

Informatique
& sciences numériques
(chaire annuelle 2021-2022)

Wendy E. Mackay

Informatique
& sciences numériques
(chaire annuelle 2021-2022)

Wendy E. Mackay

Interagir avec l'ordinateur

Cinquième Leçon
29 mars 2022

Wendy E. Mackay

L'interaction multimodale : Interagir avec tout le corps

Cinquième Leçon
29 mars 2022

Wendy E. Mackay

L'interaction humain-machine

Objectif :

se focaliser sur la tâche
pas sur la technologie qui
accomplit la tâche

L'interaction humain-machine

Au-delà
d'un clavier,
une souris et
un écran

Interfaces « WIMP »



Interfaces « WIMP »

Interface utilisateur graphique (GUI)

W Windows (Fenêtres)

I Icons Icônes

M Menus Menus

P Pointer Pointeur

Ordinateurs portables et de bureau mais
aussi les smartphones et les tablettes

Interfaces « WIMP »

Avantages

Bon compromis entre

facilité d'apprentissage

facilité d'utilisation

transfert de connaissances

puissance d'expression

Adapté pour des tâches en 2D :

rédiger un document

manipuler un tableur

Interfaces « WIMP »

Avantages

Bon compromis entre

facilité d'apprentissage

facilité d'utilisation

transfert de connaissances

puissance d'expression

Adapté pour des tâches en 2D :

rédiger un document

manipuler un tableur

Inconvénients

Ne passe pas bien à l'échelle :

applications complexes

nombreuses fonctionnalités

plusieurs utilisateurs

autres types d'entrées, 3D

Inadapté aux signaux continus

Traumatismes musculo-squelettiques

Interfaces « Post-WIMP »

van Dam (1997)

Objectif :

intégrer tous les sens en parallèle,
y compris le langage naturel et
les utilisateurs multiples

Interfaces « Post-WIMP »

van Dam (1997)

Objectif :

intégrer tous les sens en parallèle,
y compris le langage naturel et
les utilisateurs multiples

Le système doit :

détecter divers signaux
produits par l'utilisateur,
intentionnellement ou non
parole, mouvement, etc.

générer plusieurs sorties
que les utilisateurs peuvent
détecter et comprendre
parole, mouvement, etc.

O'Sullivan & Igoe (2004)



Comment
l'ordinateur
nous voit

O'Sullivan & Igoe (2004)



Les interfaces graphique
sur les ordinateurs standards,
les tablettes et les smartphones
nous perçoivent comme ayant un œil,
des (petites) oreilles et un doigt

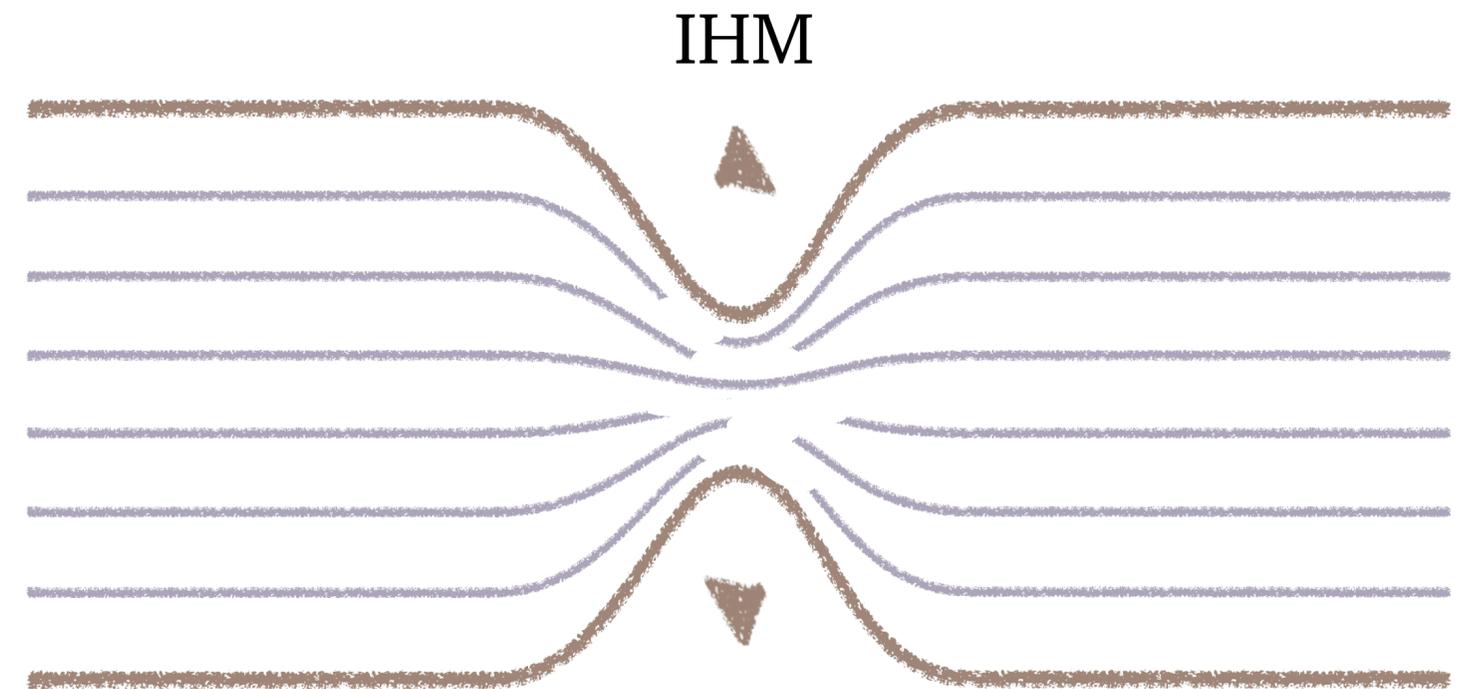
Bande passante de l'interaction

Cartes perforées

Capacités
humaines

Interaction

Capacités de
l'ordinateur



Bande passante de l'interaction

Interfaces graphiques actuelles

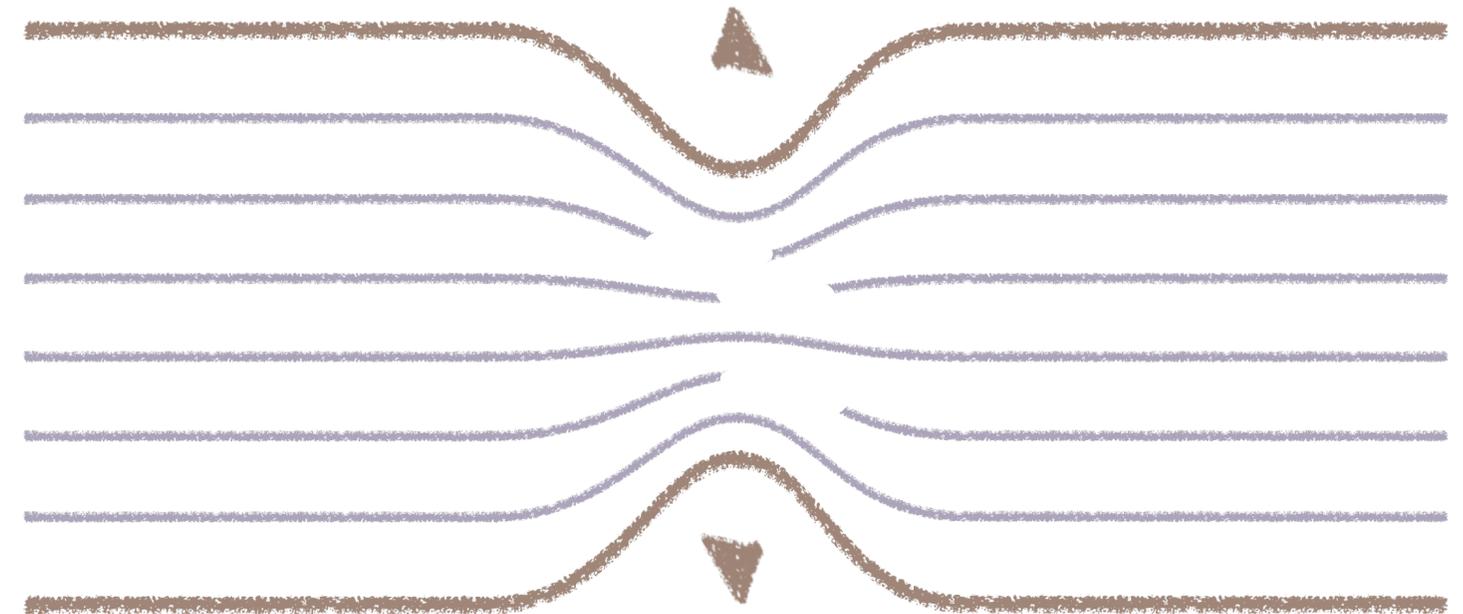
Clavier, souris, écran tactile

Capacités
humaines

Interaction

Capacités de
l'ordinateur

IHM



Bande passante de l'interaction

Interaction multimodale

Voix, geste

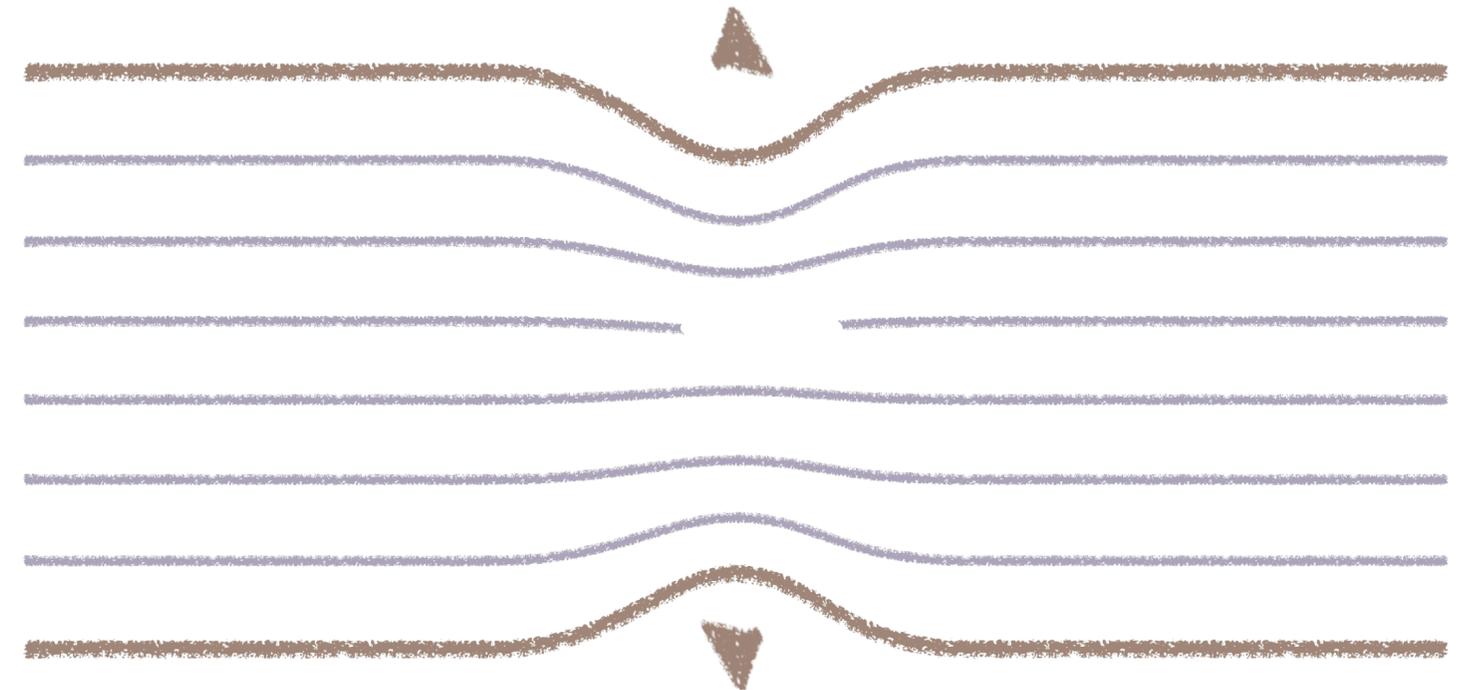
Réalité augmentée

Capacités
humaines

Interaction

Capacités de
l'ordinateur

IHM



Progrès dans les capacités de l'interface
ne suivent pas la loi de Moore

Capacités de l'ordinateur

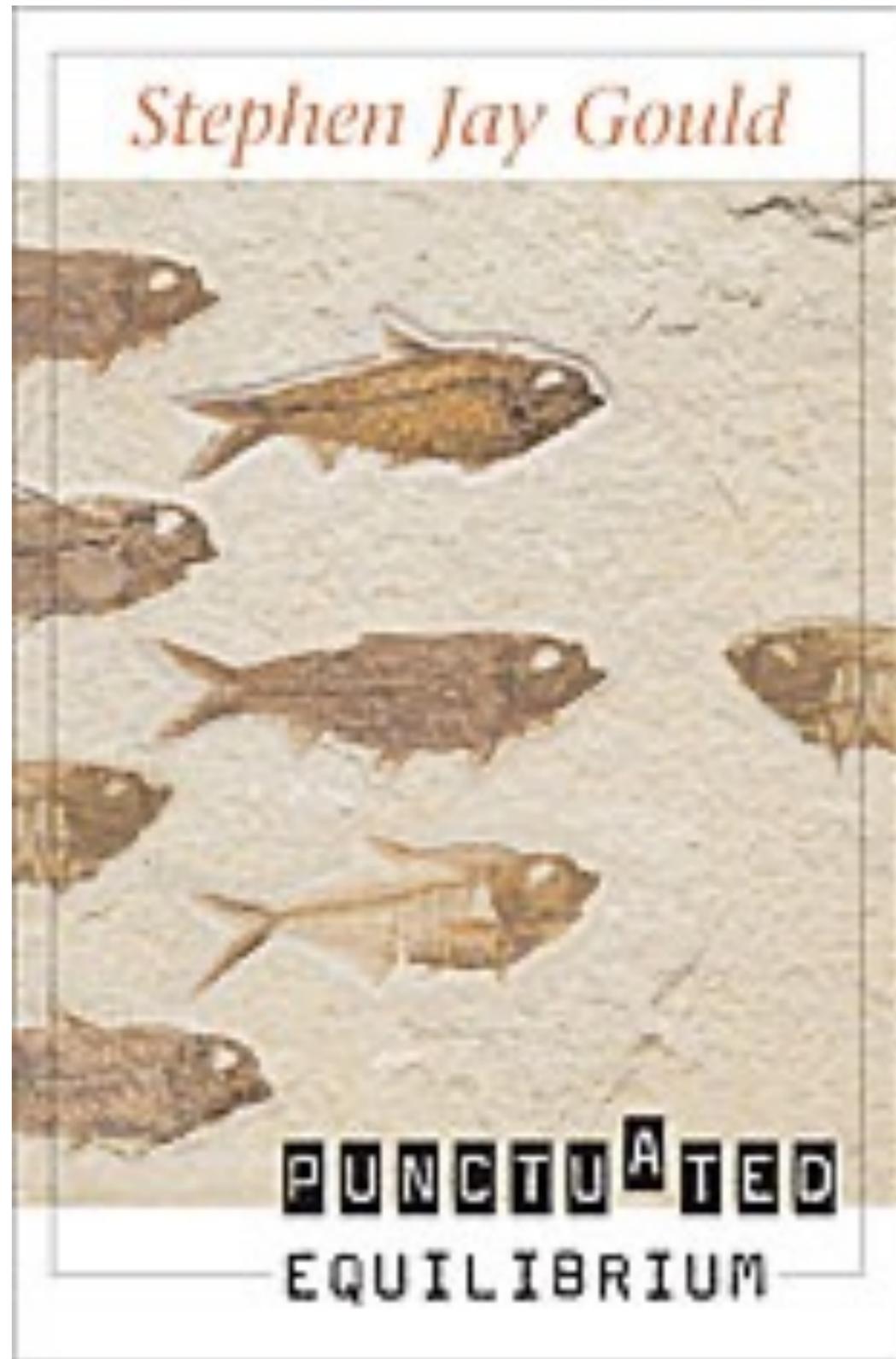
La loi de Moore (Moore, 1965)

Capacités de l'ordinateur

Processeur
Mémoire

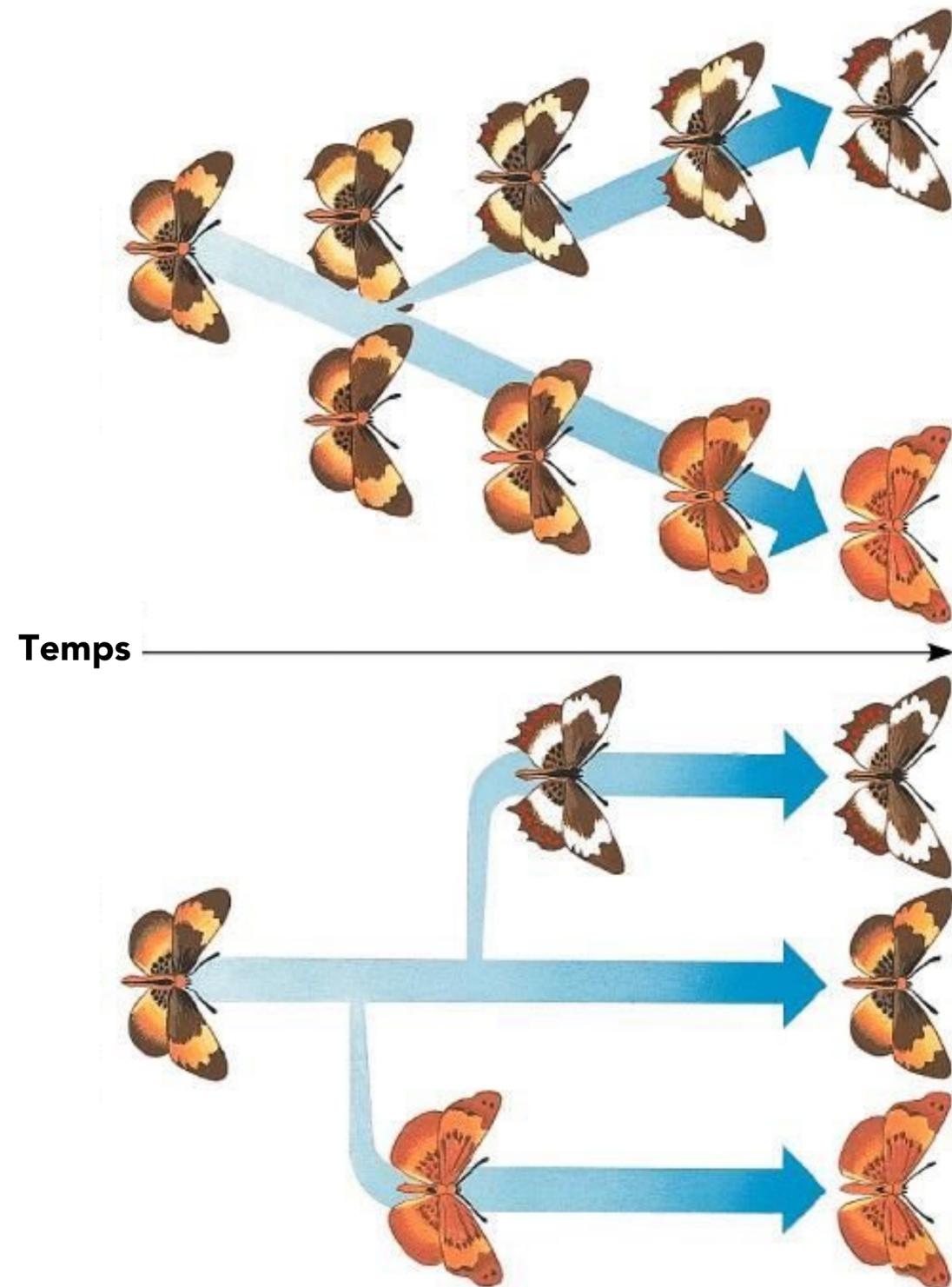
Temps

Progrès dans les capacités de l'interface
ne suivent pas la loi de Moore



Progrès dans les capacités de l'interface
ne suivent pas la loi de Moore

Plutôt un équilibre ponctué (Gould, 1972)
stable pendant une longue période
puis changement rapide



Progrès dans les capacités de l'interface
ne suivent pas la loi de Moore

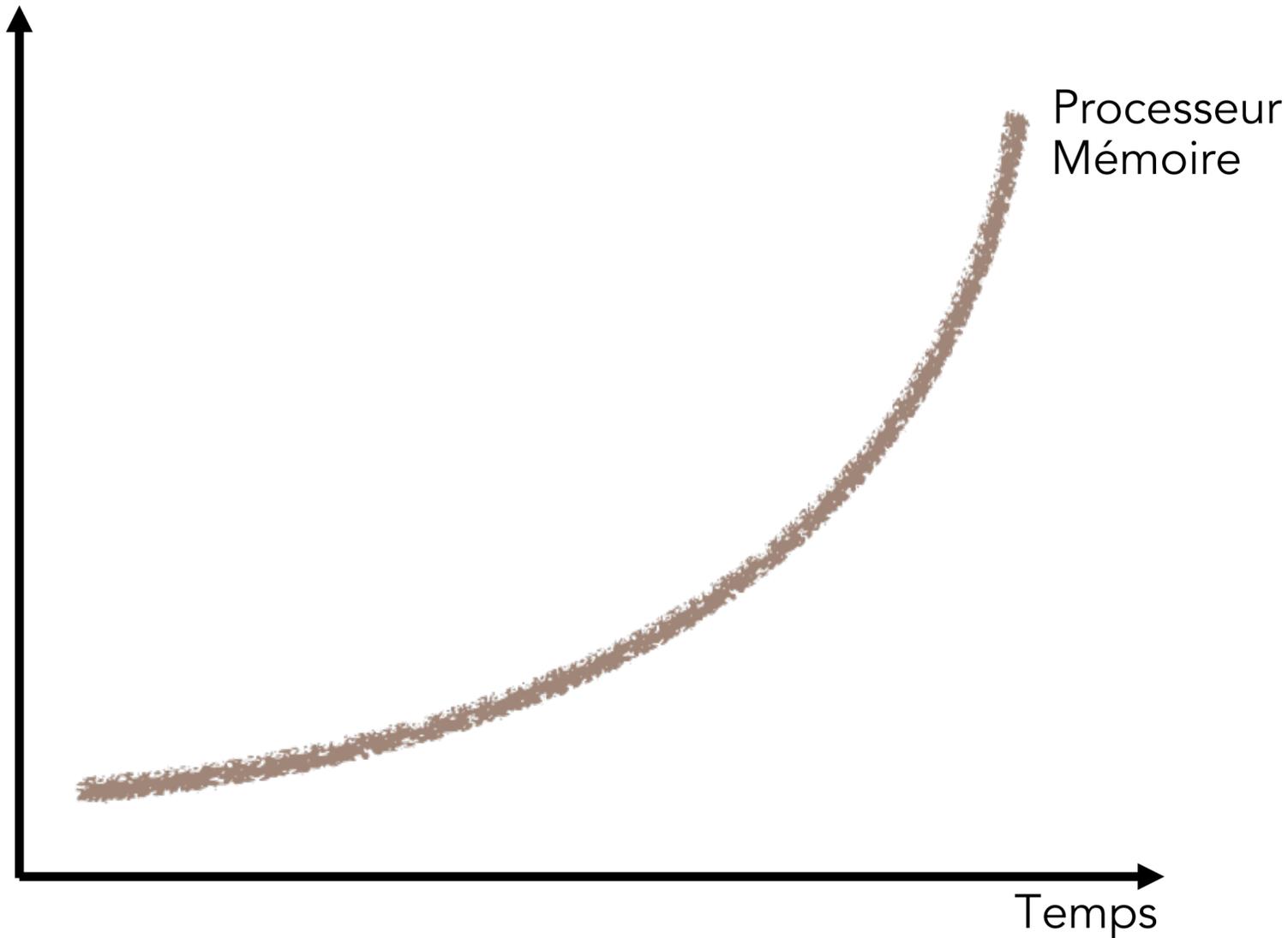
Plutôt un équilibre ponctué (Gould, 1972)
stable pendant une longue période
puis changement rapide

Biologie évolutive

Capacités de l'ordinateur

La loi de Moore (Moore, 1965)

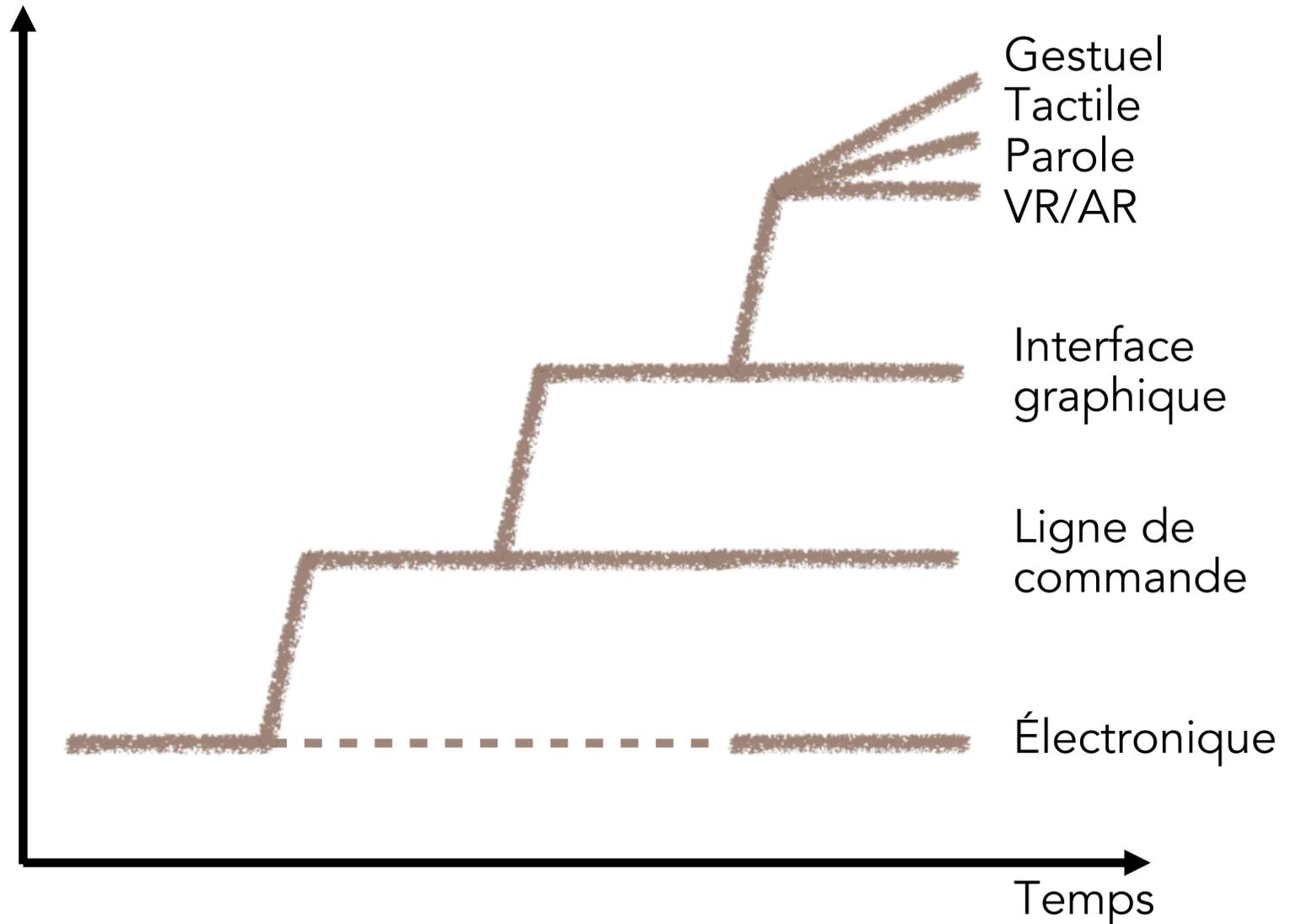
Capacités de l'ordinateur



Capacités de l'interface

Équilibre ponctué (Gould, 1972)

Capacités de l'interface



Interfaces « WIMP »

Systeme « universel »

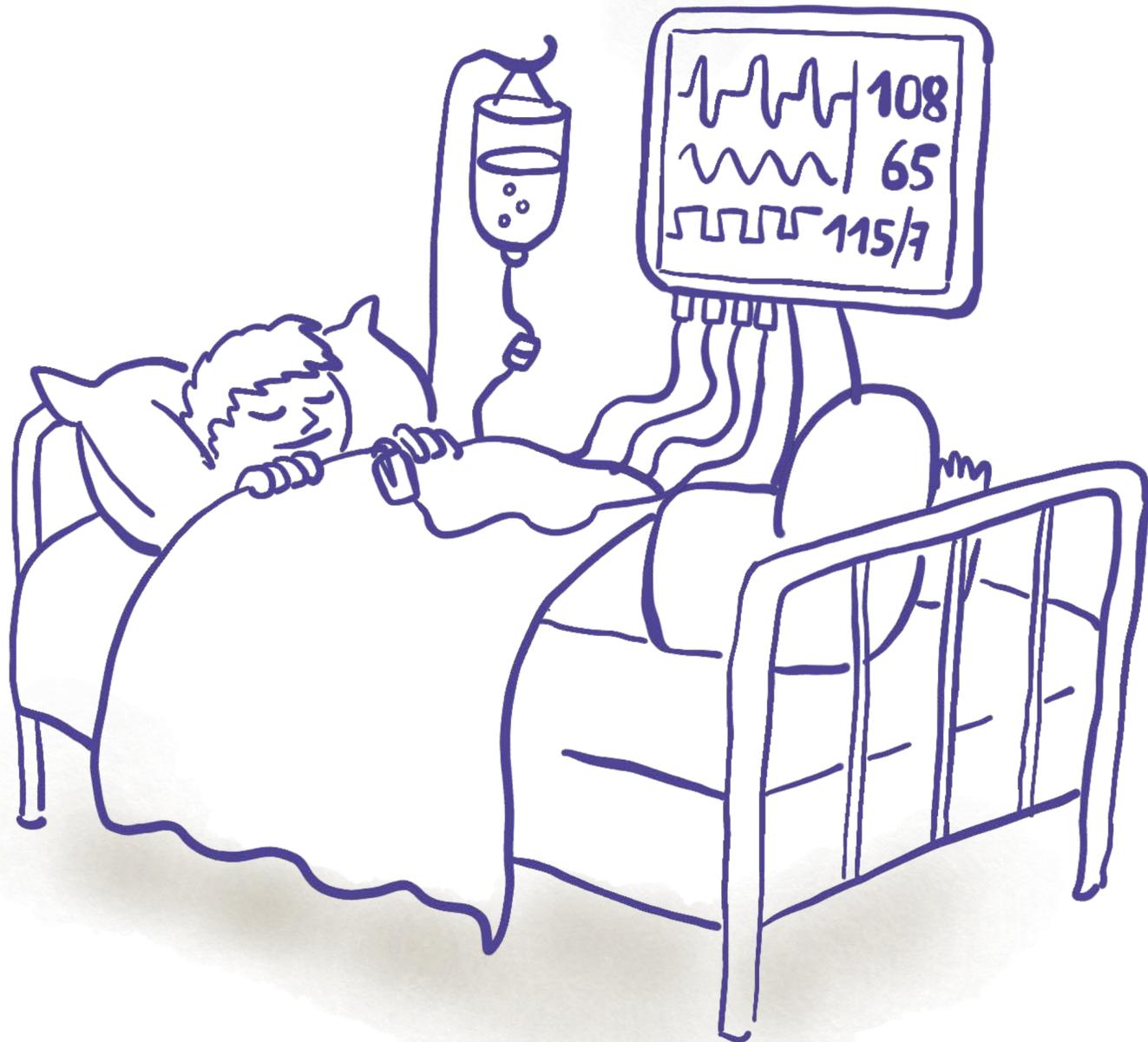
Une même forme d'interaction pour
répondre à plusieurs besoins



Interfaces « Post-WIMP »

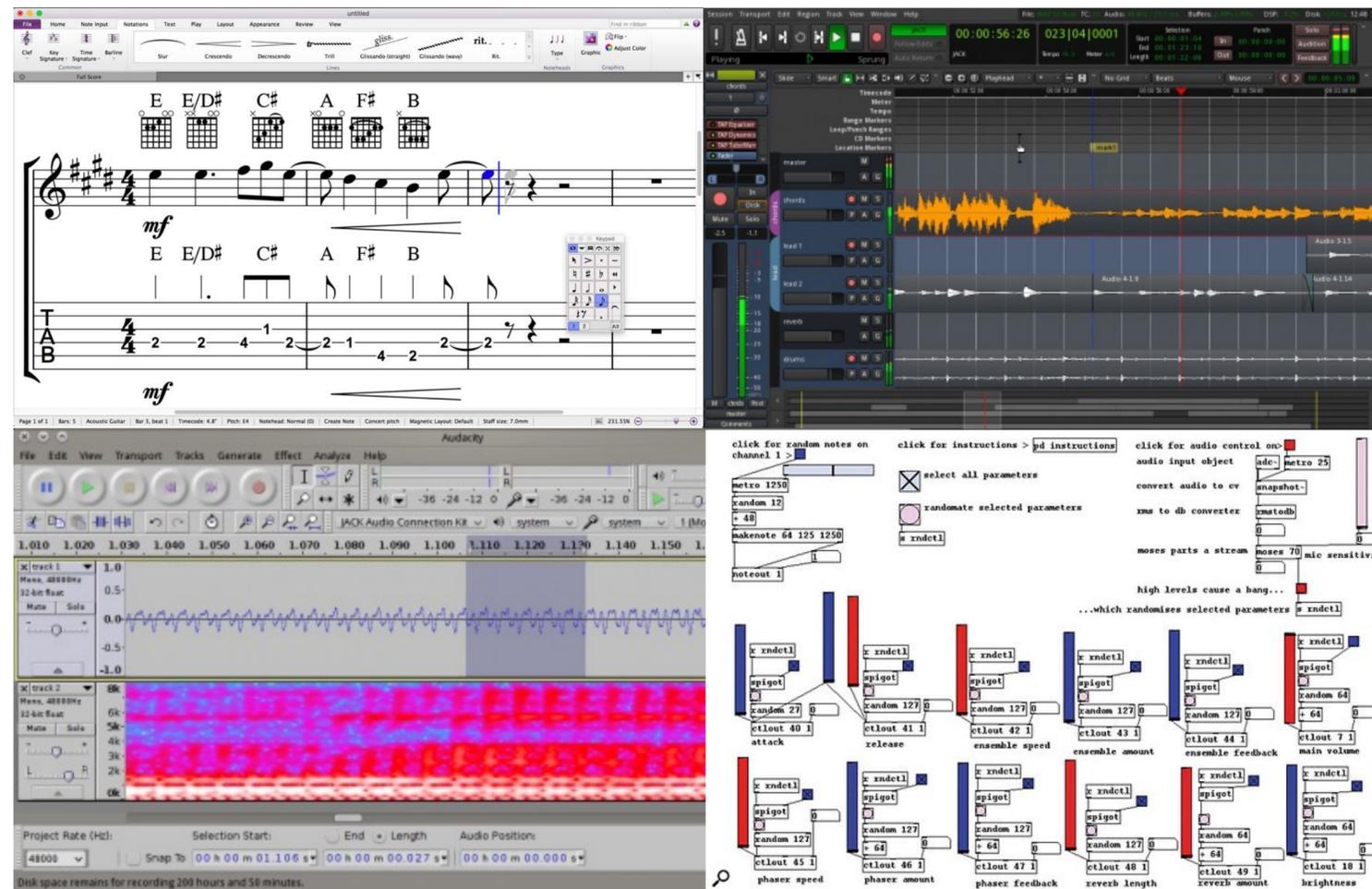
Systemes spécialisés

Plusieurs formes d'interaction pour répondre à des besoins spécifiques



Interfaces « WIMP »

Exemple : Composer de la musique



Nombreux types d'applications :

- notation musicale
- signaux audio
- structures sonores

Générer et organiser

le son et la musique

Créer des effets spéciaux pour

la composition ou le spectacle vivant

Interfaces « Post-WIMP »

Reactable (2003)



Un système conçu explicitement pour la performance musicale

Entrées : gestes, objets tangibles

Sorties : musique, table lumineuse

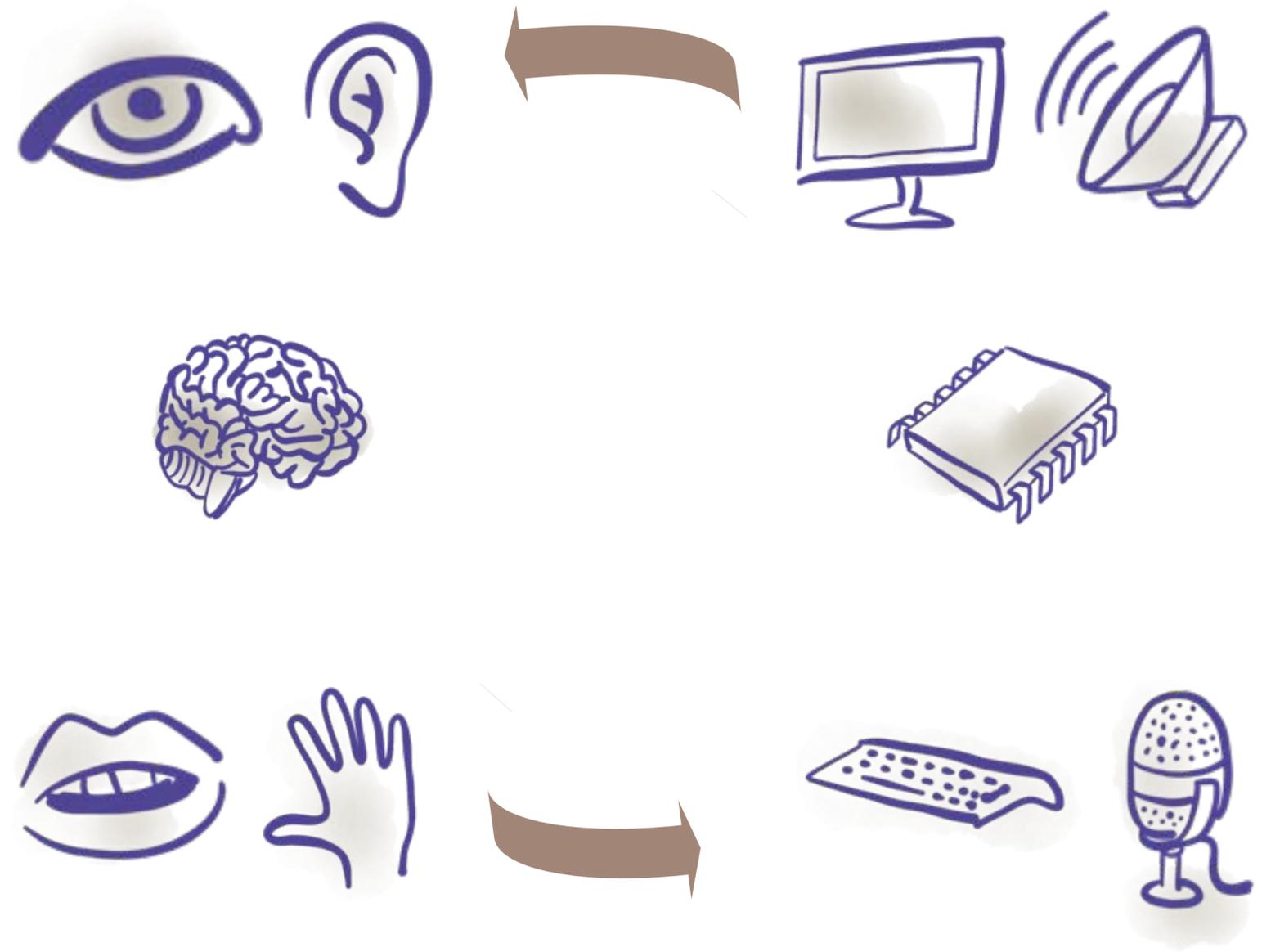
Parfait pour la musique, mais pas bon pour rédiger un document ou équilibrer un budget !

Boucle d'interaction

L'humain génère des informations
détectables par l'ordinateur

L'ordinateur génère des informations
détectables par l'humain

Mais qu'elle sont les possibilités ?



L'interaction multimodale

Interaction avec tout le corps

Modalités d'entrée et de sortie

Cinq sens : la vue, l'ouïe,
le toucher, l'odorat, le goût

Action : mouvement, parole

Modalité d'interaction :

Utilisation d'un canal sensori-moteur
en entrée et en sortie

Modalité d'entrée : humain -> ordinateur

Modalité de sortie : ordinateur -> humain

Modalités d'entrée et de sortie

Cinq sens : la vue, l'ouïe,
le toucher, l'odorat, le goût

Action : mouvement, parole

Modalité d'interaction :

Utilisation d'un canal sensori-moteur
en entrée et en sortie

Modalité d'entrée : humain -> ordinateur

Modalité de sortie : ordinateur -> humain

Modalités d'entrée

Clavier / souris

Détection de mouvement

Reconnaissance d'image

Reconnaissance de la parole

Modalités de sortie

Affichage (2D, 3D)

Son, parole, musique

Haptique

Interfaces multimodales

Dimensions clés

Nombre et types de canaux de communication

Nombre et types de modalités d'entrée

Choix d'utilisation des modes :

en parallèle, en série ou les deux

Interfaces multimodales

Dimensions clés

Nombre et types de canaux de communication

Nombre et types de modalités d'entrée

Choix d'utilisation des modes :

en parallèle, en série

Taille et type de vocabulaires de reconnaissance

Méthodes d'intégration des signaux

capteurs et canaux

Interfaces multimodales

Dimensions clés

Nombre et types de canaux de communication

Nombre et types de modalités d'entrée

Choix d'utilisation des modes :

en parallèle, en série

Taille et type de vocabulaires de reconnaissance

Méthodes d'intégration des signaux

capteurs et canaux

Types d'applications prises en charge

Interfaces multimodales

Dimensions clés

Nombre et types de canaux de communication

Nombre et types de modalités d'entrée

Choix d'utilisation des modes :
en parallèle, en série

Taille et type de vocabulaires de reconnaissance

Méthodes d'intégration des signaux
capteurs et canaux

Types d'applications prises en charge

Avantages :

flexibilité des modes de saisie

énoncés vocaux plus courts et simples
moins d'erreurs, plus simple à corriger

éviter de sur-utiliser un mode

Interfaces multimodales

Dimensions clés

Nombre et types de canaux de communication

Nombre et types de modalités d'entrée

Choix d'utilisation des modes :
en parallèle, en série

Taille et type de vocabulaires de reconnaissance

Méthodes d'intégration des signaux
capteurs et canaux

Types d'applications prises en charge

Avantages :

flexibilité des modes de saisie

énoncés vocaux plus courts et simples
moins d'erreurs, plus simple à corriger

éviter de sur-utiliser un mode

Adaptables à :

tâches, utilisateurs divers,

différences individuelles,

environnements en évolution

Interfaces multimodales

Dimensions clés

Nombre et types de canaux de communication

Nombre et types de modalités d'entrée

Choix d'utilisation des modes :
en parallèle, en série

Taille et type de vocabulaires de reconnaissance

Méthodes d'intégration des signaux
capteurs et canaux

Types d'applications prises en charge

Inconvénients :

analyse de signaux complexes

erreurs de reconnaissance

découvrabilité

comment savoir ce qui est possible ?

anthropomorphisme

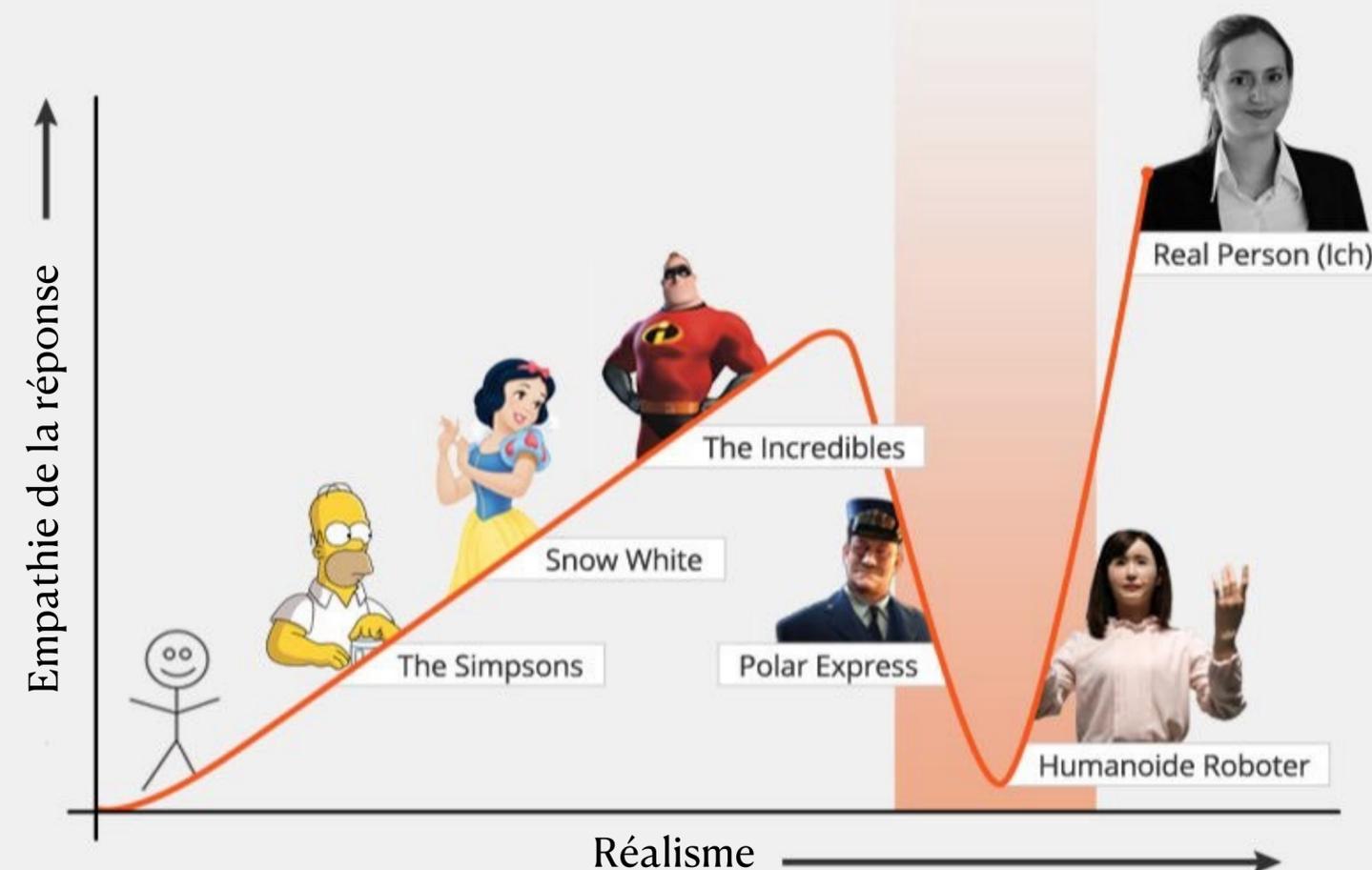
ressembler à la communication humaine

Uncanny valley (vallée de l'étrange)

Lorsque les robots ressemblent de plus en plus à des humains, ils semblent plus familiers jusqu'à ce que de subtiles imperfections d'apparence les rendent dérangentant

Vallée de l'étrange :
visuelle (avatar) ou
auditive (assistant vocal)

Masahiro Mori, 1970



Parole en entrée et en sortie

La parole est éphémère :
une fois prononcée, elle n'est plus là

L'interaction basée sur la parole nécessite
un contexte et un mode d'interaction

Comprendre l'intention signifie comprendre
ce que la personne veut vraiment et
ce qu'elle fait actuellement, et
nécessite une réponse adapté à la situation

Parole en entrée et en sortie

La parole est éphémère :

une fois prononcée, elle n'est plus là

L'interaction basée sur la parole nécessite

un contexte et un mode d'interaction

Comprendre l'intention signifie comprendre

ce que la personne veut vraiment et

ce qu'elle fait actuellement, et

nécessite une réponse adaptée à la situation

« Quand est le prochain bus ? »

Le système doit savoir :

où se trouve l'utilisateur maintenant

où se trouve le bus maintenant

combien de temps avant son arrivée

La réponse du système est importante :

réciter l'horaire complet des bus ?

dire : « Le prochain bus est à 9:12. » ?

dire : « Dans 12 minutes. »

ou une carte dynamique avec l'emplacement
du bus dans l'espace et le temps ?

Put That There

Schmandt (1979)

Première interface vocale multimodale
conversationnelle et collaborative

Reconnaissance des gestes dans l'espace,
Désignation sur un grand écran

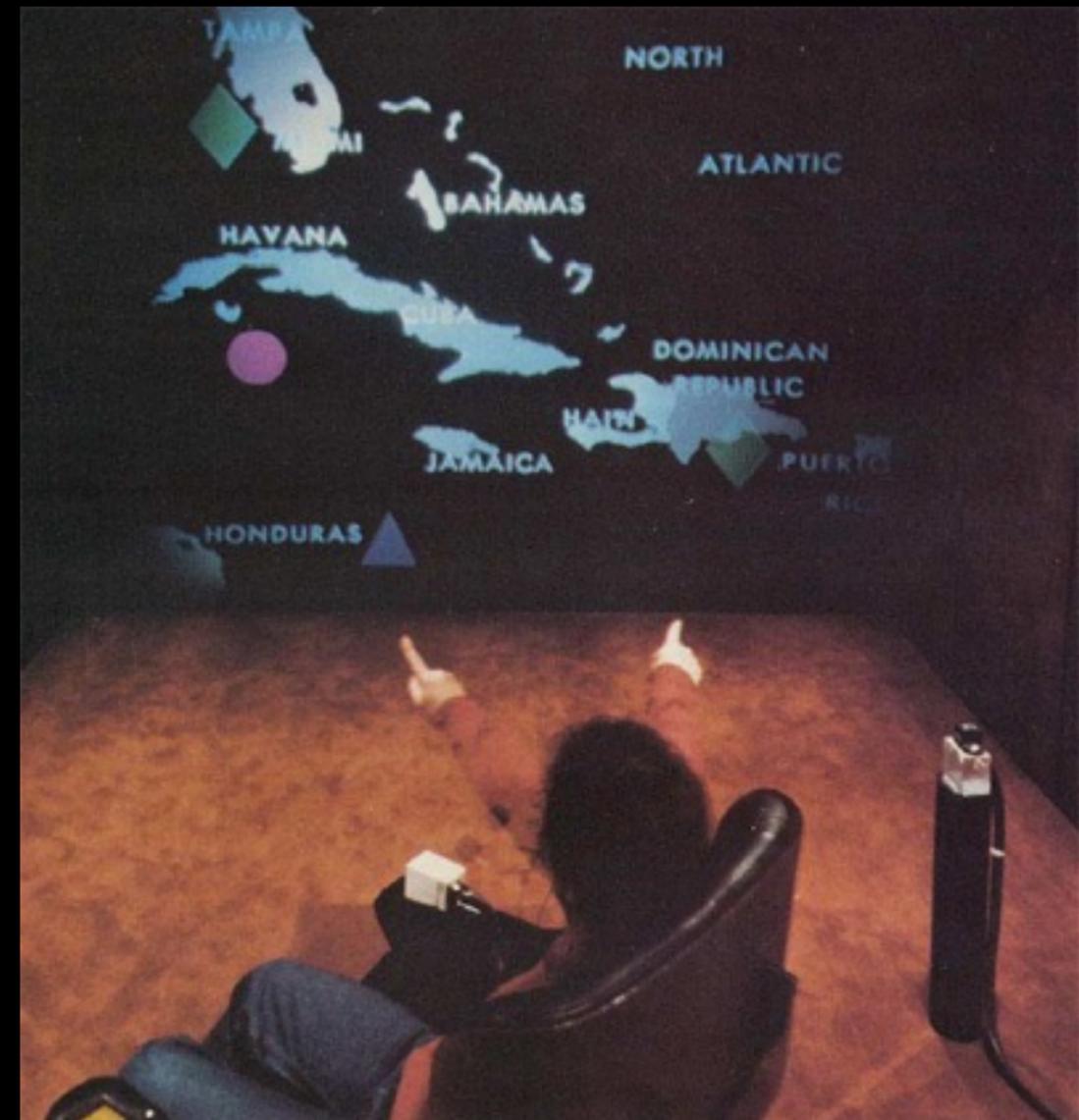
grâce à un capteur à six degrés de liberté

Reconnaissance de la parole

Montrer et dire :

mets ça	(l'objet)
là	(l'emplacement)

« Mets ça là »



Put That There

Schmandt (1979)

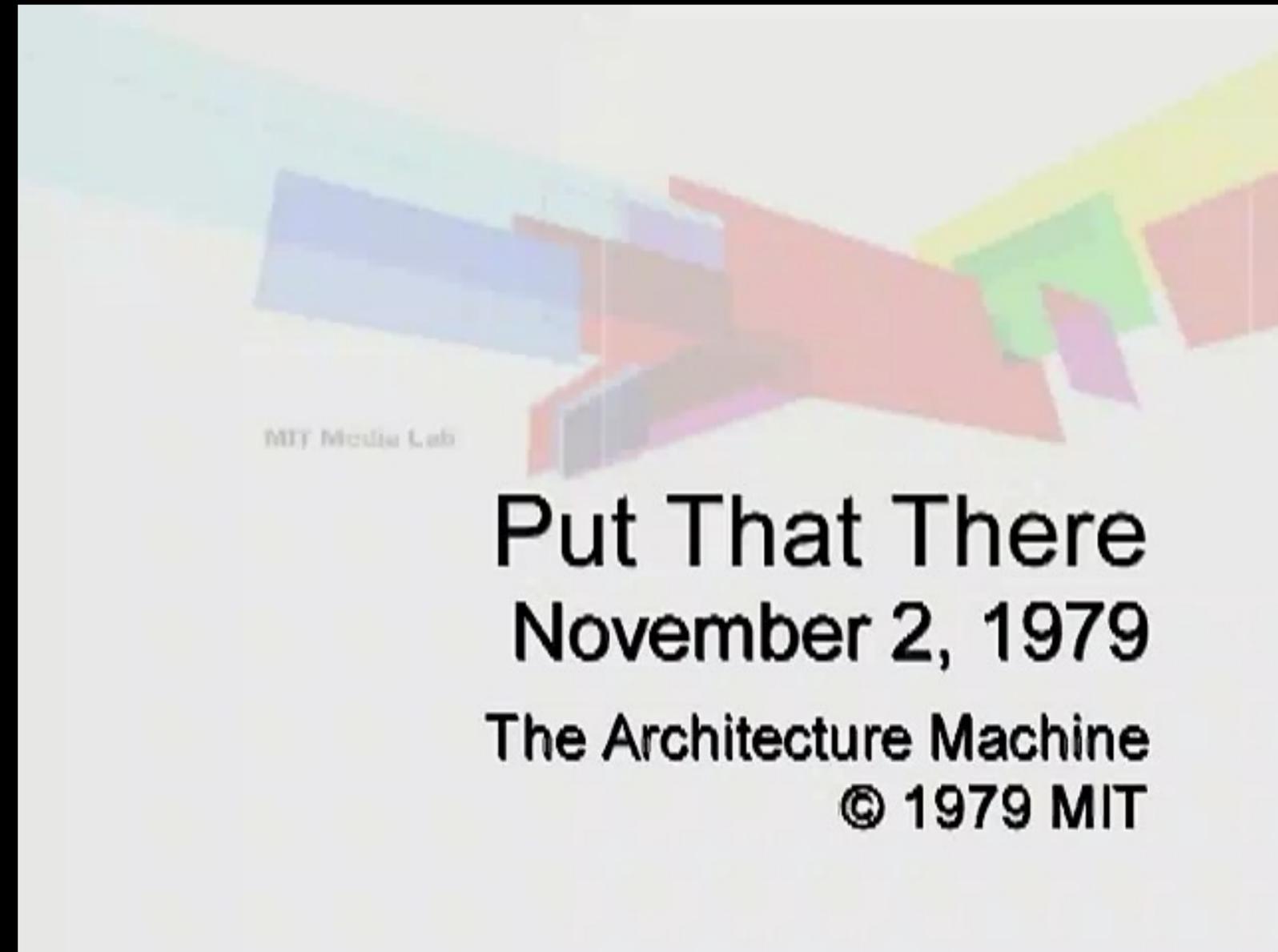
Put that there	Mets ça là
Carribbean	Caraïbes
Pay attention	Soyez attentif
Oil tanker	Pétrolier
Sail boat	Bateau à voile
Ship	Navire
Freighter	Cargo
Cruise ship	Navire de croisière



Put that there

Schmandt (1979)

Cela ne marche
pas toujours parfaitement !



Synthèse vocale

DECTalk (Klatt 1983)

Technologie « texte vers parole »

Digital Equipment Corporation



DECTalk Voices

Gendered Innovations Project

genderedinnovations.stanford.edu

Phone slave

Schmandt & Aarons (1984)

Guider l'utilisateur

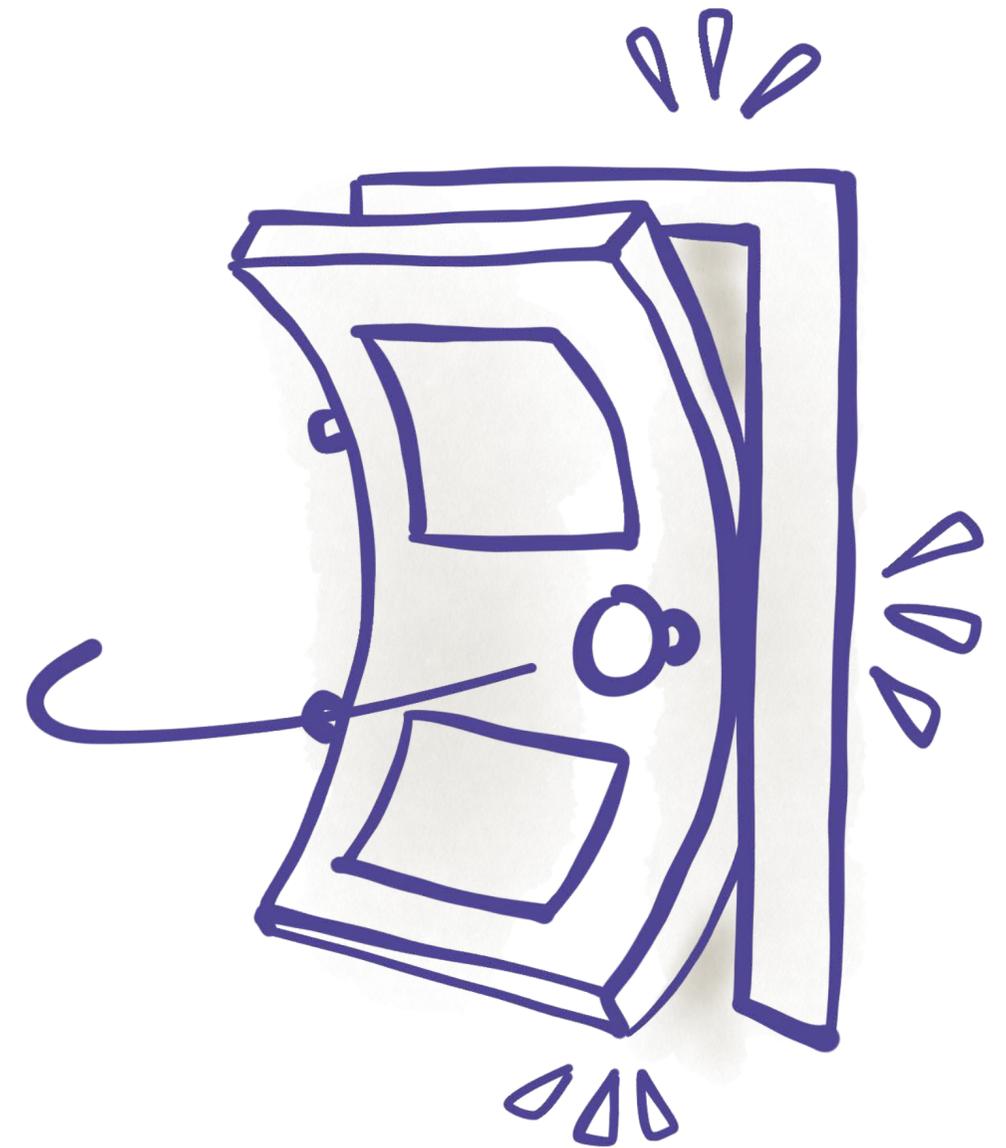
pour qu'il donne l'information
de manière à ce que l'ordinateur
puisse la restituer intelligemment,
tout en restant naturel



Perception directe

Gibson (1971)

Nous percevons l'interaction
qui provoque le son, pas le son lui-même



Perception directe

Sonic Finder (Gaver, 1989)

Icônes sonores (auditives):

Les sons correspondent
aux caractéristiques des objets
informatiques et aux interactions
qui les ont produit

Exemple : la taille d'un fichier

Son aigu petit

Son gravegrand

The image shows a blue rectangular area with the text "The SonicFinder" centered in a white, serif font. The word "The" is on the top line and "SonicFinder" is on the bottom line. The text has a slight shadow effect.

The
SonicFinder

Entrée vocale non verbale

Hum-a-Song (Kotsifakos et al., 2012)

Recherche de musique

« Requête par fredonnement »

Maintenant :

« Hey Google, quel est ce morceau ? »

la fredonner

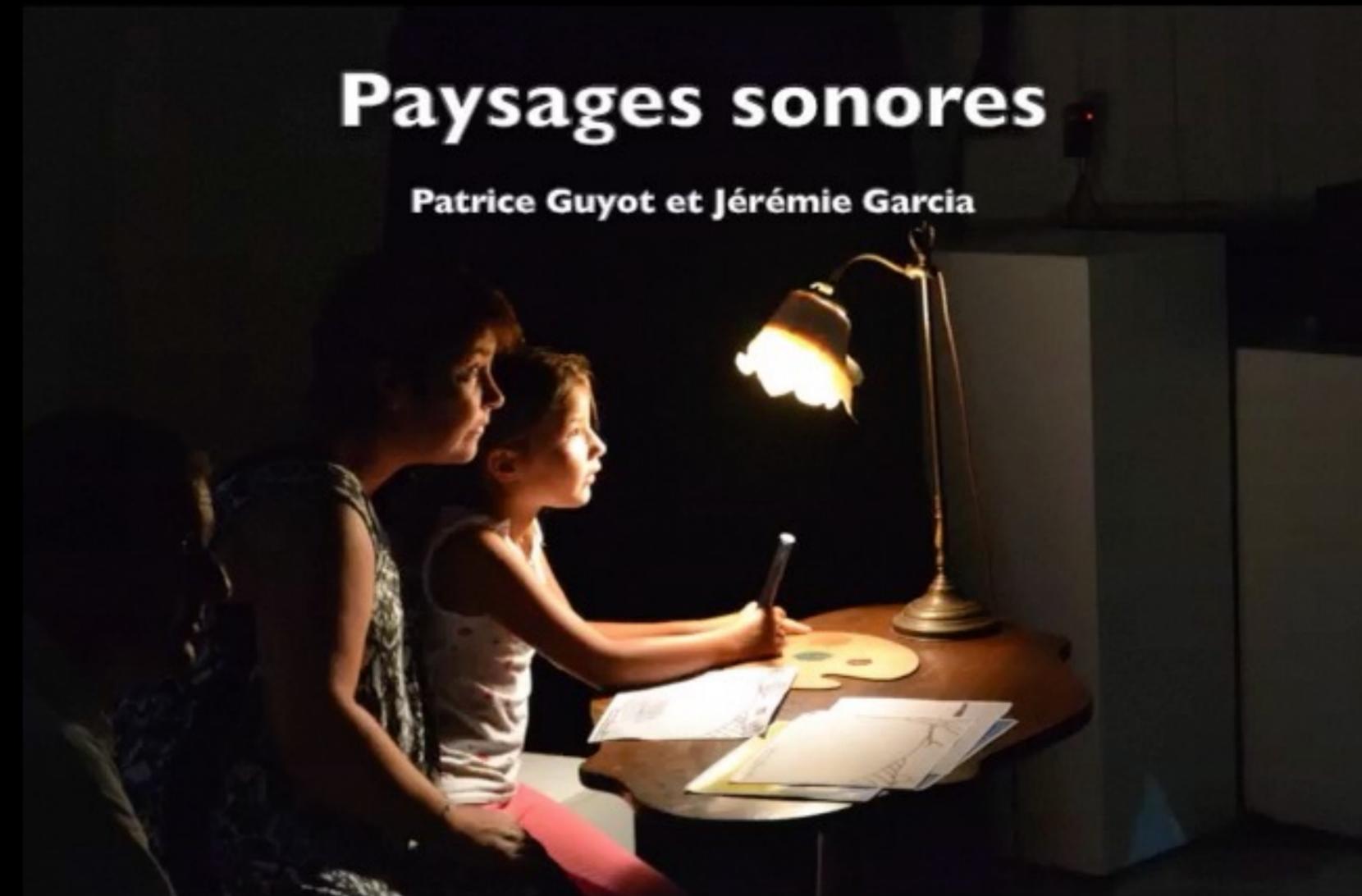


Paysages Sonores

Guyot et Garcia (2016)

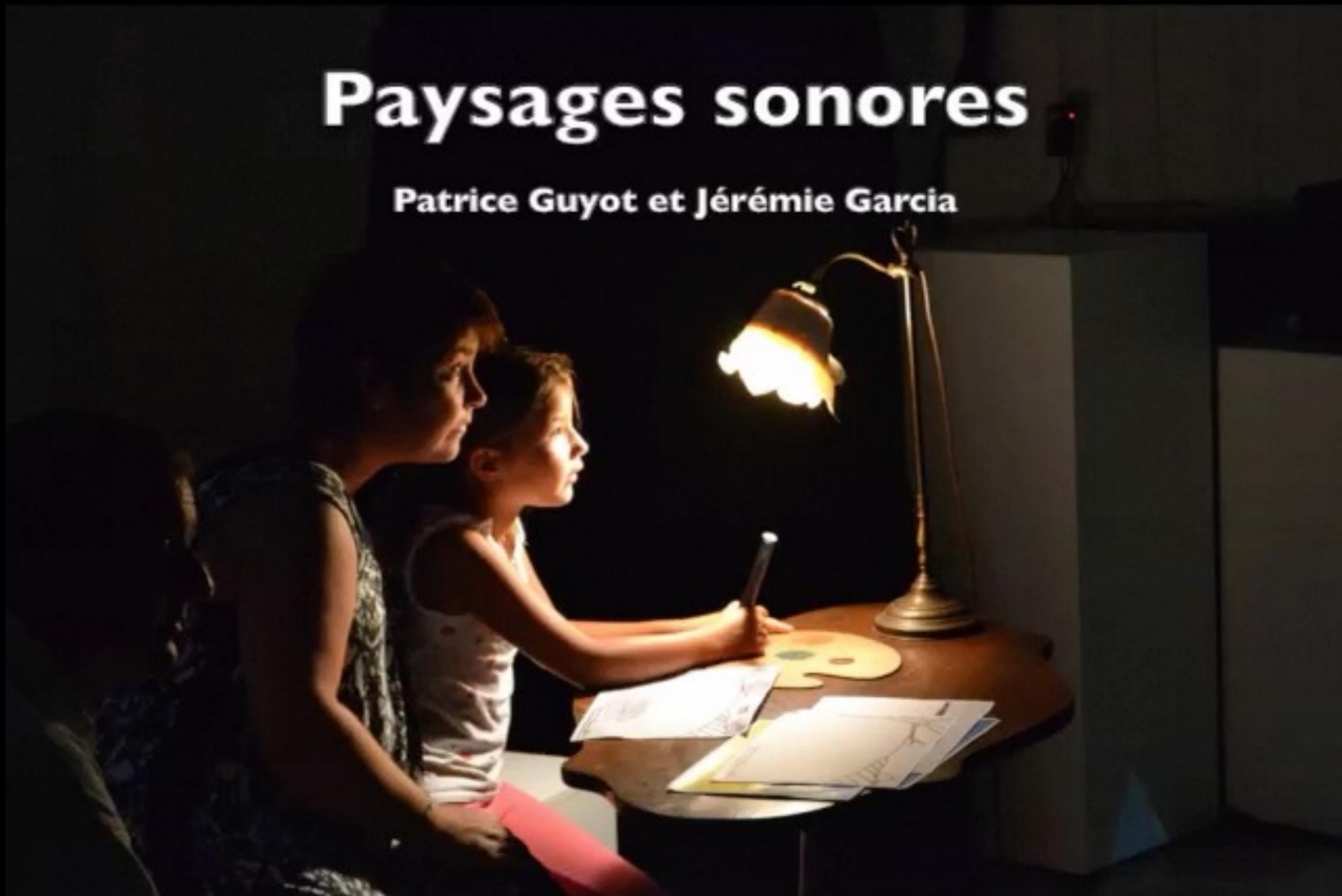
Un enfant dessine un paysage sonore :
sélectionner la « peinture » sonore
comme le vent ou l'eau
dessiner l'espace

Un autre enfant explore le paysage sonore
et crée les sons associés



Paysages sonores

Patrice Guyot et Jérémie Garcia



Interaction avec le corps

un doigt	toucher, pression
plusieurs doigts	multitouch (tables +)
une main	unimanuel (souris, stylo)
deux mains	bimanuel (clavier +)

DiamondTouch table interactive



Interaction avec le corps

Suivi du regard (Tobii)

un doigt	toucher, pression
plusieurs doigts	multitouch (tables +)
une main	unimanuel (souris, stylo)
deux mains	bimanuel (clavier +)
les yeux	regard (Tobii)

Suivi du regard (Tobii)



Interaction avec le corps

un doigt	toucher, pression
plusieurs doigts	multitouch (tables +)
une main	unimanuel (souris, stylo)
deux mains	bimanuel (clavier +)
les yeux	regard (Tobii)
les membres	muscles (Myo)

Myo Armband



Interaction avec le corps

un doigt	toucher, pression
plusieurs doigts	multitouch (tables +)
une main	unimanuel (souris, stylo)
deux mains	bimanuel (clavier +)
les yeux	regard (Tobii)
les membres	muscles (Myo)
le squelette	position (Kinect)
le corps entier	mocap (Vicon)

Suivi du squelette par Kinect (microsoft)



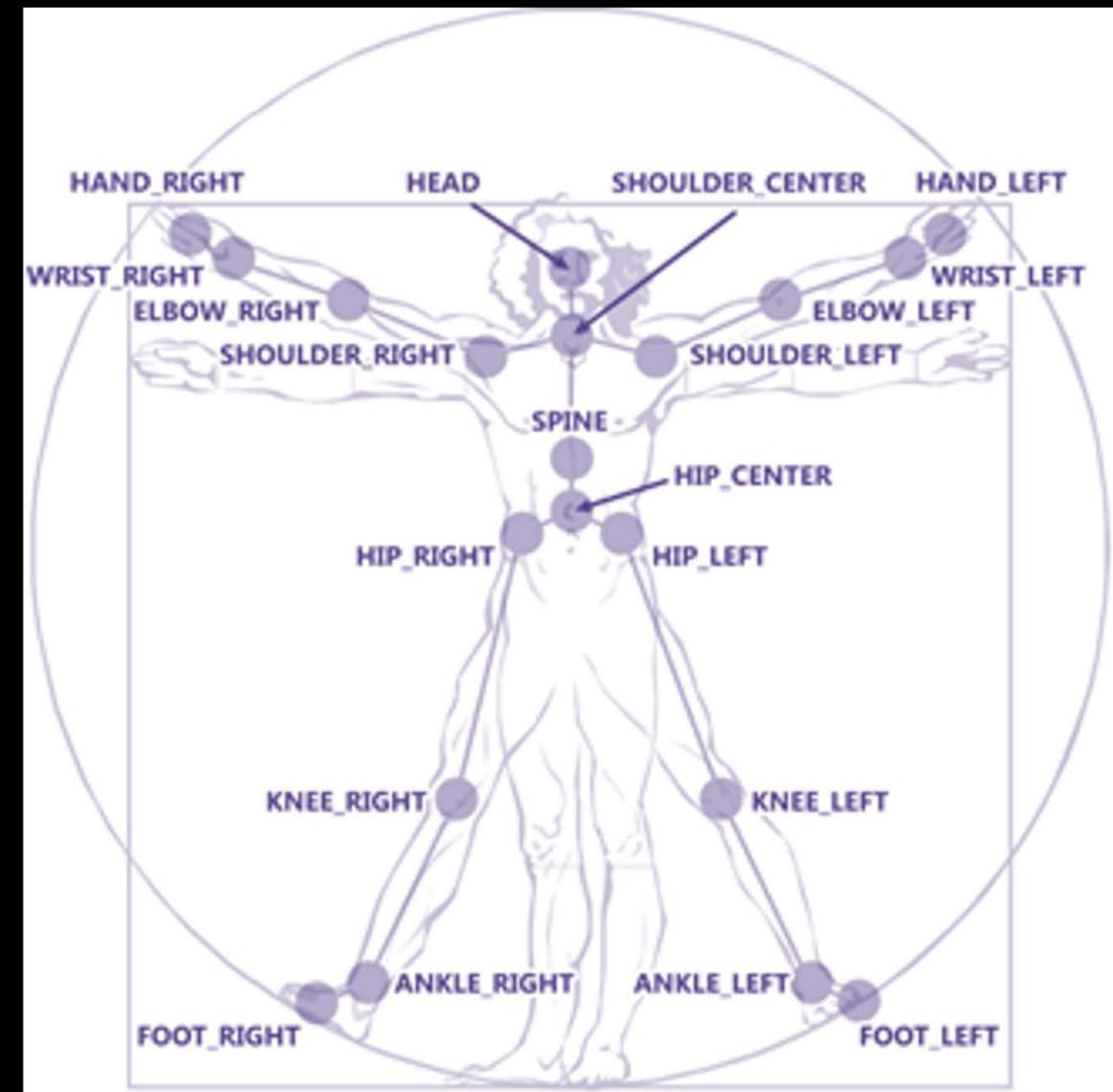
Interaction avec le corps

un doigt	toucher, pression
plusieurs doigts	multitouch (tables +)
une main	unimanuel (souris, stylo)
deux mains	bimanuel (clavier +)
les yeux	regard (Tobii)
les membres	muscles (Myo)
le squelette	position (Kinect)
le corps entier	mocap (Vicon)
le cerveau	BCI



Interaction avec le corps

un doigt	toucher, pression
plusieurs doigts	multitouch (tables +)
une main	unimanuel (souris, stylo)
deux mains	bimanuel (clavier +)
les yeux	regard (Tobii)
les membres	muscles (Myo)
le squelette	position (Kinect)
le corps entier	mocap (Vicon)
le cerveau	BCI



L'interaction gestuelle

PLATO IV (1972)

Enseignement assisté par ordinateur,
écran à plasma de 512 x 512 pixels
grille infrarouge tactile de 16 x 16 cases

Buxton (1985)

Première tablette capable de détecter
plusieurs points de contact simultanés
avec leur pression



L'interaction gestuelle

PLATO IV (1972)

Enseignement assisté par ordinateur,
écran à plasma de 512 x 512 pixels
grille infrarouge tactile de 16 x 16 cases

Buxton (1985)

Première tablette capable de détecter
plusieurs points de contact simultanés
avec leur pression

30 ans plus tard : iPhone 6S

écran multitactile sensible à la pression



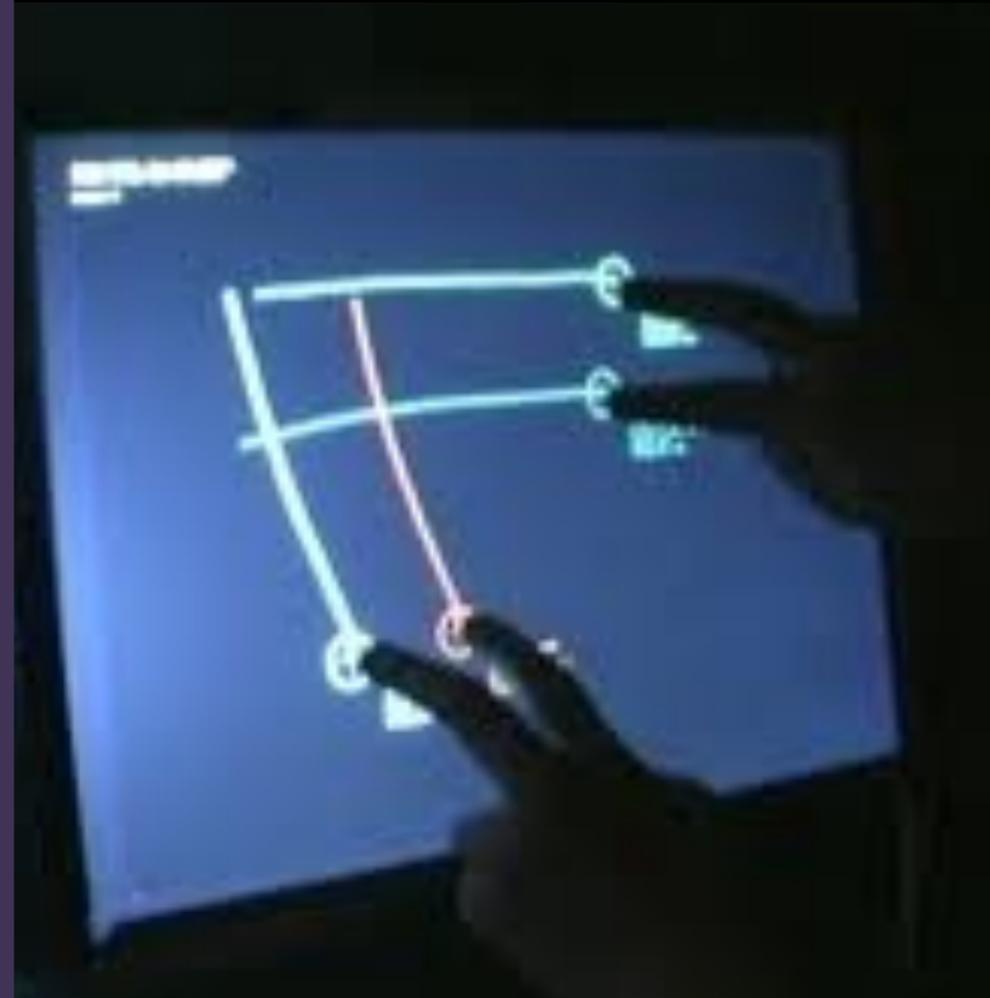
Entrée multitactile

Han (2006)

Table tactile

interaction « en grand »

très peu de latence



Entrée multitactile

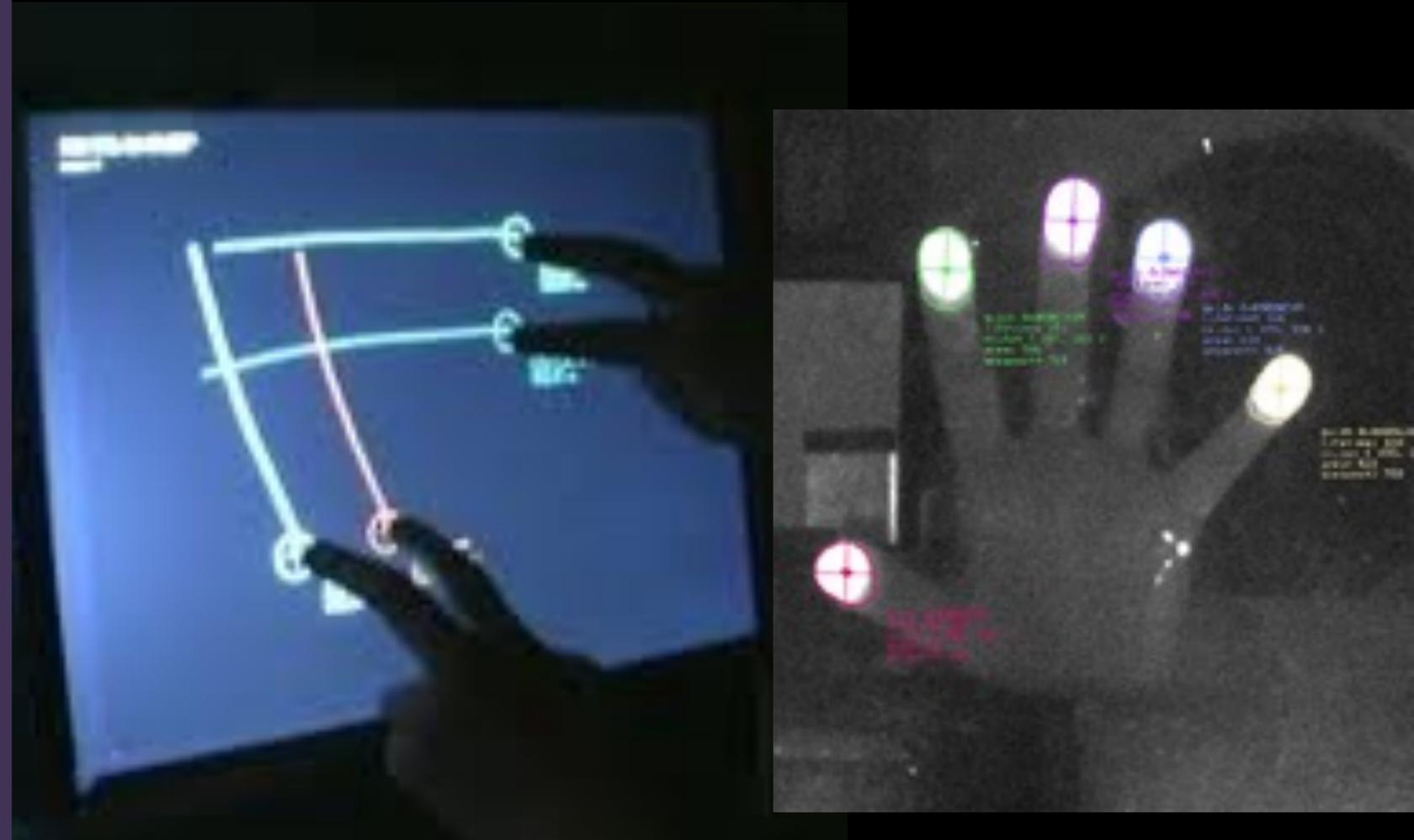
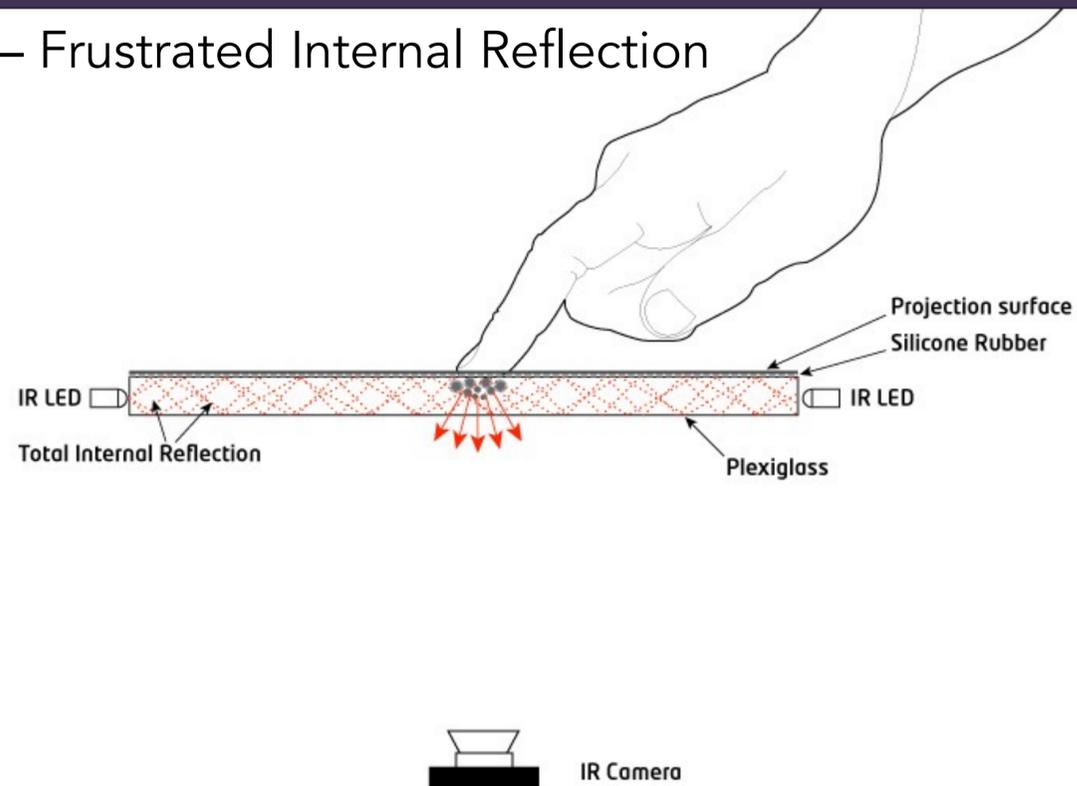
Han (2006)

Table tactile

interaction « en grand »

très peu de latence

FTIR – Frustrated Internal Reflection



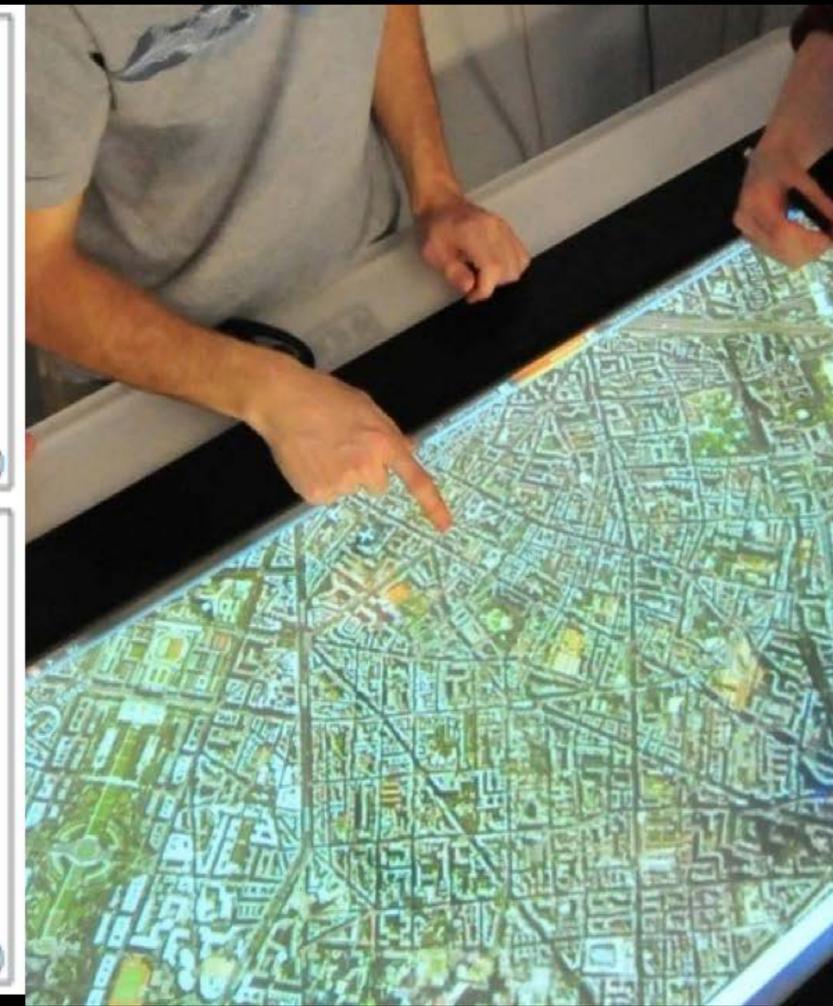
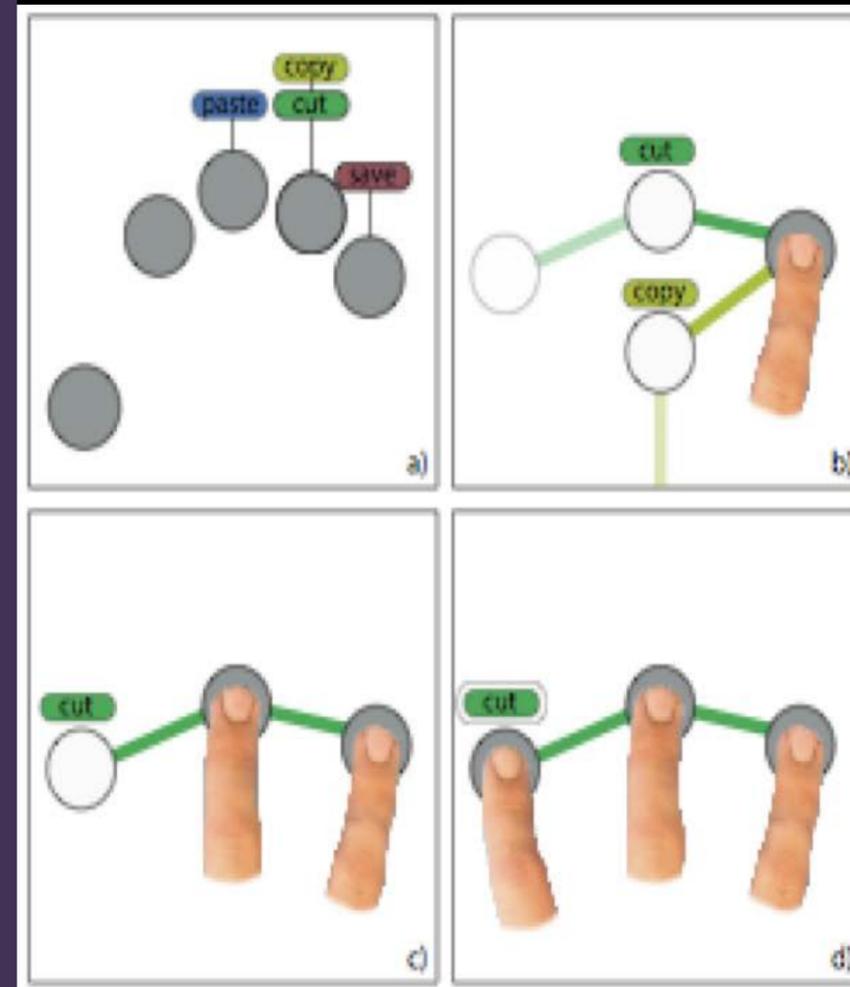
Multi-doigts

Arpège (Bau et al., 2014)

Au-delà des gestes à un et deux doigts
les accords offrent un vocabulaire
de commande beaucoup plus large

Mais comment les apprendre ?

Guide dynamique pour l'apprentissage
des commandes d'accords



Multi-doigts

Arpège (Bau et al., 2014)

Au-delà des gestes à un et deux doigts
les accords offrent un vocabulaire
de commande beaucoup plus large

Mais comment les apprendre ?

Guide dynamique pour l'apprentissage
des commandes d'accords



"relaxed" chords only involve fingers in their relaxed positions

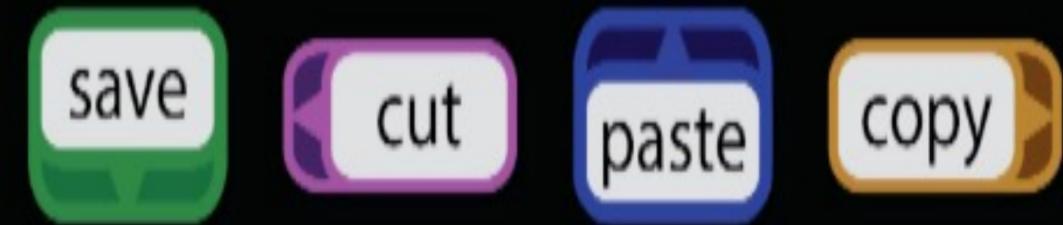
Multi-doigts

Arpège (Bau et al., 2014)

Au-delà des gestes à un et deux doigts
les accords offrent un vocabulaire
de commande beaucoup plus large

Mais comment les apprendre ?

Guide dynamique pour l'apprentissage
des commandes d'accords



Labels indicate what fingers are involved & their appropriate positions

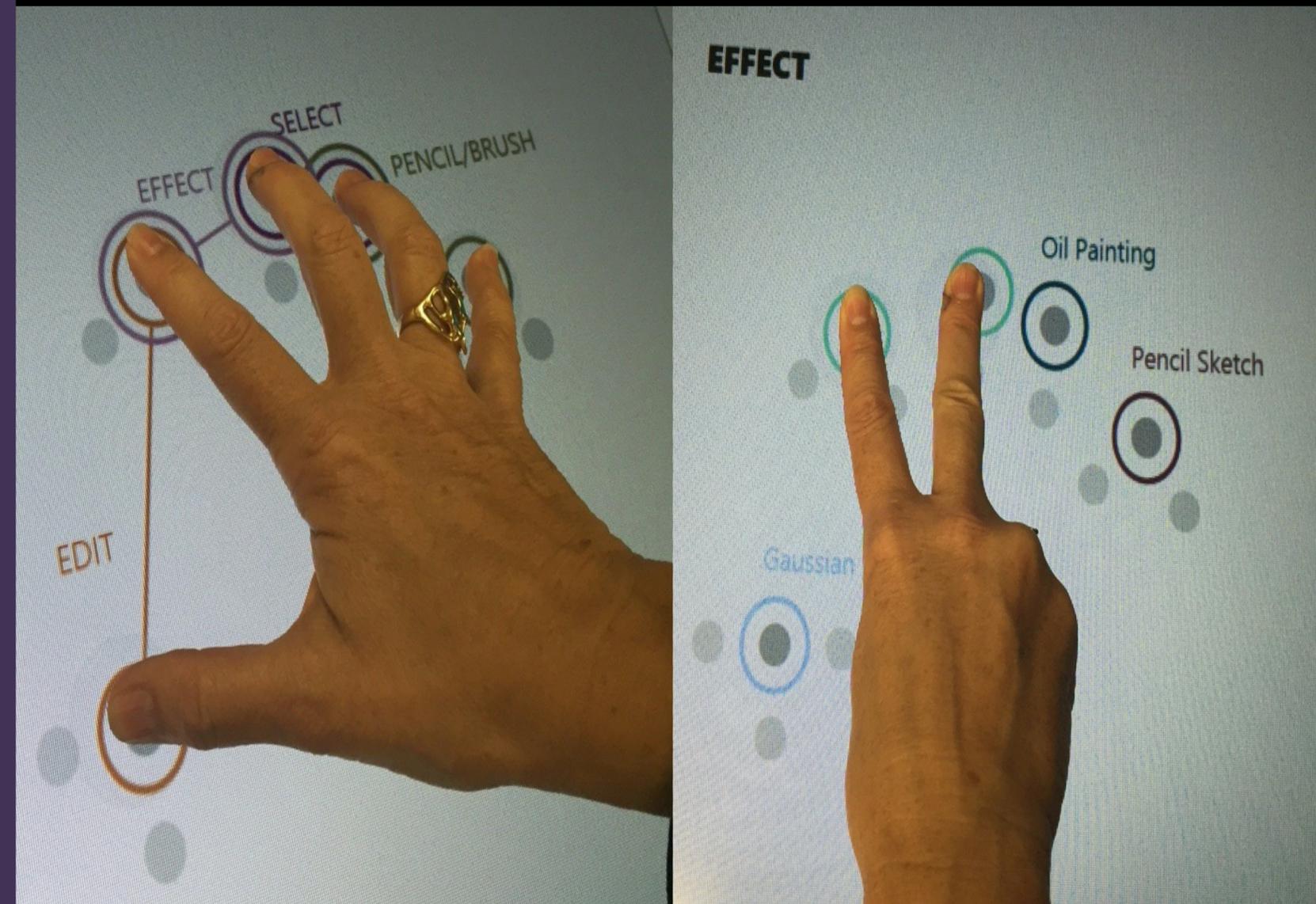
Interaction bimanuelle par accords

Dynachord (Li et al., 2016)

Séquences d'accords donnent
un plus grand vocabulaire d'accords

Ajustement dynamique des paramètres

Versions tablette et table interactive



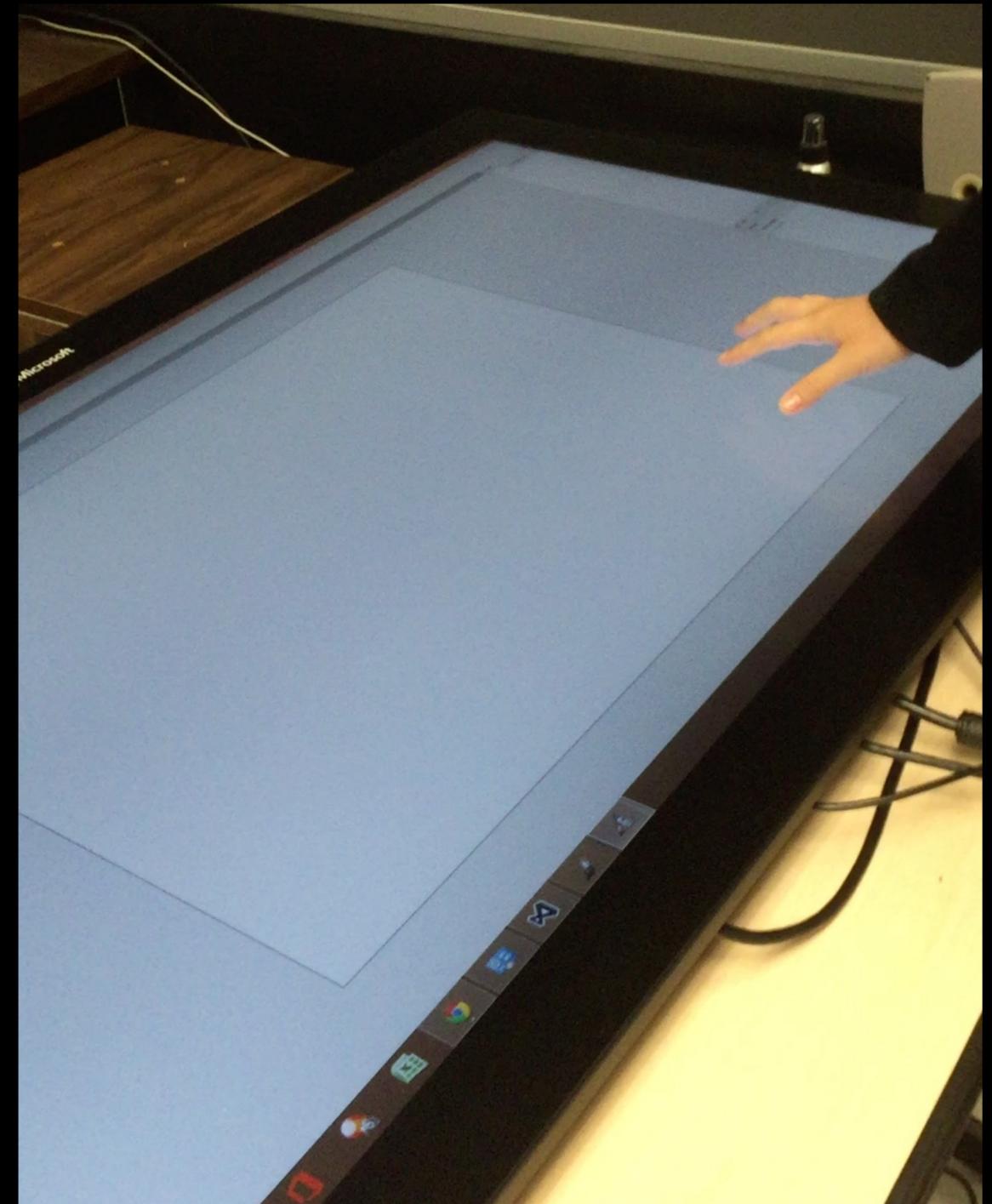
Interaction bimanuelle par accords

Dynachord (Li et al., 2016)

Séquences d'accords donnent
un plus grand vocabulaire d'accords

Ajustement dynamique des paramètres

Versions tablette et table interactive

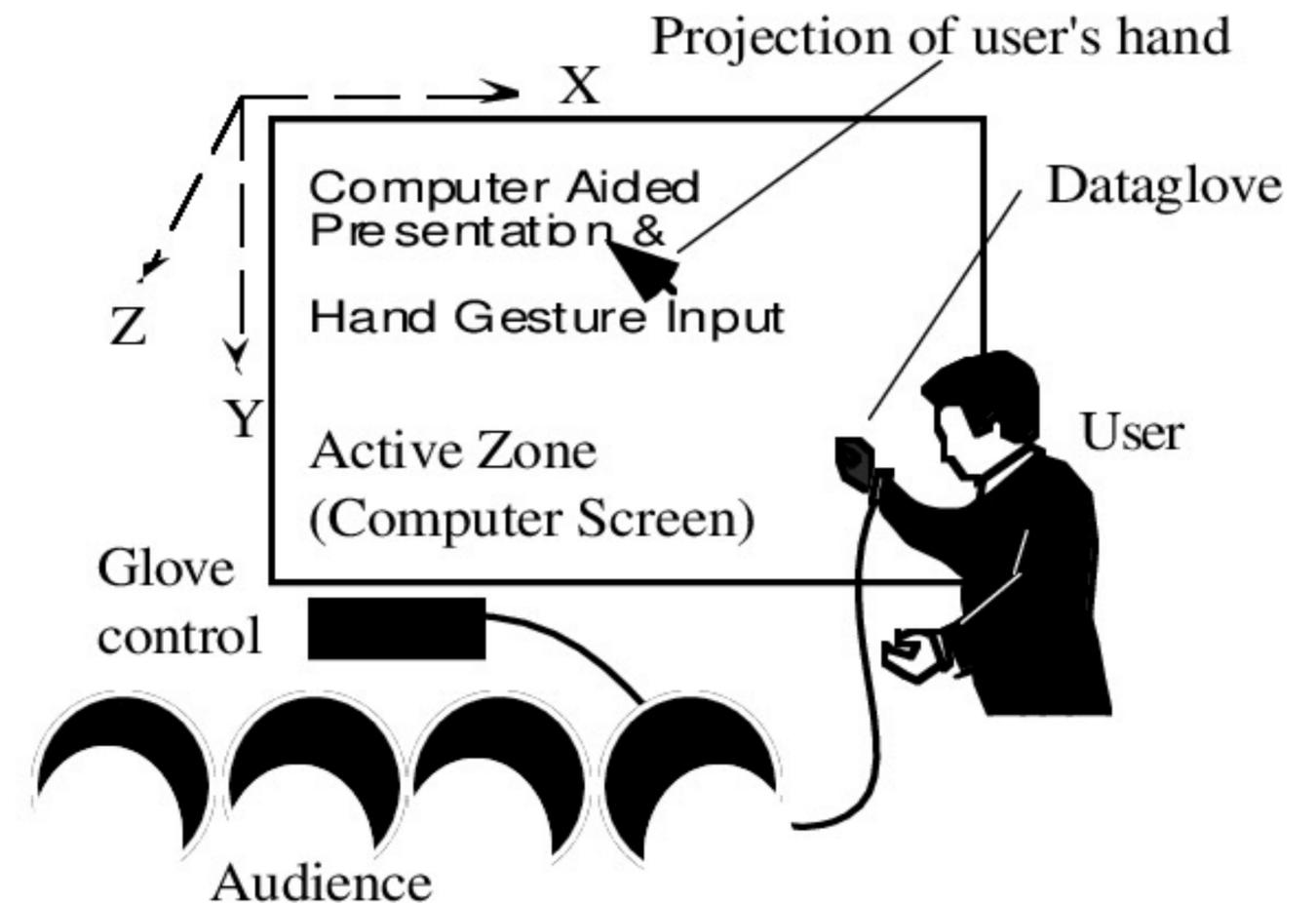


Éntrée gestuelle « mains libres »

Charade (Baudel & Beaudouin-Lafon, 1994)

Différents types de gestes à main levée :
aspect naturel de la parole
gestes délibérés

Contrôler les diapositives d'une présentation



Éntrée gestuelle « mains libres »

Charade (Baudel & Beaudouin-Lafon, 1994)

Différents types de gestes à main levée :
aspect naturel de la parole
gestes délibérés

Contrôler les diapositives d'une présentation :
créer un vocabulaire de gestes de la main
faciles à distinguer des gestes naturels

Gestes vers l'écran et position « tendue »

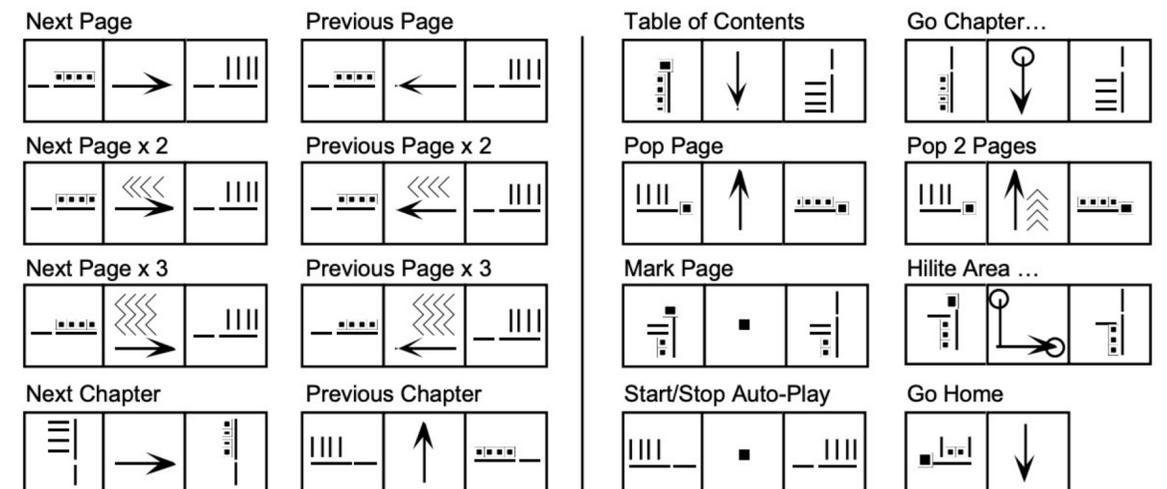
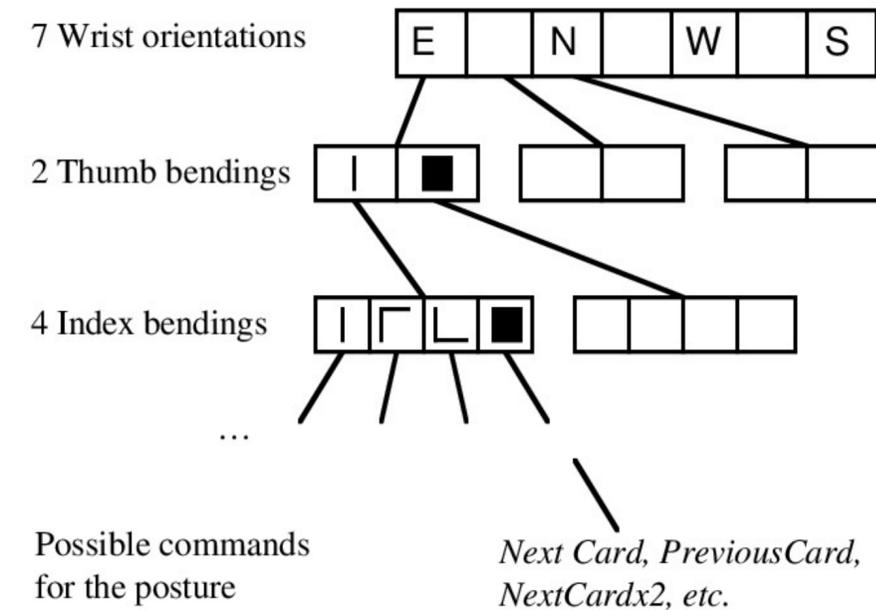
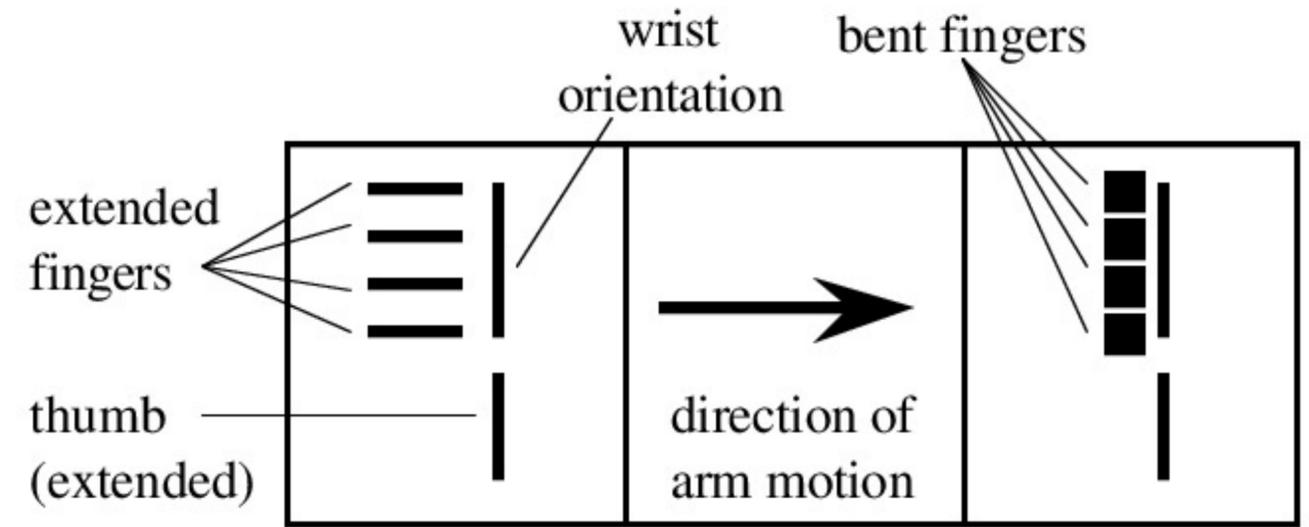


Figure 3 - Gestural command set for the prototype application.

Interaction « mains libres »

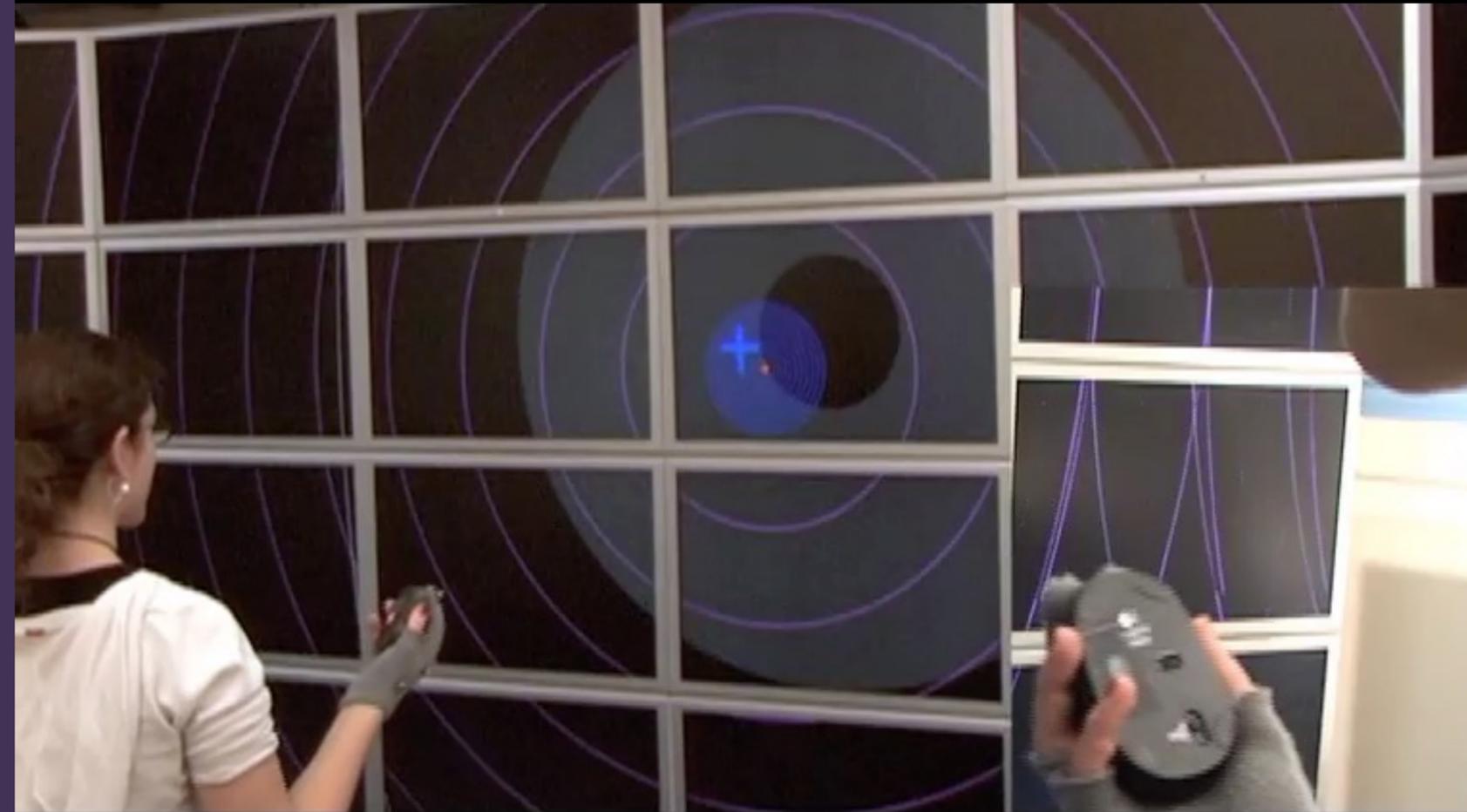
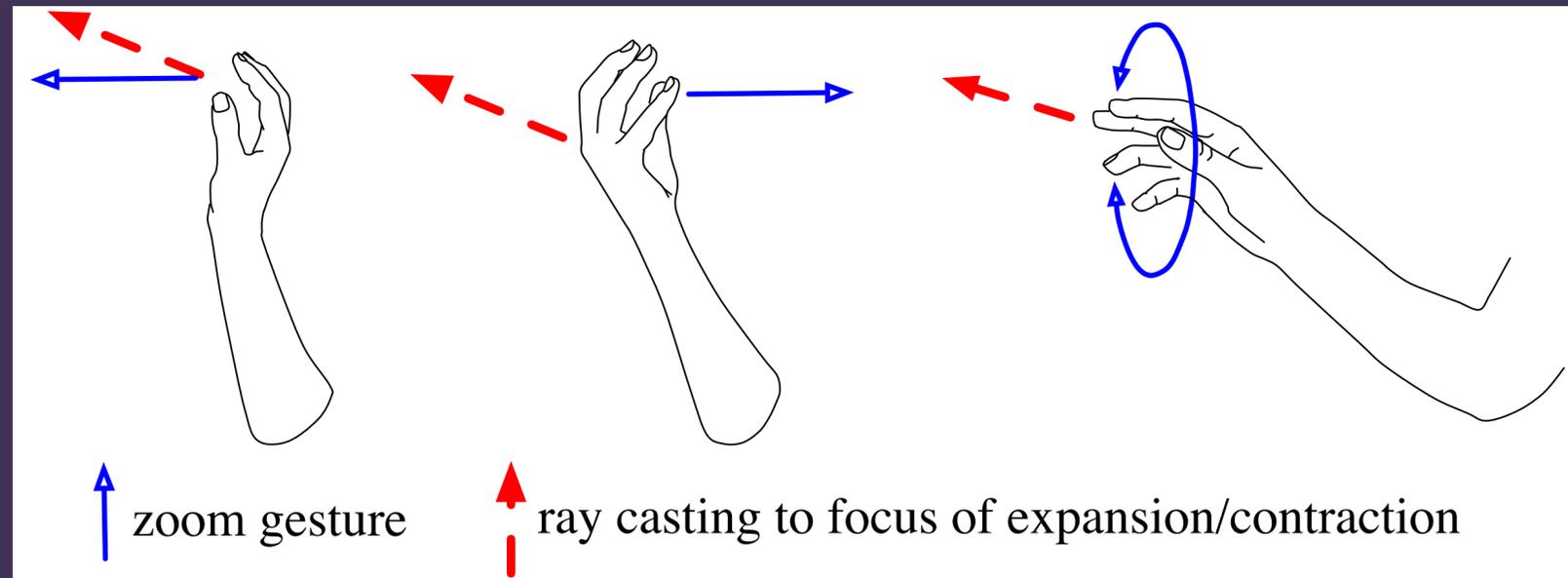
Le rêve de l'interaction avec
les mains libres

Minority Report (2002)



Interaction « mains libres »

La réalité est plus complexe



Unimanual – Linear – 3D Free

Interaction par le corps entier

Video Place (Krueger, 1974)

L'utilisateur interagit
avec l'ombre de son corps

Inspiration pour :

la caméra Eye Toy de Sony (2003)

la Kinect de Microsoft (2009)

Jeux : vecteur d'innovation pour l'interaction
périphériques d'entrée dédiés
utilisation de la 3D



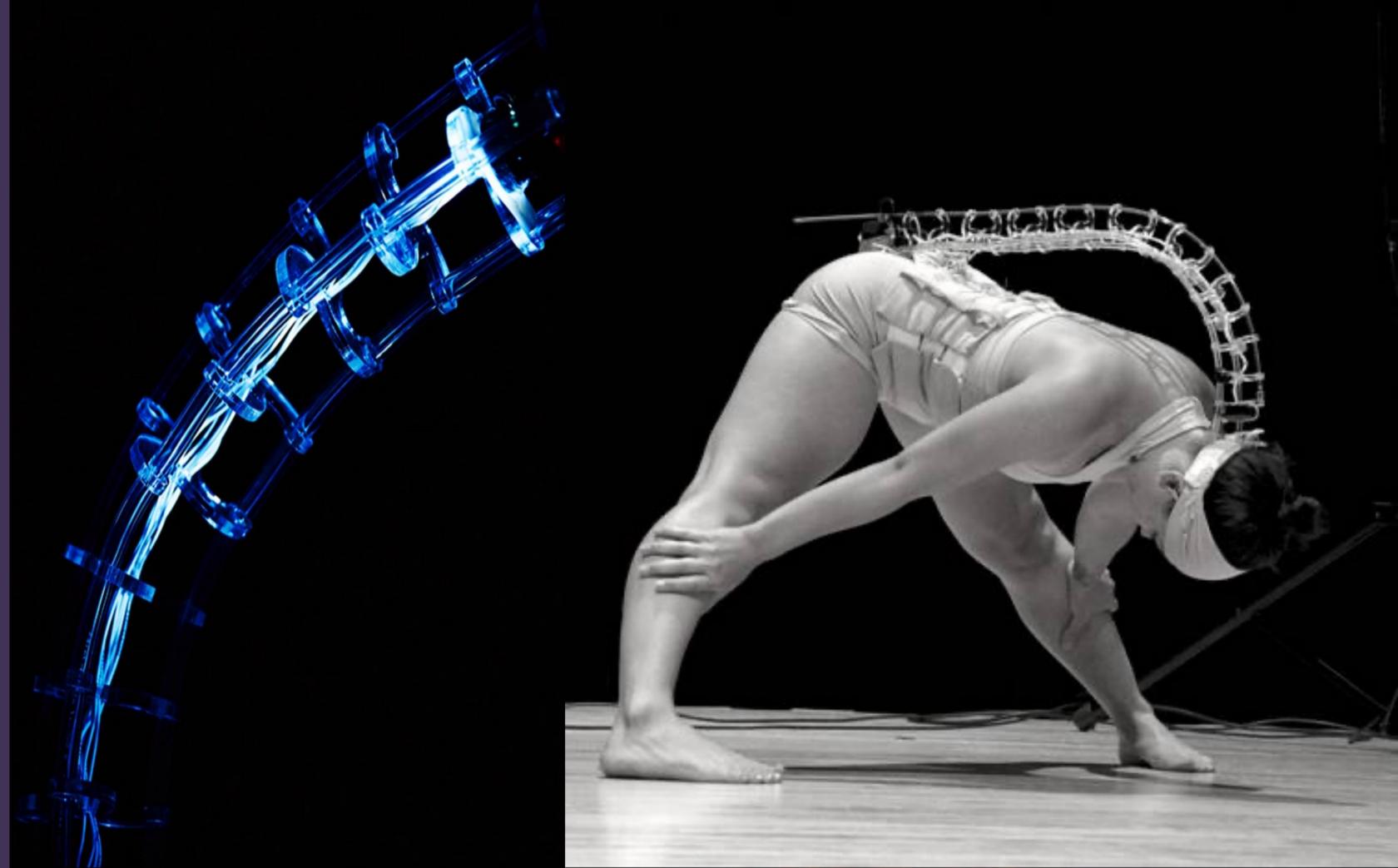


Instruments de musique numériques

Les Gestes (Malloch et al., 2013)

Instruments "prothétiques" qui détecte le
mouvement et le multitouch :
une colonne vertébrale,
une visière, ou une côte

Chaque dispositif produit
neuf paramètres différents
disponibles pour le mappage



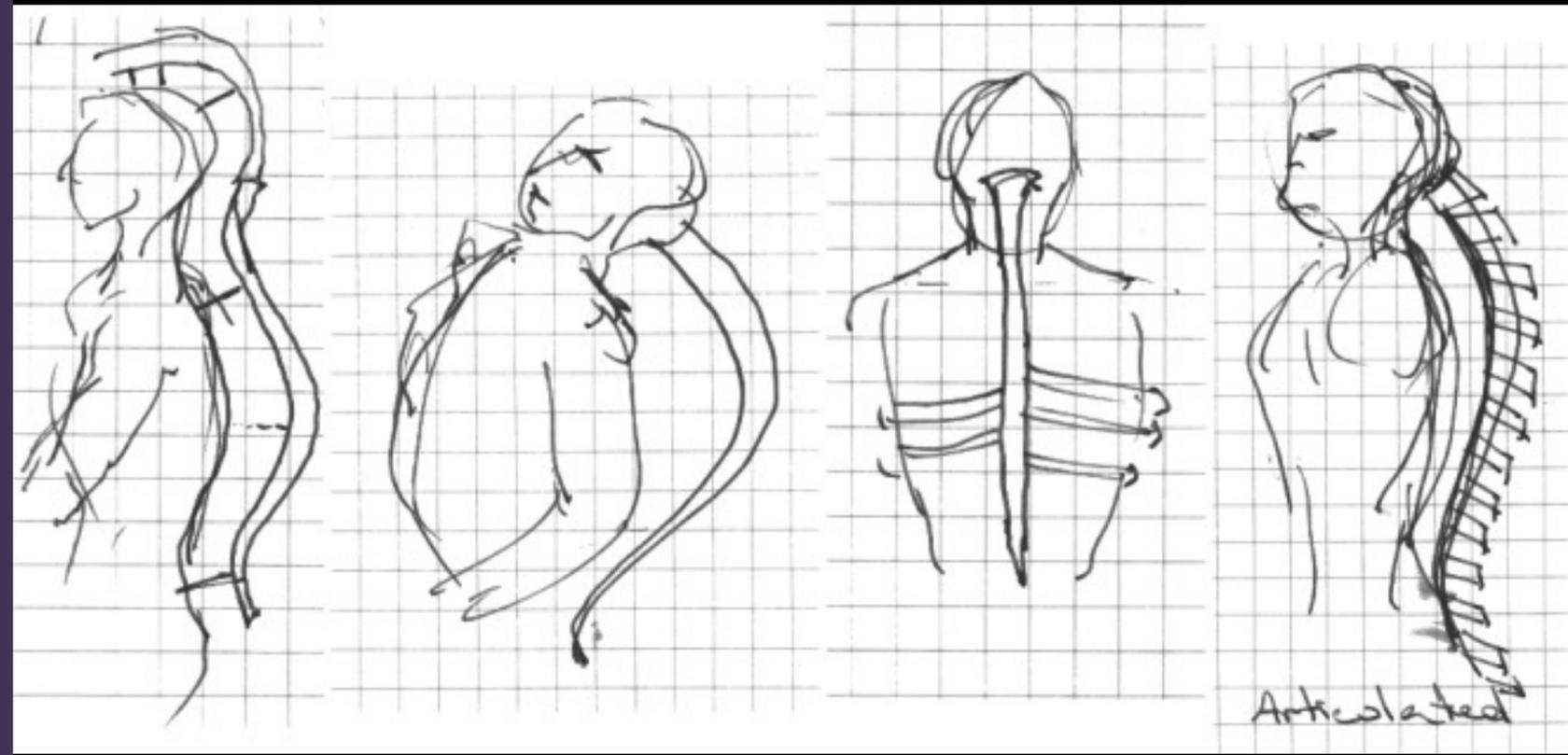
Instruments de musique numériques

Les Gestes (Malloch et al., 2013)

Les compositeurs établissent la
correspondance entre le mouvement,
le geste et le son

Chaque danseur porte un dispositif :
colonne vertébrale, visière ou côte

Interaction comme avec un partenaire
mais aussi création de la musique



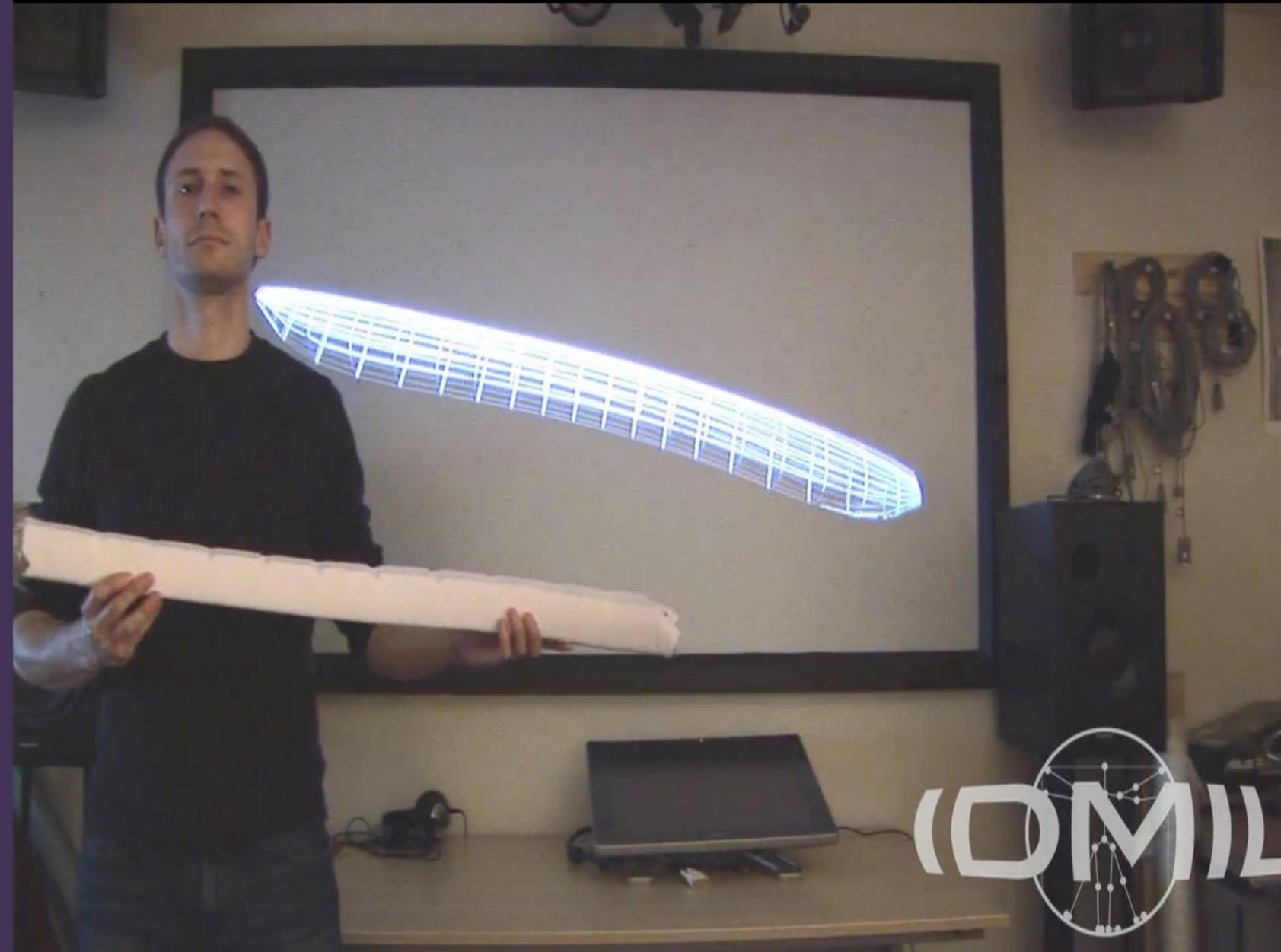
Instruments de musique numériques

Les Gestes (Malloch et al., 2013)

Les compositeurs établissent la
correspondance entre le mouvement,
le geste et le son

Chaque danseur porte un dispositif :
colonne vertébrale, visière ou côte

Interaction comme avec un partenaire
mais aussi création de la musique



Instruments de musique numériques

Les Gestes (Malloch et al., 2013)

Les compositeurs établissent la
correspondance entre le mouvement,
le geste et le son

Chaque danseur porte un dispositif :
colonne vertébrale, visière ou cône

Interaction comme avec un partenaire
mais aussi création de la musique





Interaction haptique

Le sens du toucher humain est hiérarchisé

Capteurs tactiles dans la peau :

chaleur, pression, vibration,
glissement, douleur

Proprioception

position du corps, force, mouvement

La plupart des interfaces haptiques se concentrent sur l'un ou l'autre.

Interaction haptique

Dispositifs haptiques actifs

échanger l'énergie :

force, vibration, chaleur

par contact avec une partie du corps

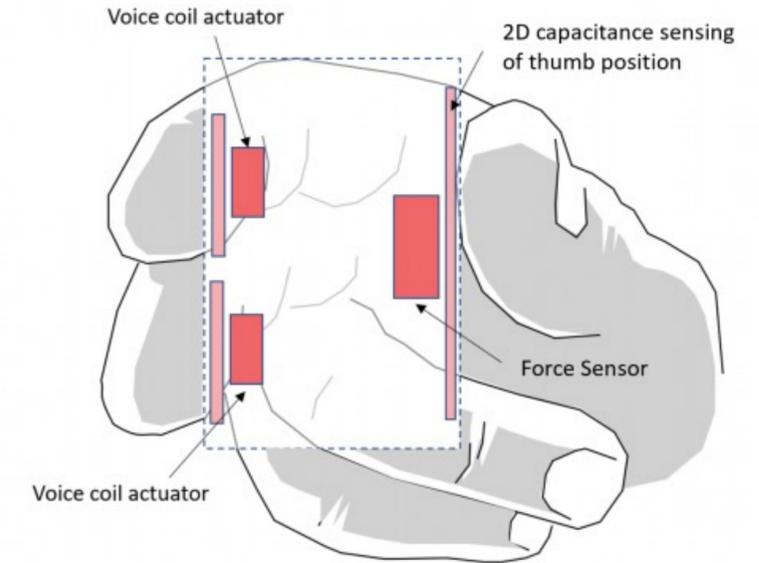
selon un algorithme interactif

Exemple : vibration sur un smartphone

Manettes de jeu à retour de force sont

à la fois un dispositif d'entrée et de sortie

TORC (Microsoft Research)



Interaction haptique

Azmandian et al. (2016)

Reciblage haptique :

Profiter de la perception humaine

utiliser un seul objet

le percevoir comme plusieurs objets

Haptic Retargeting

Dynamic Repurposing of Passive Haptics
for Enhanced Virtual Reality Experiences

Mahdi Azmandian, Mark Hancock
Hrvoje Benko, Eyal Ofek, Andy Wilson
Microsoft Research

SIGCHI 2016

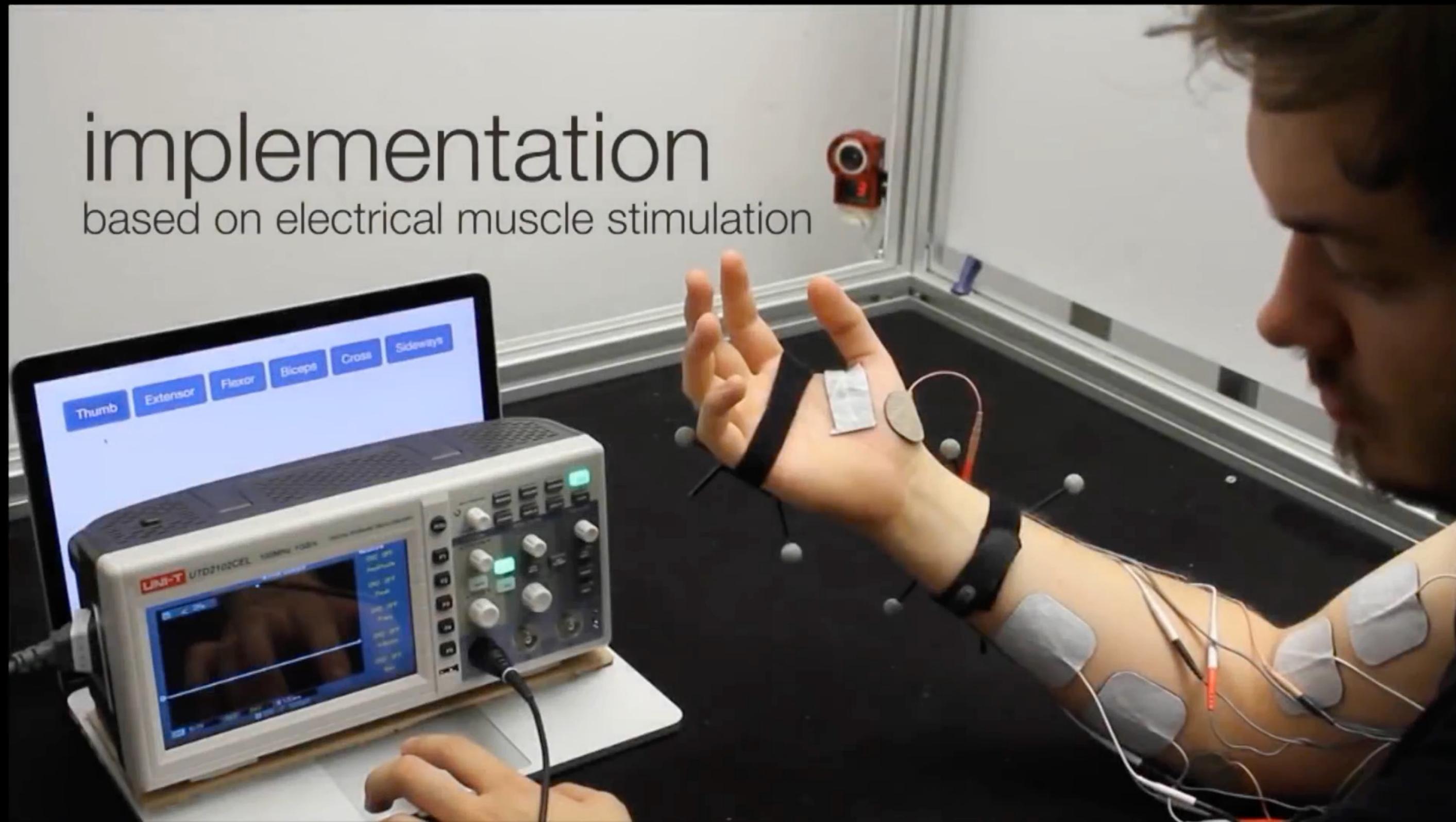
Electro-stimulation musculaire

Canal haptique en sortie

Contrôler les mouvements
en stimulant directement les muscles

implementation

based on electrical muscle stimulation



we asked users to paint this drawing



Interfaces cerveau-ordinateur

Capturer et analyser
des signaux cérébraux
Les traduire en commandes
Les relayer à un dispositif de sortie
Effectuer l'action souhaitée



Interfaces cerveau-ordinateur

Brainball (2007)

Chaque joueur

porte un détecteur des ondes alpha

essaie de déplacer la balle

vers son adversaire

en restant plus calme que lui



Interfaces accessibles

Rendre les interfaces perceptibles
et utilisables par les personnes
ayant des capacités limitées

Nous avons tous
des limitations fonctionnelles :
visuelle, auditive, physique,
vocale, cognitive, neurologique

Contrôle par l'eyetracker Tobii



Beaucoup de progrès
dans le matériel qui
détecte les signaux humains

Beaucoup reste à faire pour créer
des formes d'interaction adaptées
à des contextes divers

Le grand défi :

Créer des interfaces qui
combinent avec succès
plusieurs modalités

Leçons 5 & 6

L'Interaction multimodale :

Comment interagir avec tout le corps

La réalité augmentée et virtuelle :

Comment intégrer l'informatique
avec le monde réel

Leçons 7 & 9

La communication médiatisée :

Comment concevoir les systèmes
collaboratifs

Les partenariats humain-machine :

Comment interagir avec l'intelligence
artificielle



Leçon Cinq

29 mars 2022

Modalités d'interaction et interaction multimodale au-delà de la souris, clavier et écran

*Invitée : Laurence Nigay
Université de Grenoble*