quadrat dels dissenys convencionals, com ara la GUI, en l'orifici rodó que permet d'assolir les necessitats i les aplicacions reals.

La tecnologia s'escalfa

Podem, sense dificultats, situar el model de Weiser en la perspectiva històrica per mitjà de l'analogia amb els sistemes de calefacció. En temps antics, l'arquitectura (o almenys en els climes freds) s'havia d'ajustar a la necessitat de contenir l'escalfor. Es van construir estructures especials per contenir foc sense cremar-ho tot. De la mateixa manera, durant la primera època, es van construir estructures per acollir la informatització. Van rebre el nom de *centres informàtics*.

A mesura que l'arquitectura va progressar, es van construir edificis on els focs estaven continguts en llars de foc, de manera que es podia escalfar més d'una cambra. De tota manera, només en algunes habitacions especials s'hi podia fer foc, ja que calia que la llar de foc fos al costat d'una xemeneia. De la mateixa manera, la generació següent d'ordinadors va poder arribar a cambres de fora dels centres informàtics. Malgrat tot, calia que aquestes cambres tinguessin uns cables elèctrics especials i aire condicionat. Per tant, la informàtica continuava restringida a unes «cambres informàtiques» especials.

La generació següent de sistemes de calefacció va permetre les estufes Franklin i fins i tot els radiadors. Ja podíem produir escalfor a totes les habitacions. Però això requeria que les «canonades» distribuïssin el sistema. La intromissió de les «canonades» en l'espai habitable era vista com un preu menor que calia pagar a canvi de distribuir l'accés a l'escalfor. Una vegada més, aquesta situació no és gaire diferent de la que es va produir en la generació següent d'ordinadors (la generació que tenim actualment), on tenim accés a la informàtica distribuïda, sempre que estiguem connectats a una infraestructura de «canonades». I com passa amb el sistema de calefacció, això ha implicat la intromissió dins de l'espai i una «àncora» que limita la mobilitat.

Això ens condueix a la pròxima generació de sistemes de calefacció (l'actual): el control climàtic. En aquest moment, els aspectes de clima interior (l'escalfor, l'aire condicionat, la humitat, etc.) es poden controlar cambra per cambra. Allò que ho permet és invisible i probablement desconegut (bombes de calor, gas, gasoil, electricitat?). Només hi ha present el mecanisme que permet de controlar l'ambient segons les preferències personals. És el tipus de calefacció que equival a UbiComp: el servei és omnipresent, tot i que la distribució és invisible. En aquesta fase de maduresa, la tecnologia és perfecta i s'integra en l'estructura del lloc de treball.

Per això, dins del model d'UbiComp, no hi ha cap ordinador al meu escriptori, perquè l'escriptori és l'ordinador. Com passa actualment, hi ha una gran pissarra blanca a la paret, però amb UbiComp és activa, es pot connectar a la vostra, que potser és a tres mil quilòmetres. Hi veig molta menys tecnologia. Aconsegueixo menys intromissions (soroll, calor, etc.) i molta més funcionalitat i adaptabilitat. I amb els blocs de notes i els tabuladors, i les xarxes sense fils que faig servir, també aconsegueixo molta més mobilitat sense convertir-me en un «orfe» informàticament.

Vídeo omnipresent

El vídeo omnipresent (UbiVid) és el vídeo complementari d'UbiComp, en tant que comparteix els dos trets d'*omnipresència* i *transparència*. És al vídeo de sobretaula el que UbiComp és al GUI.



FIGURA 1. Una configuració típica de videoconferència de sobretaula. La conferència es canalitza habitualment per mitjà d'una càmera de vídeo al damunt d'un monitor situat a l'escriptori de l'usuari. Fotografia: Projecte de Telepresència d'Ontario.

Per exemple, en les videoconferències de sobretaula, allò que veiem és un usuari darrere d'una taula que parla amb algú que apareix en un monitor que té una càmera de vídeo al damunt. La figura 1 il·lustra aquesta situació. En aquests sistemes, les interaccions per vídeo es limiten a aquest parell únic de càmera i monitor.

L'UbiVid permet de trencar aquesta imatge, de la mateixa manera que UbiComp permet deixar de centrar tota l'activitat mediada per un ordinador en un únic ordinador de sobre la taula. En comptes d'aquesta situació, la premissa és tenir un ampli ventall de càmeres de vídeo i monitors en el lloc de treball, i tot disponible. Gràcies a la disponibilitat d'entrades i sortides de vídeo de mides diferents i en diversos llocs, es pot aconseguir el concepte més important que hi ha al darrere d'UbiVid: aprofitar la relació entre la funció social i el lloc físic.

A continuació explorarem la importància d'aquesta relació. Començarem per articular alguns dels principis subjacents del disseny i, en acabat, presentarem alguns exemples.

- Principi de disseny 1: preservar les relacions funció/lloc tant per a activitats locals com distants.
- Principi de disseny 2: tractar les «presències» físiques i electròniques o els visitants de la mateixa manera.

— Principi de disseny 3: utilitzar els mateixos protocols socials per a les interaccions socials electròniques i físiques.

L'anatomia social i espacial de la meva oficina

Podem desenvolupar aquests principis a través d'alguns exemples.

La figura 2 correspon a la meua antiga oficina de la Universitat de Toronto. S'hi descriuen alguns dels espais concrets:

- A. La cadira darrere l'escriptori.
- B. La cadira davant de l'escriptori.
- C. Al costat de la cadira.
- D. Cadires al voltant de la tauleta.
- E. L'entrada.

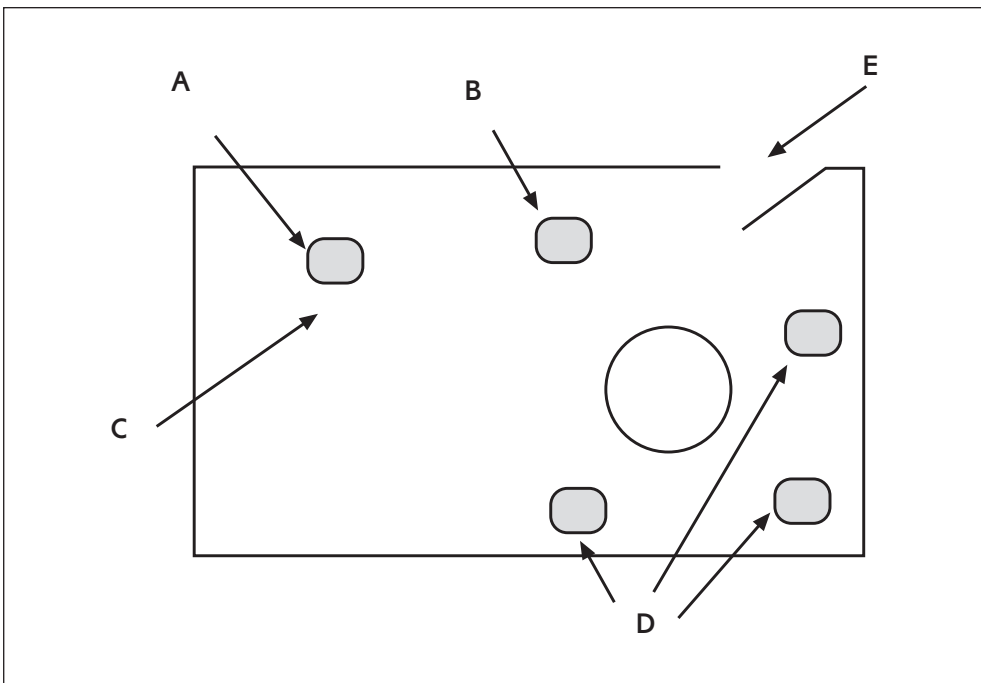


FIGURA 2. Esquema del meu despatx. S'hi indiquen diversos llocs, inclosa la cadira darrere de l'escriptori (A), la cadira davant de l'escriptori (B), espai per poder-me posar dret darrere de l'escriptori (C) i les cadires al voltant de la tauleta (D). S'associen diverses funcions socials a cada posició. La utilització de qualsevol tecnologia en aquest despatx que permeti la col·laboració o les activitats socials ha de reflectir aquestes diferències i respectar-les.

Fins i tot en aquest espai relativament simple, s'associen diverses interaccions socials i protocols amb cadascun d'aquests espais. Prenem com a exemple la reunió amb un estudiant.

Primer, em puc asseure a la cadira (A) i que les visites s'asseguin a l'altra banda de la taula. En aquest cas, sóc el professor Buxton i ells, no. Puc fer servir aquesta posició si he de dir a un estudiant que ha suspès o si el felicito formalment per haver trobat una feina molt bona.

Segon, si volgués treballar conjuntament amb l'estudiant, o en una situació similar, potser aquest estudiant vindria al darrere de la taula, a la posició C, mentre jo continuo assegut a la cadira. De tota manera, no seria gaire habitual que una persona desconeguda, o amb qui no treballés de prop, se situés darrere de l'escriptori.

Tercer, si fóssim al mig d'una reunió informal, o bé potser només xerrant, potser estaríem al voltant de la tauleta. Això passaria si la reunió fos informal i indicaria que la relació és entre iguals i no pas subordinada. Seria una reunió amb en Bill i no amb el professor Buxton.

Quart, podria ser a l'escriptori treballant i algun estudiant podria treure el cap per la porta per preguntar alguna cosa. Si no el faig passar, l'estudiant entén que tinc feina i que la reunió ha de ser breu.

Finalment, puc decidir de fer certes reunions fora del despatx si considero que el lloc no és l'adient. Això podria passar si hagués d'impartir un curs en un seminari, o bé anar-hi d'oient, i el lloc més adient seria una sala de conferències (que té unes convencions pròpies associades amb els diversos indrets dins de la cambra).

La premissa és que qualsevol tecnologia que s'hi introdueixi ha de reflectir aquest respecte per les relacions espai-funció. Conseqüentment, hem de traslladar-nos al lloc apropiat a l'activitat, més que no pas traslladar l'activitat allà on som nosaltres. L'assumpció que ho justifica és que la incomoditat d'haver d'anar al lloc adient és molt menys elevada que fer alguna cosa en el lloc erroni. Aquest criteri és coherent amb la majoria de les nostres experiències en el món físic i força divergent de la línia habitual en el camp del disseny informàtic.

A continuació, desenvoluparem alguns exemples de disseny d'entorns que segueixen alguns dels principis per a reunions mediades tecnològicament.

Exemple: al voltant de l'escriptori o al voltant de la tauleta

Podem començar amb el cas més simple d'haver de trobar-te amb algú al despatx mitjançant una connexió de vídeo. Tal com hem argumentat anteriorment, la majoria de les oficines equipades per poder realitzar aquesta funció tenen un disseny similar a l'il·lustrat a la figura 1.

Però una configuració així incompleix els principis 1 i 2, ja que totes les transaccions per vídeo es produeixen a partir d'un únic monitor situat en una posició fixa. Això significa que no es pot treure profit de les relacions lloc/funció. També provoca algunes contradiccions quan s'encavalllen peticions de serveis (com ara quan algú vol posar-se en contacte amb mi a través d'una videoconferència i joestic mirant un vídeo).

Això no vol dir pas que mai no funcioni. Fora dels casos que analitzarem més avall, aquesta configuració pot ser l'adequada per a reunions en què jo sóc a l'escriptori i la funció social de la persona remota és semblant a les associades amb els indrets *B* o *C* segons l'esquema de la figura 2.

Però què passaria en la situació d'una reunió informal d'un grup al voltant de la tauleta? En la configuració estandarditzada del vídeo de sobre la taula, com podria assumir la persona remota el lloc que hauria d'ocupar a *C*?

La nostra proposta va ser col·locar el vídeo «terminal» en cadascun dels indrets que el participant remot podria ocupar dins de l'espai de la sala si hi fos físicament. Per això, hi ha un sistema de vídeo a l'escriptori (figura 1), de manera que pugui llegir els videodocuments i fer treball «de prop» amb el col·lega remot. En acabat, també hi ha un sistema a la tauleta (figura 3), on el visitant «s'asseuria» i participaria de les converses al voltant de la tauleta.

A l'exemple, es poden conservar les relacions entre la funció i l'espai. El visitant «electrònic» s'asseu al lloc que ocuparia com a visitant físic. De la mateixa manera, el company virtual de despatx s'asseu al lloc on seria el físic. Si l'equipament se situa de manera adient, el visitant pot veure el company de despatx, el qual pot veure el visitant, etc. Gràcies a aquest ús distribuït de l'espai, els conflictes de recursos es redueixen i les convencions socials es conserven.



FIGURA 3. Participació remota en un grup informal. Aquí, un grup (incloent-hi el participant remot que apareix a la pantalla) és assegut al voltant de la tauleta del meu despatx enmig d'una reunió no formal. Segons l'esquema de la figura 2 es troben a la posició *D*.

Fotografia: Projecte de Telepresència d'Ontario.

Exemple: vista de la paret des de la porta

Fins ara els exemples que hem analitzat estaven relacionats amb posicions concretes. A més del lloc i la distància, però, també compta la possibilitat de moure's d'un lloc a l'altre dins del mateix espai. Hi ha unes pautes socials molt ben determinades sobre la manera com ens podem moure, com per exemple com ens podem acostar els uns als altres, o bé quan ens n'anem. L'espai físic implica una continuïtat, mentre que les tècniques convencionals amb un vídeo damunt una taula, com a la figura 1, defineixen un únic punt. Amb un sistema així, ningú no se'm pot acostar des de lluny. En totes les transaccions veig l'altra persona davant dels meus ulls. Els altres arriben i se'n van de manera abrupta i són molt més a prop del que serien si no ens coneguéssim molt. En resum, s'incompleix el comportament social normal. I no caldria que fos així.

Premissa: la possibilitat de moure's entre els espais és tan important com el suport que reben els espais individualment.

La figura 4 mostra com hem solucionat aquest problema. Quan entreu al meu despatx, ho feu per la porta (lloc E de la figura 2). Si entreu físicament, tot és normal. Si entreu de manera electrònica, apareixeu en el monitor que hi ha al costat de la porta, us sento des de l'altaveu del



FIGURA 4. Mantenir una distància social. A l'hora d'establir el contacte, algú apareix per la porta i té una visió frontal per mitjà de la càmera, tant si aquesta persona s'acosta pel corredor físic (imatge de l'esquerra) o pel corredor electrònic (imatge de la dreta). Es mantenen les convencions socials d'apropament i es fan servir les mateixes convencions socials per als visitants físics i electrònics.

Fotografia: Projecte de Telepresència d'Ontario.

costat de la porta i em podeu veure des d'una càmera de baixa resolució i d'angle ample que també és al costat de la porta. Així, el primer cop d'ull des del corredor electrònic és essencialment el mateix que si entréssiu per la porta del corredor físic.

Si tinc una altra feina, o bé sóc amb algú altre, potser no us veuré o no us faré cas, de la mateixa manera com passaria si entréssiu des del passadís (encara que sí que sentiria que hi ha algú o que algú ha entrat). Es manté la distància apropiada. Si truqueu a la porta i em saludeu, us puc fer passar. En aquest cas, ocuparíeu la cadira dels «visitants», és a dir, el monitor de visitant que hem vist a la figura 3.

Exemple: posició de la porta i accessibilitat

L'exemple anterior ens ha demostrat com es pot conservar la distància tant dels visitants electrònics com dels físics i mantenir la distància social de la porta. Encara podem concretar més.

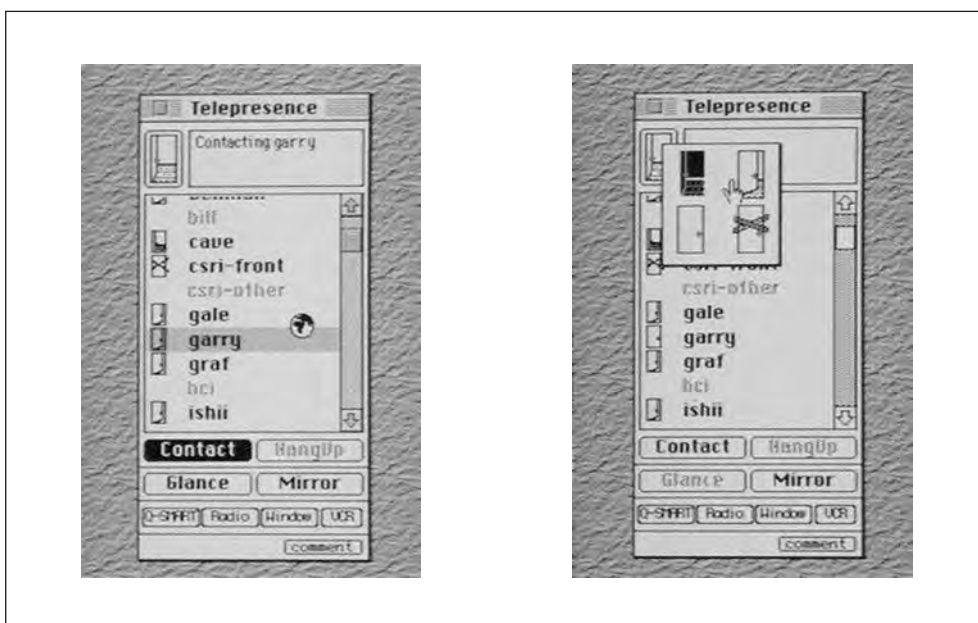


FIGURA 5. La condició «posició de la porta» especifica l'accessibilitat. La figura il·lustra la tècnica dels usuaris de l'espai mediàtic per controlar la pròpia accessibilitat a partir de les mateixes estratègies que es fan servir en els espais físics, és a dir, per mitjà de la posició de la porta. En aquest cas, hi ha la icona d'una porta que pot aparèixer en alguna de les quatre posicions següents: oberta, entreoberta, tancada o pas barrat. Cada posició indica un nivell d'accessibilitat diferent. Els visitants potencials poden determinar l'accessibilitat personal per mitjà de la posició de la porta que s'indica al costat del nom, tal com es pot veure a la imatge de l'esquerra. Cadascú marca la posició de la porta i, per tant, la disponibilitat per mitjà d'un menú senzill, tal com es pot veure en la imatge de la dreta.

Fotografia: Projecte de Telepresència d'Ontario.

La mateixa porta exerceix el control de l'accessibilitat dels visitants físics. Si és oberta, no hi ha cap problema perquè entreu. Si és entreoberta, podeu entrar al despatx i comprovar si estic enfeinat. Segurament, trucareu a la porta i esperareu una resposta abans d'entrar. Si és tanca-da, trucareu i esperareu una resposta abans d'entrar. Si hi ha un cartell de «No molesteu», no trucareu, però potser deixareu un missatge.

Premissa: és tan important entendre les diferències de la distància social i respectar-les com ho és comprendre les diferències de la distància física.

D'acord amb el principi 3, així hauria de ser per a les visites electròniques, independentment de si l'aproximació es fa per connexió telefònica o de vídeo.

La figura 5 representa la interfície, que va ser inicialment proposada per Abi Sellen, que utilitzem per transferir els protocols al domini electrònic. Amb aquesta interfície, cadascú determina la pròpia accessibilitat per mitjà de la selecció d'una de les quatre possibles posicions de la porta. Es pot fins i tot deixar una «nota» a la porta virtual per tal de fer arribar un missatge als visitants.

Tot i que es conserven els protocols del món físic per mitjà d'una metàfora, aquest disseny, tanmateix, encara no permet de complir el principi 3 en la seva totalitat. La raó és que, tot i que els protocols són paral·lels, no n'hi ha únicament un.

Es podria aconseguir si la porta física controlés l'estat de la disponibilitat personal tant per als visitants electrònics com per als físics alhora. Així (evidentment segons l'habilitat d'evitar errades), l'ordinador hauria de detectar que tanco la porta i evitar que entrés ningú ni físicament ni electrònic (per telèfon o per vídeo). El mètode que va implementar una de les col·laboradores del meu equip, Andrea Leganchuck, es pot veure a la figura 6. D'aquesta manera vam poder aplicar el principi 3: un protocol que ho controla tot.¹

Una bona part dels conceptes anteriors es basa en la noció que el lloc físic dels participants influeix considerablement en les interaccions socials durant les reunions personals. L'objectiu és, des d'un punt de vista de la perspectiva del disseny, demostrar que aquesta mateixa concepció pot ser emprada, i aprofitada, en la telepresència. Per tant, quan parlem de distància entre els participants, cal distingir entre la distància física que mantenen vers meu i la distància entre la seva persona per vídeo i jo. Aquesta darrera, més que no pas la primera, determina la distància social.

Premissa: la distància física i el lloc de la presència per vídeo respecte de la meva situació té el mateix pes/funció/aspecte socials que si fóssiu físicament en aquell lloc.

1. De fet, no és tan senzill. El significat de la posició de la porta és cultural, per tant cal considerar una problemàtica més àmplia a l'hora d'aplicar aquest criteri si volem que funcioni. De la mateixa manera que la capacitat de qualsevol sistema augmenta quan és sensible al context en què ha de reaccionar, també la qualitat i la flexibilitat de les eines de l'usuari han de permetre de concretar aquestes accions. Els exemples que presentem volen establir una manera diferent de pensar en els sistemes. No tenim la voluntat que esdevinguin un dogma per a dissenys concrets.

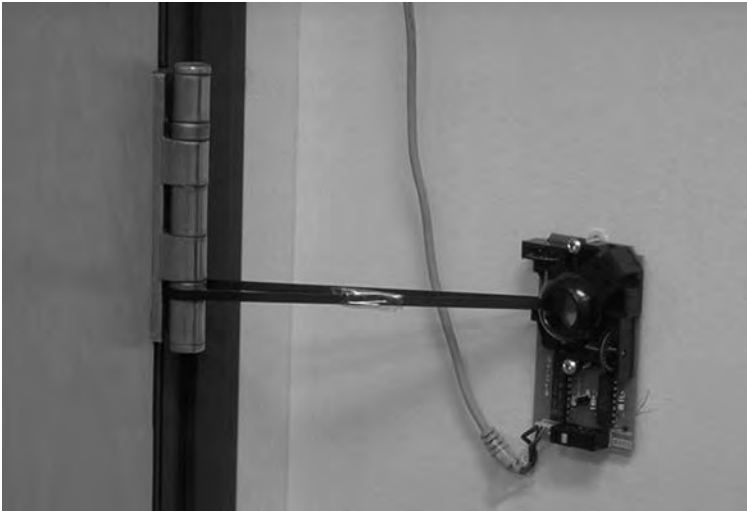


FIGURA 6. La «porta ratolí». És un ratolí MacIntosh sense la part exterior i que s'ha muntat a la paret del costat de la porta. Un mecanisme de cinta connecta la frontissa de la porta a un dels codificadors d'arbre del ratolí. D'aquesta manera el ratolí reconeix l'obertura de la porta i trasllada aquesta informació a l'ordinador per tal que estableixi l'estat de la porta (vegeu la figura 5).

Fotografia: Projecte de Telepresència d'Ontario.

A més, l'assumpció és que això és cert, independentment de la distància física real respecte a mi.

Qualificació: aquesta equivalència depèn de si el disseny és adequat. Estableix els estàndards i els criteris de disseny i avaluació.

Exemple: videoconferència d'extrem a extrem

De la mateixa manera que hi ha diversos llocs apropiats, segons les diverses interaccions dins del meu despatx, les diverses cambres són adients per a interaccions diferents dins dels edificis. Voler acomodar tots els tipus d'interaccions en una mateixa sala és tan poc apropiat com voler permetre tots els tipus d'interaccions de cada cambra en un mateix lloc. Tenint present l'arquitectura física, hem ampliat el lloc de treball a altres sales, cadascuna instrumentada, i compatible, de manera coherent amb el meu despatx.

Una d'aquestes cambres és la sala de conferències on es fan reunions d'unes dotze persones. Com qualsevol altra habitació, aquesta sala de conferències té unes funcions socials diferents associades amb els diversos indrets. Per exemple, durant una presentació, qualsevol que entri a la cambra sap qui és el conferenciant i qui és l'audiència, encara que ningú no parli, només pel lloc on és físicament cadascú. En general, el confe-



FIGURA 7. Videoconferència convencional amb el presentador en un «extrem de la sala». Totes dues fotografies mostren una videoconferència en una sala de reunions on el participant remot és el conferenciant. A la imatge de l'esquerra, la persona remota només parla i apareix en un monitor més gran. A la fotografia de la dreta, el conferenciant parla d'un document. Com que aquest és l'element més important, el document apareix en un monitor més gran i el conferenciant en un monitor adjacent més petit.

Fotografia: Projecte de Telepresència d'Ontario.

renciant és dret en un extrem de la sala, al costat de la pissarra, amb l'audiència asseguda al voltant de la taula.

Per això, per tal de permetre les presentacions de participants remots, la majoria de sales de conferència equipades per a videoconferències tenen l'aparell de vídeo en un extrem de la sala. La figura 7 és un exemple d'aquest tipus de configuració.

Però què passa si algú que hi és físicament fa una presentació i el participant remot forma part de l'audiència? Si la sala té la configuració de la figura 7, aleshores l'espai que ocupa el participant remot envaeix el del conferenciant. A més, el participant remot serà al darrere del conferenciant, no li veurà més que l'esquena i el conferenciant no el podrà veure.

Malauradament, aquesta és la situació en què m'he trobat virtualment en tots els espais on he fet una presentació en viu davant d'una audiència física i remota alhora. És un error des de qualsevol perspectiva social imaginable i es pot solucionar amb la mateixa celeritat com es comet l'error.

La solució és el que anomeno *videoconferència d'extrem a extrem*. Com a contrast de les sales de videoconferència tradicionals, la càmera i els monitors són al fons de la sala, tal com il·lustra la figura 8.²

2. De fet, la sala també permet la conferència des d'un únic extrem, només que així l'omnipresència es mou cap a una fase ulterior.



FIGURA 8. Videoconferència d'extrem a extrem. Els assistents remots a una reunió ocupen un lloc al voltant de la taula per mitjà de monitors de vídeo situats a la paret del fons. Hi veuen per la càmera adjacent, hi senten per mitjà dels micròfons i parlen per l'altaveu del monitor. El conferenciant utilitza les mateixes estratègies convencionals per interaccionar amb assistents físics i els electrònics. No calen habilitats especials.

Fotografia: Projecte de Telepresència d'Ontario.

Com en la meua oficina, la sala de conferències permet de fer les videoconferències des de qualsevol extrem. Com en l'espai físic, l'entorn hauria de permetre als participants d'assumir qualsevol paper, des del lloc adient, tant si hi són físicament com si hi participen per telepresència. A més, com en el despatx, la sala s'hauria de poder acomodar a les transicions, o moviment, de papers (i per tant de llocs).

Finalment, en el nostre exemple de videoconferència d'extrem a extrem, es pot veure que hi ha transparència en les interaccions socials entre els participants. Pel fet que es manté una reciprocitat entre l'àudio i el vídeo, juntament amb el respecte a l'espai personal, el conferenciant fa servir els mateixos mecanismes quan interacciona amb els assistents locals i els remots. Dit d'una altra manera, fins i tot quan el conferenciant no té experiència amb les videoconferències o amb la tecnologia, no cal que aprengui cap nova «interfície d'usuari». Si algú alça la mà, és evident que vol fer una pregunta. Si algú fa cara de no entendre alguna cosa, se li poden fer aclaraments. Més que aprendre noves tècniques, el disseny fa ús de les habilitats que ja existeixen i que s'han anat adquirint al llarg dels anys de viure en el món quotidià.

Exemple Hydra: reunió a quatre bandes en una taula rodona

Conjuntament, l'oficina i la sala de conferències ofereixen la possibilitat de realitzar un ampli ventall de reunions diferents. Però de la mateixa manera que hi ha més dos tipus de sales de reunions en un edifici, també hi ha d'haver altres tipus d'espais habilitats tecnològicament que permetin totes les interaccions socials que es poden trobar de manera habitual dins de les organitzacions.

En aquest exemple, introduïm una tècnica que facilita les reunions a quatre bandes, on cadascun dels participants ocupa un lloc diferent. Va ser dissenyada per tal de captar la majoria d'insinuacions espacials causades per les mirades, els moviments del cap, la percepció de les mirades i el torn de paraules que es produeixen en les reunions personals. Amb la intenció de ser coherents amb els principis dels dissenys descrits més amunt, partim del respecte per les relacions espacials «al voltant de la taula».³ La figura 9 il·lustra aquesta proposta.

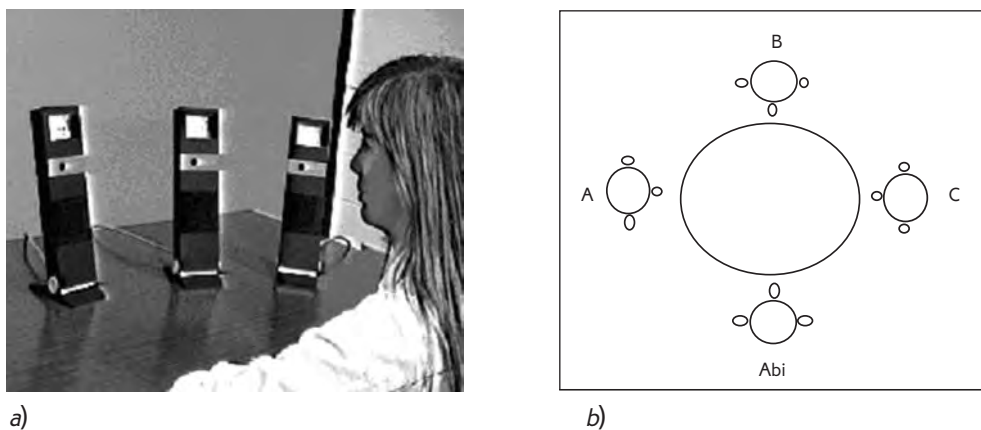


FIGURA 9. Utilització de «substituts» per vídeo per facilitar les videoconferències a quatre bandes. La figura a) mostra una videoconferència a quatre bandes, on cadascun dels tres participants remots participa a través d'un «substitut» per vídeo. Pel fet de conservar les relacions de «taula rodona», tal com il·lustra la figura b), es mantenen els actes conversatoris de les reunions presencials, com ara la percepció de les mirades, els moviments del cap, etc.

Fotografia: Projecte de Telepresència d'Ontario.

3. Aquesta idea de fer servir el vídeo com a substitut en aquestes reunions a diverses bandes resulta que no és nova. Després que nosaltres l'apliquéssim, vam saber que havia estat proposada anteriorment per Fields (1983).

Tal com hem indicat a la figura de l'esquerra, aconseguim aquest objectiu amb cadascun dels tres participants remots representats en una petita unitat de vídeo, que anomenem col·lectivament *Hydra* (Sellen, Buxton i Arnott, 1992). Cada unitat d'*Hydra* consisteix en una càmera incrustada, un altaveu i un monitor de vídeo que permet que cada participant remot ocupi un espai físic únic. Asseguda en un extrem de la taula hi ha una de les investigadores que va desenvolupar el projecte, Abi Sellen.

Amb la càmera incrustada, cada unitat d'*Hydra* facilita que cada participant remot que hi està associat tingui una visió única d'Abi. Com que cadascun dels participants remots té un sistema *Hydra*, Abi també té una perspectiva única de cada participant. Així, es poden conservar les relacions espacials de la «taula rodona» virtual, il·lustrades a la figura de la dreta. Abi veu les persones *A*, *B* i *C* en les unitats *Hydra* de la seva esquerra, la seva dreta i davant seu, respectivament. De la mateixa manera, la persona *A* la veu a la dreta, la *B* a l'esquerra, etc.

Malgrat que Abi té tres monitors, càmeres i altaveus a la taula, la totalitat del material ocupa menys espai que el telèfon i li deixa lloc per als documents o altres materials rellevants de la reunió.

Aquestes unitats *Hydra* són un bon exemple de la transparència de l'omnipresència, perquè cadascuna té una font de veu diferent per a cada participant remot. Com a resultat, es crea l'entorn que permet de mantenir conversacions en paral·lel. Així es va demostrar en un estudi formal que comparava diverses tecnologies per a reunions multipersonals (Sellen, 1992). Les unitats *Hydra* van ser l'única tecnologia analitzada que permetia d'establir el paral·lelisme amb les converses durant les reunions personals.

Encara que no es va aplicar, les unitats també podien incorporar sensors de proximitat que haurien permès de realitzar comentaris a banda de la mateixa manera que en les reunions personals: acostant-nos cap a la persona a qui es vol fer el comentari a banda. Gràcies a la percepció de mirades⁴ d'aquestes unitats, es conserven les comprovacions i les conclusions regulars de les reunions personals, ja que tots els participants poden veure que es fa un comentari a banda, entre quines persones en concret i durant quanta estona.

Cap d'aquests actes de parla quotidians no és possible amb els dissenys convencionals, tot i que amb aquesta proposta, no calen noves habilitats substancials. Una vegada més, no cal aprendre cap «interfície d'usuari». Es pot interaccionar amb els substituïts de vídeo amb les mateixes habilitats socials, o convencions, que faríem servir en una situació personal.

Concepte: substituït de vídeo. No pensis en la càmera com a càmera. Pensa que és el substituït de l'ull. De la mateixa manera, no pensis en l'altaveu com a altaveu. Pensa que és el substituït d'una boca. Integrat en una única unitat, un vehicle que aconsegueix els principis de disseny 1 i 2.

4. La percepció de mirades és un aspecte important de la interacció humana. És tan rellevant que alguns (Russ, 1925) han argüït que els ulls emeten raigs actius, o «raigs oculars» que es poden notar, com quan algú ens mira des de darrere.

L'escala i l'indret defineixen l'espai d'interacció

En tot l'article hem recalcat la importància de les consideracions espacials a l'hora d'emprar la tecnologia i dissenyar-la. Gairebé tots els nostres exemples han tractat només amb l'indret, però com a component de l'espai, l'escala pot ser tan important com l'indret en termes de la capacitat d'afectar la qualitat de la interacció.

Considerem l'impacte de seure de manera electrònica davant per davant en un escriptori, de la manera com il·lustra la figura 1, en comparació amb la figura 10. A la figura 10, per mitjà de la projecció posterior, el participant remot adquireix la mida real quan apareix a l'altra banda de l'escriptori. Hi ha tot un nombre significatiu de punts que deriven d'aquest exemple.

Primerament, no és com mirar la televisió. Per l'escala de la imatge, el contorn de la pantalla és fora del contorn principal de la visió. La persona remota és definida pel perímetre de la silueta i no pel bisell del monitor.



FIGURA 10. Situació presencial. En aquest escenari, cada participant té una sobretaula informatitzada on apareix la mateixa informació. La intenció és captar l'essència de treballar davant per davant en una taula. Cadascú veu el participant remot de la mida real. La càmera de vídeo (d'una unitat Hydra) es manifesta de manera no intrusiva a l'escriptori. Els participants interaccionen amb l'ordinador amb un punter. Quan un dels participants mira cap a l'escriptori, sembla que els ulls es projectin en l'espai de l'altre i, així, s'incrementa la sensació de telepresència. Encara que implica l'ús de molta tecnologia, s'integra en l'ecologia arquitectònica. S'aconsegueixen molts serveis i molt espai i no pas aparells o estris.

Fotografia: Projecte de Telepresència d'Ontario.

En segon lloc, pel fet de tenir la mida real, s'estableix un equilibri entre el pes o el poder que exerceix cada participant.

Tercer, i potser el més important, la percepció del participant remot pot arribar en el nostre espai físic. Quan la part remota mira cap a l'escriptori, la nostra sensació de percepció de la mirada (vegeu també Ishii, Kobayashi i Grudin, 1992) ens fa sentir que miren la nostra taula. La seva mirada traspasa la distància cap al nostre lloc de treball compartit, amb la qual cosa s'enforteix la impressió de telepresència.

Resulta cabdal en aquest exemple el contrast entre la simplicitat i la naturalitat de l'ambient i la potència de la funcionalitat. Pel fet que es manté el principi de la invisibilitat, es crea una situació de treball potent i no intrusiva.

Principi de disseny 4: el continent on dissenyem les solucions és la sala en què treballem/representem/aprenem, no en una capsa que puguem tenir damunt l'escriptori. Aquesta és la diferència entre el disseny ecològic de l'espai mediàtic omnipresent i el disseny dels aparells.

Espai mediàtic omnipresent, sensació de proximitat i entorn reactiu

Hem començat amb una presentació breu de la informàtica omnipresent. Després hem vist com alguns dels mateixos conceptes es poden aplicar a la interacció mediada per vídeo. Per recalcar-ne el paral·lelisme, ho hem anomenat *vídeo omnipresent*. Malgrat tot, no es tracta de dos universos paral·lels i la distinció entre tots dos s'anirà desdibuixant. En primer lloc, els senyals d'àudio i de vídeo emprats en aquest tipus de conferències serà digital.⁵ En segon lloc, fins i tot en els sistemes de videoconferència, els ordinadors formen part de l'equació, per exemple quan s'estableixen les connexions, com hem il·lustrat a la figura 5.

En aquesta secció, argumentarem que els grans avantatges apareixen quan pensem en tots dos d'una manera integrada. En comptes del vídeo o de la informàtica omnipresent, la nostra aportació és que el terme *espai mediàtic omnipresent* és més apropiat.

Ja hem vist dos exemples de com podem treure profit de la integració de les tecnologies informàtiques i de telecomunicacions. Per una banda ens hem referit als sistemes Hydra, on hem argüït la integració de sensors de proximitat per tal de mediar tant si hi havia públic o si es tractava d'una conversa privada entre dos individus. Aquí veurem un exemple d'entorn que reacciona a cada gest quotidià (en aquest cas del llenguatge corporal que significa inclinar-se cap a algú) i que, de manera transparent, canvia la modalitat del sistema adequadament.

5. En la majoria de prototipus vam fingir que ho era per mitjà d'una tecnologia anàloga, ja que el nostre centre d'interès era l'ús més que no pas els detalls d'enginyeria de la infraestructura subjacent. El nostre punt de vista és que si el model d'ús era correcte, aquesta informació arribaria a l'enginyeria. Sembla que aquest és l'ordre correcte per aconseguir els objectius: primer les persones, l'enginyeria de telecomunicacions després.

Un altre exemple és el que hem il·lustrat per mitjà de la figura 6, on l'ordinador nota la posició de la porta física. D'aquesta manera es trenca amb les pràctiques informàtiques convencionals, però això és totalment coherent amb el món quotidià. La integració apropiada de l'ordinador i el vídeo omple el buit entre la interacció persona-persona i persona-ordinador.

Observació: la porta és un element d'entrada tan legítim cap a l'ordinador com el ratolí o el teclat.

L'habilitat necessària perquè els ordinadors siguin més «conscients» de l'entorn, com mostren aquests dos exemples, és una part important del nostre treball. Comparem la nostra proposta amb el que passa en el reconeixement remot, on els sensors dels satèl·lits recullen informació sobre l'ecologia de la Terra. Com que es tracta de la mateixa estratègia, només una mica més propera, descrivim els nostres plantejaments (com en els darrers dos exemples) com a «reconeixement pròxim». També fem servir el terme *entorn reactiu* (Buxton, 1995; Cooperstock, Fels, Buxton i Smith, 1997).

La noció que les coses que hi ha al nostre entorn poden tenir reaccions vers nosaltres no és nova. Només cal pensar en la porta del supermercat que, tenint en compte que potser vas carregat de bosses, s'obre de manera automàtica quan t'acostes a la sortida. És tan habitual que ja ho tenim assumit. Però pensem-hi un moment. Els ordinadors són fets de milers d'interruptors, malgrat l'IA, l'interruptor de reconeixement de moviment de la porta del supermercat és força més espavilat que qualsevol dels interruptors de l'ordinador personal.

Quan ens acostem a l'ordinador, s'atura el protector de pantalla i apareixen les finestres obertes? Sap, almenys, que som al davant? Seria gaire difícil canviar aquesta realitat? Com podem fer coincidir la propaganda entre la intel·ligència dels ordinadors i la celeritat amb què avancem, quan l'electrònica d'una porta de supermercat és de manera significativa més avançada que el nostre ordinador?

Allò que permet que algunes coses, com els llums que perceben moviment, les portes dels supermercats, el nostre ratolí de porta i les unitats Hydra, responguin de manera útil als esdeveniments que es produeixen a prop seu és que es tracta d'aparells especialitzats creats per uns dissenyadors que eren tan conscients d'allò que no havien de fer els aparells com d'allò que havien de fer. Perquè són especialistes, els dissenyadors poden formular assumpcions sobre on seran utilitzats i com, i sobre les necessitats i tipus d'activitats associades. Els ordinadors personals convencionals no poden recrear aquestes accions perquè són aparells destinats a propòsits generals.

Si es fan servir i es dissenyen de manera apropiada, l'espai mediàtic omnipresent es converteix en una alternativa valuosa a la situació actual.

Cal reconèixer que quan combinem les tecnologies del vídeo i les informàtiques, ens trobem amb les tecnologies que permeten el reconeixement pròxim de manera gairebé gratuïta. Les mateixes càmeres que fem servir per a les videoconferències es poden convertir en els «ulls» del meu ordinador. El mateix micròfon amb el qual parlo als meus col·legues pot també esdevenir una «orella» per a l'ordinador. Les pantalles per on veig el vídeo també poden visualitzar dades i viceversa: quan el món és digital, l'àudio, el vídeo i les dades són una mateixa cosa.

Principi de disseny 5: tots el aparells que es fan servir en la interacció humana (càmeres, micròfons, etc.) són candidats legítims per a la interacció humà-ordinador (i sovint simultàniament).

Tornem a la qüestió de si el meu ordinador surt del protector de pantalla quan m'hi acostot. Amb l'espai mediàtic omnipresent, la càmera de sobretaula pot detectar si sóc a l'escriptori. Si no hi sóc, però la càmera de la porta nota que sóc a la cambra, aleshores l'ordinador pot canviar de la sortida visual a l'auditiva per comunicar-se amb mi. Al mateix temps, com que analitza l'entrada cap al micròfon (per mitjà d'un senyal senzill de detecció), sap si enrao no o no. Si parlo, esperarà fins que acabi i així no m'interromprà.

Aquest repertori estès de tecnologies pot servir de base per a una interfície encara més perfecta entre els móns físics i els electrònics. Krueger (1983, 1991) ha demostrat com les càmeres de vídeo poden ser aparells d'entrada efectius a l'hora de controlar els sistemes informàtics. La figura 11 il·lustra un exemple del seu treball.

Resulta un punt central del sistema (a l'inrevés del que acostuma a passar en els sistemes de realitat virtual) el fet que no provoca cap intromissió i, a diferència de la majoria de les interaccions informàtiques, té lloc al fons (Buxton, 1995). No cal dur guants especials ni sensors. El sistema veu i entén els gests amb les mans de la mateixa manera que les persones: mirant les mans.

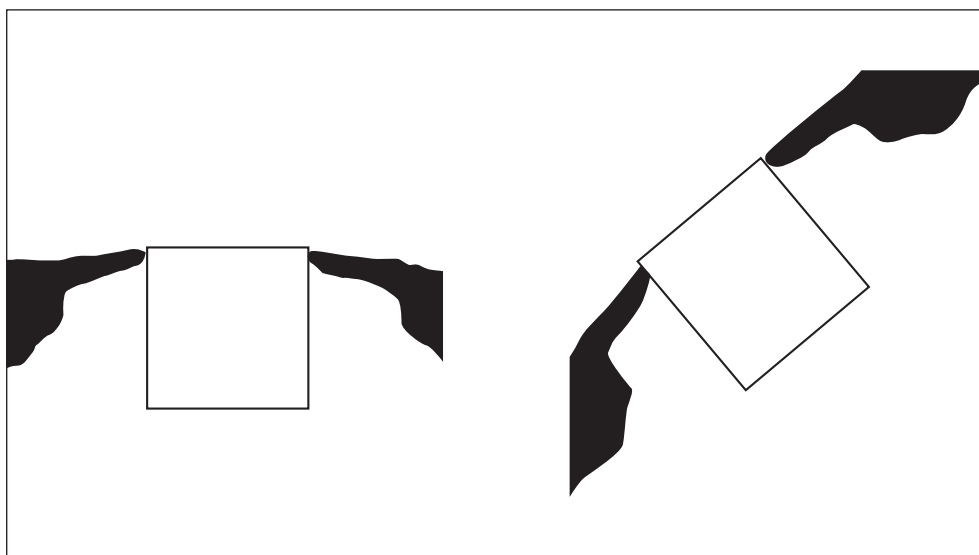


FIGURA 11. El vídeo d'escriptori de Myron Krueger. Les mans de l'usuari poden ser «vistes» per l'ordinador i queden sobreposades a la pantalla. El sistema reconeix cada mà, la posició i la forma (oberta, tancada, si indica algun lloc, etc.). D'aquesta manera, podem manipular els objectes de l'escena. En aquest cas, l'usuari reorienta un quadrat.

Aquestes tecnologies no faciliten només la relació entre l'humà i la màquina. Així també es pot accedir a una connexió més perfecta entre els aparells dels mons físic i electrònic. A mesura que les tecnologies són més íntimes, o pròximes a la persona, cada vegada tindran més capacitat d'esdevenir un pont entre aquests dos mons. Alguns dels ordinadors de la mida d'un tabulador ara s'assemblen més a una càmera que a una calculadora, per exemple (vegeu la figura 12, que pot servir com a exemple).



FIGURA 12. El visor PDA Handspring amb la càmera com a mòdul d'ull. Aquest aparell d'ús manual esborra la distinció entre aparell informàtic, assistent personal digitalitzat (PDA) i càmera digital. <<http://www.handspring.com>>

Un dels millors exemples a l'hora d'emprar aquests mitjans de manera combinada i que permeten d'establir aquesta mena de pont és l'escriptori digital de Wellner (1991), tal com l'hem il·lustrat a la figura 13. Aquest sistema va més enllà tant dels ordinadors de sobretaula com de la metàfora que implica la paraula *sobretaula*. En aquest cas, l'escriptori és l'ordinador.

Tal com demostra la figura, hi ha un projector i una càmera muntats al damunt de cada escriptori. El primer projecta la visualització de l'ordinador cap a l'escriptori. La càmera permet que l'ordinador «vegi» què hi ha damunt de la taula. Així, els documents electrònics poden ser projectats, com altres estris actius, com ara la calculadora o un navegador. I, com en l'exemple de Krueger, la càmera permet que l'ordinador vegi les accions de les mans damunt la taula i pugui utilitzar-les com a dades d'entrada. També permet que l'ordinador pugui «veure» els documents i els objectes de l'escriptori. Hi torna a haver un gran potencial de reconeixement. En el prototipus de treball, per exemple, la càmera es pot fer servir per escanejar dades alfanú-

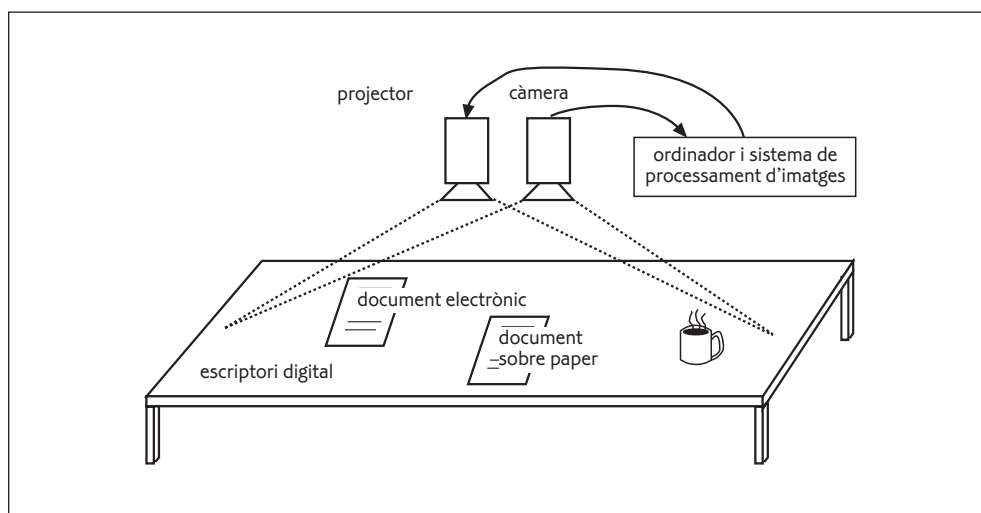


FIGURA 13. L'escriptori digital (Wellner, 1992). Amb aquest sistema, els documents digitals es projecten damunt la taula. De manera similar, una càmera permet que l'ordinador vegi què hi ha damunt de la taula. Pot veure els documents i «llegir-los» per mitjà de tècniques de reconeixement de caràcters òptics (OCR). També pot «veure» les mans de l'usuari i reconèixer gests, com ara quan indica cap algun lloc, selecciona i activa els botons gràfics dels aparells que es projecten damunt de la superfície de l'escriptori.

meriques on s'apliquen tècniques de reconeixement de caràcters, de manera que l'ordinador pot «llegir» què hi ha al damunt de la taula.

Resum i conclusions

Hem topat amb la barrera de la complexitat. Per mitjà de les tècniques de disseny, no podem augmentar de manera significativa la funcionalitat dels sistemes sense traspasar el llinard de la frustració dels usuaris. Més que afegir complexitat, la tecnologia hauria de reduir-la i incrementar la nostra capacitat de funcionar en el món emergent del futur.

L'espai mediàtic omnipresent de disseny representa un trencament de les pràctiques anteriors, un canvi en el disseny que es basa en les habilitats adquirides dels usuaris, que no demana que se n'hagin d'aprendre de noves. És un enfocament madur de disseny que trenca amb la mentalitat d'un únic aparell al qual s'han aplicat tècniques de «contenedor-ficat-en-una-capsa», que és el dominant en les tendències actuals. Com la bona arquitectura i el disseny interior, és confortable, no intruïssiu i funcional.

Per poder recollir els beneficis que ofereix aquest projecte, caldrà repensar com definim, ensenyem i practiquem la ciència. Si seguim el camí assenyalat més amunt, el centre de la nos-

tra recerca actual ha de ser aplicar les nostres habilitats a la tecnologia i les ciències socials per fer més acurada la nostra comprensió del disseny i establir la validesa en els termes que són més importants: els humans.

Agraïments

La feina i les idees desenvolupades en aquest article són fruit d'un nombre incomptable de discussions amb els meus col·legues, i de les seves contribucions, del Projecte de Telepresència d'Ontario, Xerox PARC i Rank Xerox EuroPARC. Estic molt agraït a tots els que han fet que aquests entorns hagin estat tan estimulants. Vull agrair especialment les contribucions d'Abi Sellen, Sara Blym, Steve Harrison, Mark Weiser, Tom Moran, Marilyn Mantei, Brigitta Jordan i Bill Gaver.

La recerca presentada en aquest article ha rebut el suport del Projecte de Telepresència d'Ontario, Xerox PARC i del Consell de Recerca sobre les Ciències Naturals i l'Enginyeria del Canadà. Els agraeixo el suport rebut.

Referències bibliogràfiques

- BLY, S.; HARRISON, S.; IRWIN, S. (1993). «Media Spaces: bringing people together in a video, audio and computing environment». *Communications of the ACM*, núm. 36 (1), p. 28-47.
- BUXTON, W. (1995). «Integrating the Periphery and Context: A New Model of Telematics». *Actes de Graphics Interface '95*, p. 239-246.
- COOPERSTOCK, J.; FELS, S.; BUXTON, W.; SMITH, K. C. (1997). «Reactive environments: Throwing away your keyboard and mouse». *Communications of the Association of Computing Machinery (CACM)*, núm. 40 (9), p. 65-73.
- ELROD, S.; HALL, G.; COSTANZA, R.; DIXON, M.; DES RIVIERES, J. (1993). «Responsive office environments». *Communications of the ACM*, núm. 36 (7), p. 84-85.
- FIELDS, C. I. (1983). «Virtual space teleconference system». *United States Patent 4, 400, 724* (23 agost 1983).
- GAVER, W.; MORAN, T.; MACLEAN, A.; LÖVSTRAND, L.; DOURISH, P.; CARTER, K.; BUXTON, W. (1992). «Realizing a video environment: EuroPARC's RAVE System». *Actes de CHI 1992*, p. 27-35.
- ISHII, H.; KOBAYASHI, M.; GRUDIN, J. (1992). «Integration of inter-personal space and shared workspace: Clarboard design and experiments». *Actes de CSCW 1992*, p. 33-42.
- KRUEGER, MYRON, W. (1983). *Artificial Reality*. Reading: Addison-Wesley.
- (1991). *Artificial Reality II*. Reading: Addison-Wesley.
- MANTEI, M.; BAECKER, R.; SELLEN, A.; BUXTON, W.; MILLIGAN, T.; WELLEMAN, B. (1991). «Experiences in the use of a media space». *Actes de CHI 1991, ACM Conference on Human Factors in Software*, p. 203-208.

- RUSS, Charles (1925). «An instrument which is set in motion by vision». *Discovery, Series 1*, vol. 6, p. 123-126.
- SELLEN, A. (1992). «Speech patterns in video mediated conferences». *Actes de CHI 1992, ACM Conference on Human Factors in Software*, p. 49-59.
- SELLEN, A.; BUXTON, W.; ARNOTT, J. (1992). «Using spatial cues to improve videoconferencing». *Actes de CHI 1992*, p. 651-652. [També enregistrat per vídeo a les *Actes de CHI 1992*]
- STULTS, R. (1986). «Media Space». *Systems Concepts Lab Technical Report*. Palo Alto, CA: Xerox PARC.
- WEISER, M. (1991). «The computer for the 21st century». *Scientific American*, núm. 265 (3), p. 94-104.
- WELLNER, P. (1991). «The DigitalDesk Calculator: Tactile manipulation on a desktop display». *Proceedings of the Fourth Annual Symposium on User Interface Software and Technology (UIST '91)*, p. 27-33.
- WELLNER, P.; MACKAY, W.; GOLD, R. [ed.] (1993). «Computer-Augmented Environments: Back to the real world». *Communications of the ACM*, núm. 36 (7). [Número especial]