

Vers des systèmes de communication multi-échelles

Nicolas Roussel

Projet In Situ, LRI & INRIA Futurs
Bâtiment 490, Université Paris-Sud
91405 Orsay Cedex, France
roussel@lri.fr

RESUME

Quarante ans après le PicturePhone d'AT&T, la plupart des systèmes de communication vidéo sont toujours conçus comme des téléphones augmentés visant à reproduire la communication face à face. Dans cet article, je m'intéresse à ce but implicite et rarement remis en cause. Je propose une nouvelle approche de la communication médiatisée que je qualifie de *multi-échelles*. L'article présente cette approche, trois exemples de systèmes qui l'illustrent et quelques perspectives qu'elle laisse entrevoir.

MOTS CLES : Vidéo, communication médiatisée, communication multi-échelles, coordination, communication, collaboration.

ABSTRACT

Forty years after AT&T's PicturePhone, most video communication systems are still designed as augmented telephones that aim at reproducing face-to-face communication. In this paper, I discuss this implicit and hardly questioned goal. I propose the concept of multiscale communication systems as an alternative approach to computer-mediated communication. The paper presents the multiscale approach, three systems that illustrate it and some directions for future work.

CATEGORIES AND SUBJECT DESCRIPTORS: H.4.3 [Communications Applications]: Computer conferencing, teleconferencing, and videoconferencing.

GENERAL TERMS: Design, Human Factors.

KEYWORDS: Video-mediated communication, computer-mediated communication, multi-scale communication, coordination, communication, collaboration.

INTRODUCTION

Quarante ans après le PicturePhone d'AT&T [18], les systèmes de communication vidéo sont toujours conçus comme des téléphones augmentés visant à reproduire la communication face à face et présentés comme l'outil de communication idéal. Contrairement aux prédictions de nombreux futurologues du siècle passé, la visioconférence n'a pas remplacé les voyages d'affaires, le nombre de ceux-ci ayant même toujours tendance à augmenter [16]. Et bien que des outils de visioconférence soient aujourd'hui disponibles gratuitement sur les trois principales plateformes logicielles (Microsoft Windows, Linux et Apple Mac OS X), leur usage reste très limité. L'échange de texte (e.g. courrier électronique, messagerie instantanée, mini-messages) et la communication parlée restent les moyens les plus utilisés pour communiquer à distance ou de manière asynchrone.

De nombreux chercheurs en Interaction Homme-Machine (IHM) ont tenté d'expliquer les raisons de cet échec de la visioconférence et ont proposé de nouveaux usages de la vidéo pour la communication médiatisée (e.g. [8, 20, 15]). Ces travaux ont en quelque sorte culminé en 1997 avec la publication du livre *Video-mediated communication* [9]. La plupart des systèmes développés à cette époque étaient basés sur des liaisons audio-vidéo analogiques. Mais alors même que la vidéo numérique et les réseaux rapides envahissaient notre quotidien, la fin des années 1990 marqua le déclin de l'intérêt de la communauté IHM pour la communication vidéo. Comme l'avait prophétisé Karam [23], l'avènement des autoroutes de l'information fut fatal aux divers projets existants : les réseaux analogiques furent peu à peu abandonnés et les travaux associés, faute d'être adaptés aux nouveaux réseaux, furent oubliés ou glorifiés en l'état.

La réalisation d'un système de communication vidéo numérique pose de nombreux problèmes liés à la capture, au codage, à la transmission et à la synchronisation des flux. Malgré les progrès réalisés en matière de réseaux et de gestion de données multimédia, la mise en œuvre de ces systèmes reste une affaire de spécialistes et leur évolution reste de ce fait guidée par une vision technocentrique. Ainsi, bien que le nouveau codec H.264 soit

décrit comme permettant une qualité d'image "ultra-efficace et sans précédent" [1], son usage pour la communication entre individus se limite encore à la réalisation de nouveaux visiophones.

Les slogans publicitaires modernes de l'industrie téléphonique font écho à la foi inébranlable de la direction et des chercheurs d'AT&T des années 1960 pour qui l'ajout de l'image au téléphone était inévitable et nécessairement bénéfique (e.g. "C'est mieux de se voir quand on se parle", "Les appels visio ? C'est si simple de partager encore plus d'émotions"). Peu de monde s'interroge aujourd'hui sur le but de la recherche et du développement en matière de communication vidéo. Pourtant, comme le fait remarquer Lipartito [18], outre les raisons de l'échec du PicturePhone, on peut s'interroger sur celles de son invention : quel est le but fondamental recherché par ce type de technologie ?

À travers un article de 1992 intitulé *Beyond being there* [14], Hollan et Stornetta ont une première fois attiré l'attention de la communauté IHM sur le but implicite et rarement remis en cause de la plupart des travaux sur la communication médiatisée. Dans cet article, je m'intéresse également aux objectifs de ce type de recherche. Dans la section suivante, je reviens sur l'analyse de Hollan et Stornetta et sur les deux grandes interprétations que l'on peut en faire. Je propose ensuite de considérer la communication médiatisée selon une nouvelle approche, que je qualifie de multi-échelles et que j'illustre par la description de trois systèmes de communication vidéo. Je conclus enfin par une discussion sur les perspectives que cette nouvelle approche laisse entrevoir.

BEYOND BEING THERE

"Être là-bas" (*being there*) est littéralement impossible. Cette expression fait référence au concept de *présence* que Lombard et Ditton définissent comme "l'illusion perceptuelle de non-médiation" [19]. Elle symbolise également le but implicite généralement visé par les concepteurs de systèmes de communication vidéo : "atteindre la même richesse d'information que dans l'interaction en face à face" pour pouvoir "interagir avec ceux qui sont éloignés comme nous le faisons avec ceux qui sont proches" [14].

Le sentiment de présence, au sens de Lombard et Ditton, varie selon les moyens de communication utilisés. Des théories comme celles de la *présence sociale* [28] ou de la *richesse des médias* [6] ont été élaborées pour caractériser les différents médias, les comparer et déterminer celui ou ceux qui maximisent l'efficacité ou la satisfaction pour une tâche particulière. La plupart des travaux basés sur ces théories considèrent que le sentiment de présence augmente en fonction de la richesse du média utilisé [7]. Ainsi, la capacité de l'image à transmettre des informations non-verbales (e.g. postures, gestes, expres-

sions, contact visuel) est fréquemment citée pour expliquer l'accroissement du sentiment de présence provoqué par l'ajout d'une liaison vidéo à une liaison audio.

Dans leur article, Hollan et Stornetta remettent en question le but fondamental des travaux sur la communication médiatisée. Selon eux, ce but devrait être la conception d'outils qui aillent au-delà de la simple imitation des possibilités offertes par la présence physique. Cette idée, reprise par de nombreux autres chercheurs, conduit à deux types de travaux correspondant globalement aux deux sens de l'expression "au-delà" : "plus loin" et "autrement".

Au-delà, plus loin

Hollan et Stornetta posent la question suivante : "que se passerait-il si nous développions des outils de communication permettant une plus grande richesse d'information que le face à face ?". Dans la conclusion de leur article, ils ajoutent : "nous devons développer des outils que les gens préféreront utiliser même dans les cas où la proximité physique est possible". Certains voient ici une invitation à poursuivre plus avant l'approche technocentrique qui a guidé le développement des systèmes de communication et abouti aux outils de visioconférence actuels. Pour eux, si ces systèmes ne sont pas utilisés, c'est simplement qu'ils ne sont pas encore assez bons.

Selon cette approche, l'ajout de technologie est le principal moyen d'augmenter la valeur du système. L'image et le son sont considérés comme une base minimale. L'amélioration de leur qualité et l'ajout d'autres modalités (e.g. toucher, odorat, goût) sont généralement proposés pour permettre de mieux percevoir et agir à distance, pour rendre la communication "plus naturelle", "plus directe" et donc "plus réelle". Ces dernières années, on a ainsi vu se multiplier les travaux sur des systèmes permettant des prises de vue multiples, la reconstruction 3D des participants et/ou de leur environnement et un affichage de type immersif (e.g. [11, 13, 16]).

Mais en essayant de faire "mieux que la présence physique", de la dépasser, les concepteurs de ces systèmes ne font que déplacer subtilement le but initial dénoncé par Hollan et Stornetta sans le remettre en cause. Mieux encore, ils se placent dans une situation de course sans fin, la loi de Moore leur permettant d'annoncer régulièrement que la nouvelle version de leur système est plus performante que la version précédente (et donc plus naturelle, plus directe, plus intuitive...) mais bien moins que la version suivante qu'ils imaginent déjà.

Au-delà, autrement

La communication face à face qui sert de référence aux systèmes de communication vidéo traditionnels est bien adaptée aux situations de travail formelles, avec un but précis. Mais l'activité d'un groupe de personnes ne se limite pas à ces situations formelles et à des actes de

communication explicites. Dans le cas de groupes géographiquement dispersés, il est particulièrement important de ne pas négliger le besoin d'interactions informelles et spontanées. Dans ce contexte, l'analyse de Hollan et Stornetta peut être vue comme une incitation à concevoir des systèmes permettant des formes de communication différentes, plus informelles, plus légères ou plus intimes, par exemple, non nécessairement basées sur le son et l'image.

Différents projets autour des *mediaspaces* [20] ont ainsi proposé de nouveaux services de communication audio-vidéo pouvant avantageusement compléter la visioconférence traditionnelle dans le cadre d'une utilisation quotidienne (e.g. coups d'œil de quelques secondes, liaisons permanentes, vues d'ensemble montrant plusieurs bureaux). Ces travaux ont démontré qu'une utilisation périphérique de ces services contribuait à l'émergence d'une conscience de groupe, laquelle renforçait le sentiment de présence.

D'autres travaux ont par la suite visé à rendre la communication moins intrusive, plus implicite, plus personnelle. Strong et Gaver, par exemple, ont proposé trois dispositifs minimalistes utilisant la vue, l'odorat et le toucher pour exprimer un sentiment ou une émotion [29]. De nombreux autres dispositifs tangibles et/ou ambiants ont suivi, tirant souvent partie d'un contexte particulier, comme la relation intime pouvant exister entre deux personnes, pour simplifier à l'extrême la communication¹ (e.g. [2, 3, 30]).

VERS DES SYSTEMES DE COMMUNICATION MULTIECHELLES

L'évolution des systèmes de communication tend à les rendre plus divers, plus nombreux et, malheureusement, souvent plus complexes. L'idée que ces systèmes ne doivent pas se contenter de reproduire les caractéristiques de la présence physique est aujourd'hui admise par tous. Mais cette idée ne suffit pas à structurer la recherche et le développement des nouveaux systèmes. Si elle remet en cause le but implicite sur lequel les travaux antérieurs s'étaient fondés, elle n'en propose pas vraiment de nouveau.

La diversité et le nombre de services disponibles rendent le choix du mode de communication de plus en plus difficile. Le degré d'implication souhaité semble toutefois être un critère essentiel pour ce choix : en fonction du contexte particulier dans lequel elle se trouve et de la place qu'elle veut accorder à la communication distante

¹ L'idée que la prise en compte du contexte permet de réduire le message à transmettre tout en préservant le sens général et l'intention n'est pas récente. Ainsi en 1862, selon [21], Victor Hugo en Angleterre et voulant savoir comment son livre *Les Misérables* avait été accueilli en France, envoya un télégramme contenant un simple point d'interrogation à son éditeur qui lui répondit par un point d'exclamation.

dans son activité présente, une personne sera plus ou moins disposée à interagir avec le ou les correspondant(s).

Les dispositifs et logiciels de communication à notre disposition combinent souvent des services correspondant à différents degrés d'implication. Les téléphones portables modernes, par exemple, combinent des services d'identification de l'appelant, d'échange de messages textuels, de messagerie vocale, de téléphonie et éventuellement de visiofonie (Figure 1). Cet éventail de services permet de contrebalancer l'accessibilité permanente liée à la possession du téléphone. Cependant, les transitions entre les services sont généralement difficiles, voire impossibles². Les transitions entre dispositifs ou logiciels sont encore plus difficiles.

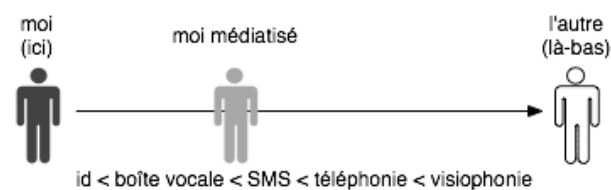


Figure 1 : Degré d'implication des services offerts par un téléphone portable.

La notion de degré d'implication souhaité peut être raffinée en différenciant par exemple le degré d'engagement (i.e. je suis prêt à m'exposer) de celui d'attention (i.e. je suis prêt à être dérangé). Ces notions ont été en partie abordées par les travaux sur les mediaspaces (e.g. [12, 10]). Toutefois, les transitions entre différents niveaux ont été jusqu'ici plutôt négligées. La fluidité de ces transitions est pourtant essentielle pour pouvoir alterner rapidement entre des activités de *coordination*, de *communication* ou de *collaboration* (ou *production*, si l'on se réfère au modèle du trèfle de [24]).

Au lieu de multiplier les services de communication indépendamment les uns des autres - qu'ils tentent d'imiter la présence physique ou non - il me semble que nous devrions concevoir des systèmes capables de proposer un niveau d'implication variable et des transitions fluides entre les niveaux. Un tel système permettrait aux utilisateurs de choisir simplement et rapidement le niveau d'implication le plus adapté au contexte particulier dans lequel ils se trouvent.

D'une certaine manière, ces notions de degré d'implication et de transitions fluides sont à rapprocher de celles de *zoom sémantique* et de *zoom continu* proposées par Perlin et Fox pour la navigation dans des mon-

² Lorsque l'on vient de manquer un appel faute d'avoir pu décrocher avant sa messagerie vocale, par exemple, l'expérience montre qu'il vaut mieux attendre avant de rappeler le correspondant si l'on ne veut pas que sa messagerie décroche pendant qu'il parle à la nôtre.

des d'information [22]. Pour cette raison, je qualifie cette nouvelle approche de la communication médiatisée de *multi-échelles*. Il me semble par ailleurs que d'autres notions utilisées dans le domaine de la Visualisation d'Information pourraient être adaptées à la communication entre individus. Ainsi, le fameux mantra de Shneiderman pour la recherche d'information (d'abord une vue d'ensemble puis du zoom et du filtrage et enfin des détails à la demande [27]) me semble particulièrement pertinent pour décrire l'engagement progressif dans une communication avec une ou plusieurs autres personnes.

Dans la suite de cette section, je présente trois systèmes de communication vidéo qui illustrent de manières différentes les principes de niveaux d'implication variables et de transitions fluides entre ces niveaux. Les détails concernant ces systèmes ayant été publiés par ailleurs, leur description se limitera aux aspects relatifs à l'approche multi-échelles.

Exemple 1 : VideoServer

VideoServer [24] est un serveur HTTP individuel qui permet à un utilisateur de rendre des images préenregistrées ou capturées en temps réel accessibles à d'autres personnes à travers de simples URLs³. Ce serveur a été développé dans le but de faciliter la création de mediaspaces numériques que les utilisateurs puissent aisément adapter à leurs besoins. VideoServer est capable de transmettre les flux vidéo sur la connexion HTTP elle-même. Les images peuvent ainsi être affichées par un moteur de rendu HTML en lieu et place d'une image traditionnelle⁴, dans un courrier électronique (Figure 2, gauche) ou dans une page Web (Figure 2, droite).

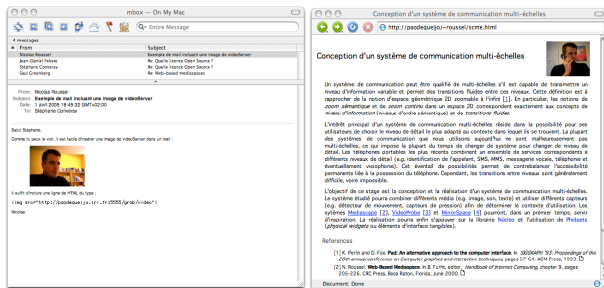


Figure 2 : Images capturées en temps réel par VideoServer et affichées dans l'application Mail d'Apple et le navigateur Web Camino.

Plusieurs services vidéo peuvent également être combinés en utilisant les URLs correspondantes dans un même document. Quelques lignes de JavaScript permettent de rafraîchir régulièrement une image fixe, ou de la transformer en un flux vidéo lorsque la souris la survole et d'ouvrir une nouvelle fenêtre montrant un autre flux lorsqu'on clique dessus. On peut ainsi passer simplement

³ e.g. <http://machine/push/video>

⁴ e.g. ``

d'une petite image fixe (e.g. 80x60 pixels) mise à jour toutes les 5 minutes à un flux vidéo de mêmes dimensions à la même place, puis à un flux plus rapide affiché dans une fenêtre que l'utilisateur peut librement déplacer et redimensionner, l'image affichée s'adaptant automatiquement à la place disponible (Figure 3).

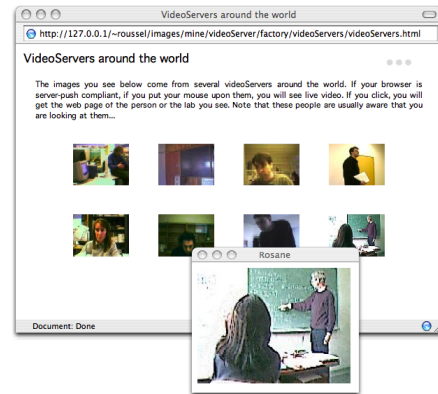


Figure 3 : De la vue d'ensemble aux détails à la demande, un exemple de mise en œuvre d'attention sélective.

Comme la plupart des mediaspaces, et à l'inverse des nombreux serveurs de webcam, VideoServer offre à son utilisateur des moyens de notification et de contrôle qui lui permettent de définir simplement différentes politiques d'accès. En fonction du contexte (e.g. identité de la personne distante, heure de la journée, document dans lequel doivent s'insérer les images), l'image capturée en temps réel peut-être envoyée telle quelle ou filtrée, mais peut aussi être remplacée par une image préenregistrée (Figure 4). À l'attention sélective offerte par la combinaison de services s'ajoute donc la notion d'engagement sélectif.

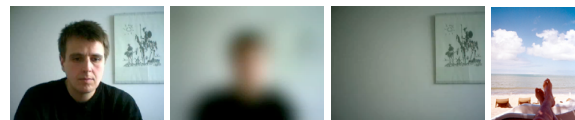


Figure 4 : Exemples de mise en œuvre d'engagement sélectif : image capturée, image filtrée, image préenregistrée ambiguë et indicateur explicite d'absence prolongée.

VideoServer permet l'intégration de services de communication vidéo synchrones au Web et au courrier électronique, deux moyens de communication asynchrones. Comme nous l'avons vu, les services vidéo peuvent être configurés de façon à couvrir un large spectre allant d'une vue d'ensemble d'un groupe de personnes mise à jour périodiquement à une vue plein écran et temps réel d'une personne particulière. Les transitions entre différents niveaux d'implications peuvent être déclenchées par de simples actions à la souris sur les images elles-mêmes et les fenêtres qui les contiennent. Pour toutes ces raisons, VideoServer peut être considéré comme un premier exemple de système de communication multi-échelles.

Exemple 2 : VideoProbe

VideoProbe [4] est un système de communication vidéo permettant à un groupe de personnes de partager leur vie quotidienne par échanges de photos. Lorsque la scène filmée par le système ne change pas, son écran est entièrement blanc (Figure 5, image 1). Lorsqu'un changement est détecté, l'écran affiche progressivement les images capturées, à la manière d'un miroir (images 2 et 3). Si le changement dans la scène persiste plus de trois secondes, une photo est automatiquement transmise aux autres systèmes. Un carré translucide recouvrant progressivement les images indique le temps restant avant la prise de vue (images 4 et 5). Lorsqu'une photo est transmise aux autres membres du groupe, elle est affichée pendant trois secondes correctement orientée et plus grande que les images précédentes (image 6). Si la scène ne change plus, le système retourne lentement à son état initial. Dans le cas contraire, d'autres photos peuvent être prises et transmises selon le même processus.



Figure 5 : Transitions entre les modes *endormi* (1), *miroir* (2 à 3) et *transmission d'image* (4 à 6).

Une télécommande permet de basculer le logiciel dans un mode de consultation des photos prises par l'ensemble des systèmes connectés et de choisir celles que l'on veut conserver. Les autres photos perdent progressivement leurs couleurs et contrastes et finissent par disparaître au bout de quelques jours (Figure 6).

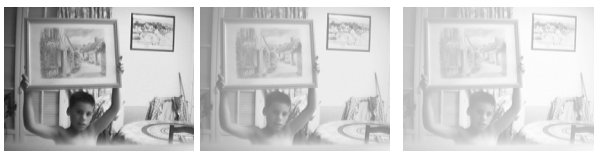


Figure 6 : Vieillesse d'une photo prise par le système.

Les transitions entre les différents modes de fonctionnement jouent un rôle extrêmement important dans l'interaction avec VideoProbe. L'utilisation combinée de la détection de mouvement et d'un retardateur permet de basculer simplement de la veille (absence de mouvement ou mouvement continu) à la prise de vue. Celle-ci peut se déclencher de manière implicite (une personne entre dans la pièce et va s'asseoir) ou explicite (une personne se place devant la caméra et attend la prise de vue). Le mouvement face à la caméra détermine indirectement la fréquence à laquelle le système transmet des images. Et bien que l'espace de manœuvre soit assez limité (au mieux, une image toutes les 15 secondes environ) on

peut néanmoins considérer qu'il s'agit là d'un système de communication multi-échelles.

La prise de vue est un processus lent pendant lequel les images affichées sont peu à peu transformées pour atteindre l'état dans lequel elles sont effectivement transmises (apparition progressive, puis recouvrement par le carré indiquant le temps restant avant la prise de vue). La lente dégradation des images reçues qui ne sont pas explicitement conservées procède de la même démarche. La modification de la présentation des images en fonction de leur statut permet de qualifier une nouvelle fois le système de multi-échelles.

Exemple 3 : MirrorSpace

MirrorSpace [25] est un système de communication vidéo qui prend en compte la notion de distance. Tandis que les systèmes traditionnels se contentent de créer un espace partagé correspondant à une distance interpersonnelle particulière, MirrorSpace est à l'inverse conçu pour offrir un continuum de distances permettant l'expression d'une grande variété de relations entre individus.

Les flux vidéo des lieux reliés par MirrorSpace sont affichés sur un écran unique fusionnant par transparence l'image des participants locaux et distants. Afin de permettre des formes de communication intimes où le regard joue un rôle très important, la caméra est placée au centre de l'écran. Un utilisateur peut ainsi se placer très près de celle-ci tout en étant toujours capable de voir la personne distante et de communiquer avec elle. Le dispositif comporte également un capteur de proximité qui mesure en continu la distance à la personne ou l'objet le plus proche. Les distances mesurées sont utilisées pour appliquer un effet de flou sur les images affichées.



Figure 7 : Diminution de l'effet de flou accompagnant l'approche d'une personne.

Le flou permet de percevoir le mouvement d'une personne éloignée avec un minimum d'implication. Il offre également un moyen intuitif pour initier ou éviter une transition vers un mode de communication plus engagé en se déplaçant simplement vers le dispositif (Figure 7) ou au contraire en s'en éloignant. MirrorSpace est sans conteste un excellent exemple de système de communication vidéo multi-échelles.

DISCUSSION

La communication par l'image est traditionnellement vue comme un enrichissement de la communication sonore. VideoServer, VideoProbe et MirrorSpace montrent que l'image peut être utilisée de manière utile et graduée sans liaison sonore associée. VideoProbe et MirrorSpace

montrent en particulier qu'il est possible de concevoir des systèmes de communication vidéo qui ne restituent pas nécessairement toute l'information disponible (i.e. toutes les images capturées, sans les modifier).

On peut vraisemblablement affirmer que les degrés d'engagement et d'attention associés à un flux d'images sont liés au niveau de détail de ce flux : plus les images sont grandes, colorées, lisibles et fréquentes, plus elles exposent celui qui est filmé et attire l'attention de celui qui peut les voir. À la manière des systèmes présentés dans la section précédente, les futurs systèmes de communication vidéo multi-échelles devront trouver de nouveaux moyens de faire varier ce niveau de détail.

Outre la taille et le taux de rafraîchissement des images, leur contraste, leur luminosité et leur netteté, d'autres caractéristiques intra ou inter-images pourraient être manipulées. Des procédés de rendu stylisé tels que ceux présentés par Karahalios et Donath [17] pourraient être utilisés pour mettre en œuvre des modes de communication légers (Figure 8, gauche). Des procédés de composition spatiale et/ou temporelle des images pourraient au contraire enrichir le message transmis, en combinant par exemple des images plus anciennes à la dernière image capturée (Figure 8, droite).

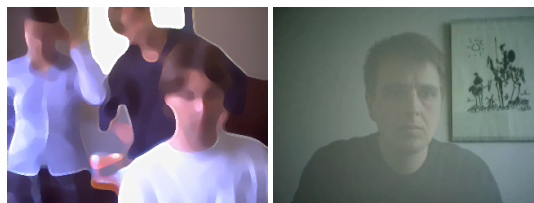


Figure 8 : Dégradation spatiale et enrichissement temporel.

La composition temporelle des images permet de faire un lien entre des communications synchrones et asynchrones. Les transitions entre ces modes devront être particulièrement étudiées. On peut par exemple imaginer que le système soit capable de produire un résumé en images de quelques secondes montrant ce qui s'est passé dans une pièce (locale ou distante) durant l'absence d'une personne.

Le déclenchement des transitions pose de nombreux problèmes d'interaction. Comment initier une communication, si elle n'est pas permanente ? Comment l'interrompre, de manière temporaire ou définitive ? Comment passer d'un mode de communication à un autre ? Dans ce contexte, on peut envisager l'utilisation de la caméra comme périphérique d'entrée [4], éventuellement accompagnée d'autres capteurs.

Enfin, un dernier axe de recherche prometteur concerne la combinaison de média de nature diverse (e.g. image vidéo, texte et dessins, son non-parlé, parole). Les transitions entre les différentes formes de communication ren-

dues possibles par la combinaison de ces média seront là encore à étudier. Il serait ainsi intéressant de permettre une continuité et une complémentarité de service entre l'échange de mini-messages par téléphone mobile, un dispositif vidéo similaire à MirrorSpace et une ligne de téléphonie fixe.

CONCLUSION

Dans cet article, j'ai décrit une nouvelle approche de la communication médiatisée qui vise à concevoir des systèmes multi-échelles, c'est-à-dire offrant un niveau d'implication variable et permettant des transitions fluides entre ces niveaux. J'ai présenté trois systèmes de communication vidéo qui illustrent cette approche ainsi que plusieurs pistes pour des travaux futurs. L'approche proposée ne se limite pas à la conception de systèmes vidéo. J'espère donc qu'elle séduira d'autres chercheurs et trouvera d'autres champs d'application dans le domaine de la communication médiatisée.

BIBLIOGRAPHIE

1. QuickTime and MPEG-4: Now Featuring H.264. Technology Brief, Apple, April 2005.
2. Brave, S. and Dahley, A. inTouch: a medium for haptic interpersonal communication. In CHI'1997 extended abstracts, pages 363-364, 1997. ACM Press.
3. Chang, A., Resner, B., Koerner, B., Wang, X. and Ishii, H. LumiTouch: an emotional communication device. In CHI'2001 extended abstracts, pages 313-314, 2001. ACM Press.
4. S. Conversy, N. Roussel, H. Hansen, H. Evans, M. Beaudouin-Lafon and W. Mackay. Partager les images de la vie quotidienne et familiale avec video-Probe. In *Proceedings of IHM 2003*, pages 228-231, Novembre 2003. ACM Press, International Conference Proceedings Series.
5. Crowley, J., Coutaz, J. and Bérard, F. Perceptual user interfaces: things that see. *Communications of the ACM*, 43(3):54-64, March 2000.
6. Daft, R. and Lengel, R. Information richness: a new approach to managerial behavior and organizational design. In L. Cummings and B. Staw, editors, *Research in organizational behavior* 6, pages 191-233, 1984. JAI Press.
7. Dennis, A. and Valacich, J. Rethinking Media Richness: Towards a Theory of Media Synchronicity. In *Proceedings of HICSS '99*, pages 1017, 1999. IEEE Computer Society.
8. Egido, C. Videoconferencing as a Technology to Support Group Work: A Review of its Failure. In *Proceedings of CSCW'88*, pages 13-24, September 1988. ACM Press.

9. Finn, K., Sellen, A. and Wilbur, S., editors. *Video-Mediated Communication*, Lawrence Erlbaum, April 1997. 584 pages.
10. Fish, R., Kraut, R., Root, R. and Rice, R. Evaluating Video as a Technology for Informal Communication. In *Proceedings of CHI'92*, pages 37-48, 1992. ACM Press.
11. Fuchs, H., Bishop, G., Arthur, K., McMillan, L., Bajcsy, R., Lee, S., Farid, H. and Kanade, T. Virtual Space Teleconferencing Using a Sea of Cameras. In *Proceedings of the first International Symposium on Medical Robotics and Computer Assisted Surgery*, pages 161-167, June 1994.
12. Gaver, W., Moran, T., MacLean, A., Lövstrand, L., Dourish, P., Carter, K. and Buxton, W. Realizing a Video Environment: EuroPARC's RAVE System. In *Proceedings of CHI'92*, pages 27-35, 1992. ACM Press.
13. Gross, M., Würmlin, S., Naef, M., Lamboray, E., Spagno, C., Kunz, A., Koller-Meier, E., Svoboda, T., Van Gool, L., Lang, S., Strehlke, K., Vande Moere, A. and Staadt, O. blue-c: A spatially immersive display and 3D video portal for telepresence. In *Proceedings of SIGGRAPH 2000*, pages 819-827, 2003.
14. Hollan, J. and Stornetta, S. Beyond being there. In *Proceedings of CHI'92*, pages 119-125, 1992. ACM Press.
15. Ishii, H., Kobayashi, M. and Arita, K. Iterative Design of Seamless Collaboration Media. *Communications of the ACM*, 37(8):83-97, August 1994.
16. Jouppe, N. First steps towards mutually-immersive mobile telepresence. In *Proceedings of CSCW 2002*, pages 354-363, 2002. ACM Press.
17. Karahalios, K. and Donath, J. Telemurals: linking remote spaces with social catalysts. In *Proceedings of CHI '04*, pages 615-622, 2004. ACM Press.
18. Lipartito, K. PicturePhone and the Information Age: The Social Meaning of Failure. *Technology and Culture*, 44(1):50-81, January 2003.
19. Lombard, M. and Ditton, T. At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication*, 3(2), September 1997.
20. Mackay, W. Media Spaces: Environments for Informal Multimedia Interaction. In M. Beaudouin-Lafon, editor, *Computer-Supported Co-operative Work*, Trends in Software Series, John Wiley & Sons Ltd, 1999.
21. Nørretranders, T. The tree of talking. In *The user illusion: cutting consciousness down to size*, Penguin books, 1991.
22. Perlin, K. and Fox, D. Pad: An alternative approach to the computer interface. In *Proceedings of SIGGRAPH '93*, pages 57-64, 1993. ACM Press.
23. Riesenbach, R., Buxton, W., Karam, G. and Moore, G. Ontario Telepresence Project: Final report, Information technology research centre, Telecommunications research institute of Ontario, March 1995.
24. N. Roussel. Mediascape: a Web-based Mediaspace. *IEEE Multimedia*, 6(2):64-74, April-June 1999.
25. N. Roussel, H. Evans and H. Hansen. Proximity as an interface for video communication. *IEEE Multimedia*, 11(3):12-16, July-September 2004.
26. Salber, D. De l'interaction homme-machine individuelle aux systèmes multi-utilisateurs. Thèse de Doctorat, Université Joseph Fourier, Grenoble (France), September 1995.
27. Shneiderman, B. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In: *Proceedings of VL '96*, pages 336-343, 1996. IEEE Computer Society.
28. Short, J., Williams, E. and Christie, B. *The Social Psychology of Telecommunications*, Wiley, New York, 1976.
29. Strong, R. and Gaver, W. Feather, scent and shaker: supporting simple intimacy. In *Proceedings of CSCW'96*, November 1996. ACM Press.
30. Suzuki, K. and Hashimoto, S. Feellight: a communication device for distant nonverbal exchange. In *ETP '04: Proceedings of the 2004 ACM SIGMM workshop on Effective telepresence*, pages 40-44, 2004. ACM Press.