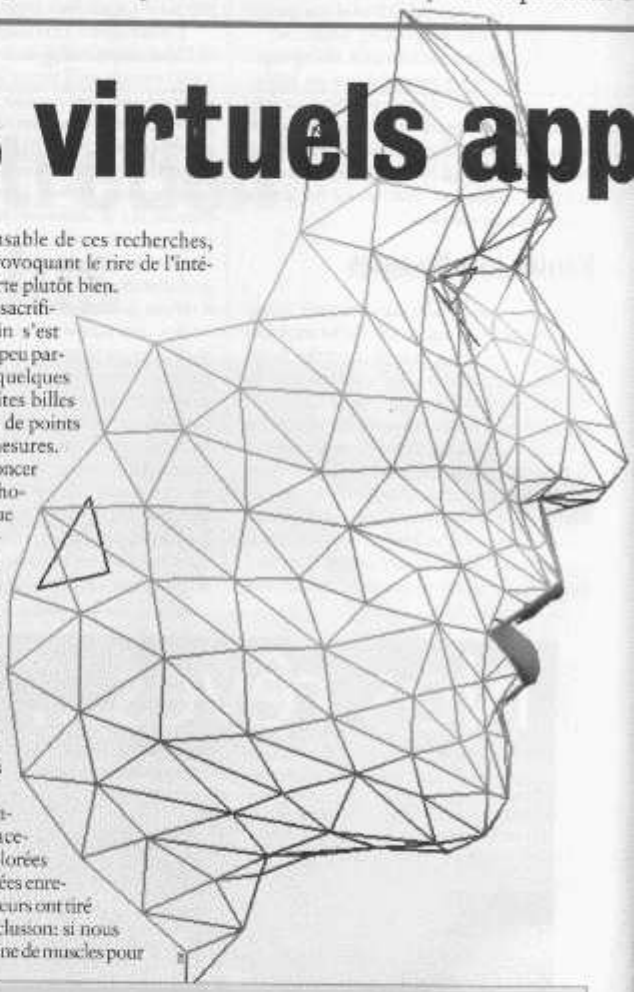
Langage Les têtes articulées numériques créées par des chercheurs français s'expriment c

Nos clones virtuels app



IMPRESSIONNANTS DE RÉALISME, AKI ROSS ET SES COMPAGNONS, LES PERSONNAGES DU LONG MÉTRAGE hollywoodien «Final Fantasy». Le nec plus ultra en la matière. Grain de la peau, plissé des tissus, finesse des cheveux: tout y est. A peine remarque-t-on une légère raideur dans la démarche. Pourtant, dès que les acteurs numériques se mettent à parler, il y a quelque chose qui cloche. Les timbres des voix, enregistrés par de vrais comédiens, n'y sont pour rien. Alors? L'impression de malaise vient en fait des mouvements, trop simplifiés, des lèvres et du visage des héros. D'une mauvaise représentation gestuelle de la parole.

Mais oui. Si la parole est le support du langage, elle est d'abord la résultante d'un ensemble de gestes, audibles mais aussi visibles (lire encadré). Prononcer des mots requiert avant tout la mise en action d'une série «d'articulateurs de la parole» qui vont de la mâchoire au larynx, en passant notamment par les lèvres et la langue. Autant de mouvements que les chercheurs de l'Institut de la communication parlée (l'ICP), à Grenoble, ont entrepris d'analyser. Ce qui les a conduits à créer de surprenants clones virtuels s'exprimant comme des êtres humains.

Pour produire un être numérique, rien de tel que de partir du réel et d'un être en chair et en os. Pierre Badin, un chercheur du laboratoire, «*s'est sacrifié*», comme l'explique le responsable de ces recherches, Gérard Bailly, provoquant le rire de l'intéressé qui s'en porte plutôt bien.

En guise de «sacrifice», Pierre Badin s'est donc vu coller un peu partout sur la face quelques centaines de petites billes colorées, servant de points de repère aux mesures. Puis il a dû prononcer les différents phonèmes de la langue française, cependant que des caméras filmaient les mouvements visibles de son visage et que des appareils du type imagerie à résonance magnétique analysaient ceux des organes cachés.

De l'analyse statistique des déplacements des billes colorées et des autres données enregistrées, les chercheurs ont tiré une première conclusion: si nous utilisons une centaine de muscles pour

La parole «vue» et «touchée»

Les sourds et les malentendants ne sont pas les seuls à lire la parole sur les lèvres de leur interlocuteur. Pour tout un chacun, la vision joue un rôle majeur dans la compréhension du langage. Nous en avons tous fait intuitivement l'expérience: il est beaucoup plus difficile de comprendre un correspondant parlant une langue étrangère au téléphone, lorsqu'on ne le voit pas. Les scientifiques ont aussi apporté leur preuve de la composante visuelle de la parole, en étudiant «l'effet cocktail party». Imaginez deux personnes qui conversent dans un environnement sonore. Lorsque le bruit de fond devient trop important, arrive un moment où, les yeux bandés, elles ne se comprennent plus du tout. Mais si elles peuvent voir la bouche de leur interlocuteur, elles gagnent 25% en intelligibilité, et environ 60% si elle peuvent observer leur visage. En regardant son interlocuteur, «on gagne environ dix décibels sur le rapport signal/bruit», résume Gérard Bailly, de l'Institut de la communication parlée à Grenoble. Tout se passe donc comme si le bruit ambiant avait baissé de 10 dB, ce qui représente, grosso modo, la différence d'intensité sonore existant entre «une voiture et un vélo qui passent». Mais la parole est aussi palpable, comme en témoigne la méthode TADOMA, mise au point à l'Ecole des aveugles Perkins, dans le Massachusetts, pour enseigner la compréhension et la production de la parole à des enfants sourds et aveugles. En plaçant leurs mains sur le visage de leur interlocuteur (sur ses joues, sa gorge, son menton), les utilisateurs de la technique peuvent «sentir» ses gestes vocaux. Ils peuvent ainsi pratiquement suivre une conversation normale.

mme de vrais êtres humains. Utiles pour les télécommunications du futur.

rennent à parler

Deux clones en grande conversation: parviendront-ils à inventer un nouveau langage?

HÉLÈNE LOEVENBRUCK, chercheuse au laboratoire de Grenoble, prête sa voix et son visage à ses collègues qui créeront son clone virtuel.

discourir, il suffit de retenir une dizaine de mouvements élémentaires pour obtenir une excellente description de la parole. Une dizaine, seulement, de *«degrés de liberté»* selon Gérard Bailly; deux affectant la mâchoire (qui peut s'ouvrir ou se fermer, se rétracter ou avancer), les autres concernant les lèvres, le larynx, la langue ou le vélum (ce petit clapet permettant d'envoyer de l'air dans le conduit nasal). Et cela ne vaut pas uniquement pour la langue française. Des expériences analogues, réalisées en anglais avec l'aide d'Hélène Loevenbruck, *«la clonette du laboratoire»* comme l'appellent ses collègues, mais aussi avec des visiteurs de langue maternelle allemande ou arabe, ont abouti aux mêmes constatations.

Dix mouvements: une sacrée simplification. Fort utile lorsqu'il s'agit de passer à la deuxième étape: la création du clone virtuel. Modélisé, le visage de Pierre Badin apparaît d'abord sur l'écran de l'ordinateur, en 3D, sous forme d'une juxtaposition de polygones (dont les sommets se trouvent là où étaient les billes). Ressemblant, le portrait, mais on peut faire mieux. Reste à *«l'habiller pour le rendre plus réaliste»*, explique Frédéric Elisei. Poursuivant sa démonstration, le chercheur *«plaque»* sur le visage artificiel *«de la texture»*, obtenue à partir d'un mélange de photos de l'original humain prises sous différents angles. Il aura fallu un mois à l'équipe pour réaliser la tête articulée. Mais pour les spectateurs que nous sommes, en quelques minutes le tour est joué. Le clone virtuel est là, il s'anime, il parle. Aussi vrai que nature.

Bureau numérique

Mais qu'on ne s'y trompe pas. Si rien n'exclut que les clones de l'ICP donnent de nouvelles idées aux créateurs de personnages numériques façon Aki Ross, ce n'est pas le premier objectif des chercheurs grenoblois. L'avenir de leurs têtes parlantes? Ils le voient plutôt dans les télécommunications du futur. Pourquoi pas participant à des téléconférences? Actuellement, constate Gérard Bailly, la visiophonie, qui transmet simultanément du son et de l'image, *«se traduit par un véritable patchwork: on voit les différents interlocuteurs sur l'écran, on entend des voix, mais on ne sait plus qui parle»*. A cette technique, on pourrait substituer la *«labiophonie»*: les participants à la conférence seraient remplacés par leurs clones, installés autour d'un *«bureau virtuel»*. Voilà qui améliorerait la visibilité de la scène et, de ce fait, l'intelligibilité des propos échangés. Les clones pourraient aussi se tailler une place de choix dans les ordinateurs communiquant vocalement avec leurs utilisateurs. Ce genre de machine *«devrait bientôt arriver sur le marché»*, pronostique le chercheur. On pourrait donc déjà voir grand: envisager qu'en tapant un texte sur le clavier, on générerait non seulement une voix de synthèse, mais on animerait les traits d'une tête virtuelle. Toujours dans le but de faciliter les communications, surtout dans un environnement bruyant.

Mais les têtes articulées pourraient aussi contribuer à l'apprentissage de langues étrangères, à la rééducation de patients ayant des problèmes orthophoniques ou encore fournir un outil de recherche pour l'étude du développement du langage chez l'enfant.

Dans la tête bien réelle de Gérard Bailly, les projets ne manquent pas. Même les plus fous, ceux qui ont trait à la vie artificielle. Le voilà rêvant à deux de ses clones, face à face et communiquant librement. *«Pourraient-ils s'inventer un langage à eux? Et si tel était le cas, est-ce que cela ressemblerait à un langage humain?»* Là, on sort de la science pour rejoindre Aki Ross dans le monde de la science-fiction.

ELISABETH GORDON

INFOS COMPLÉMENTAIRES SUR
WWW.hebdo.CH