

Interaction gestuelle

Sylvain Malacria

<http://www.malacria.com/>

<mailto:sylvain.malacria@inria.fr>

Diapositives adaptées de Géry Casiez et Thomas Pietrzak

Qui suis-je ?

Chargé de recherche au centre Inria Lille

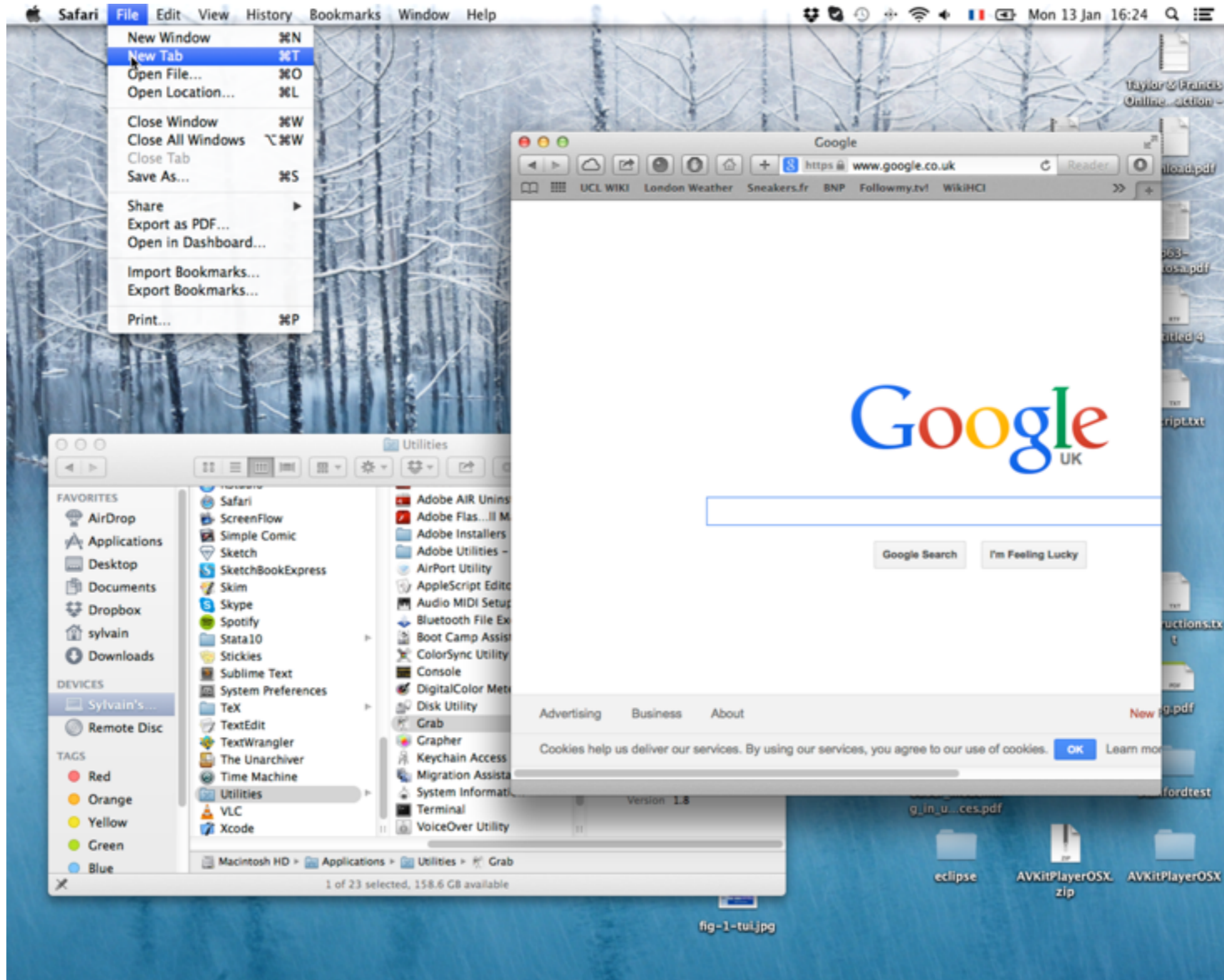
Membre de l'équipe Mjolnir (Inria & Université Lille 1)

Chercheur en Interaction Homme-Machine

Thèmes de recherche :

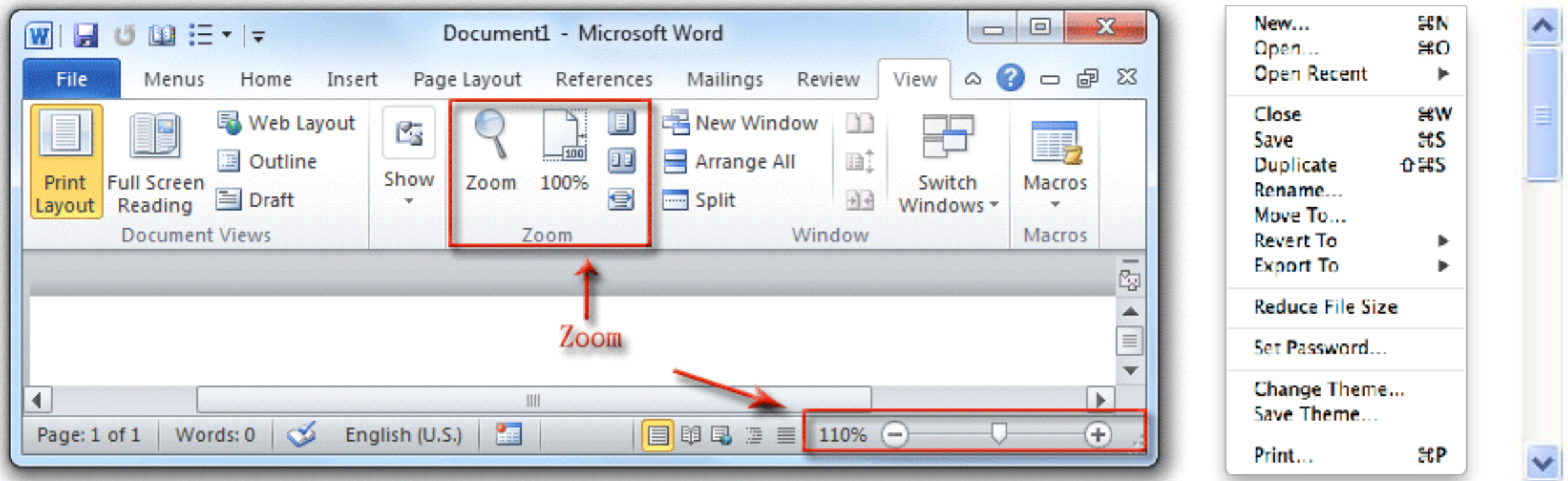
- ▶ Sélection de commande
- ▶ Interaction tactile et gestuelle
- ▶ Technique de navigation
- ▶ Développement de compétences numériques
- ▶ Fonctions de transfert

Windows Icons Menu Pointer



Pas “suffisamment directe”

Interaction *indirecte* via la manipulation d'elements graphiques





Inadaptée



Inadaptée

Interaction gestuelle





00:11:10

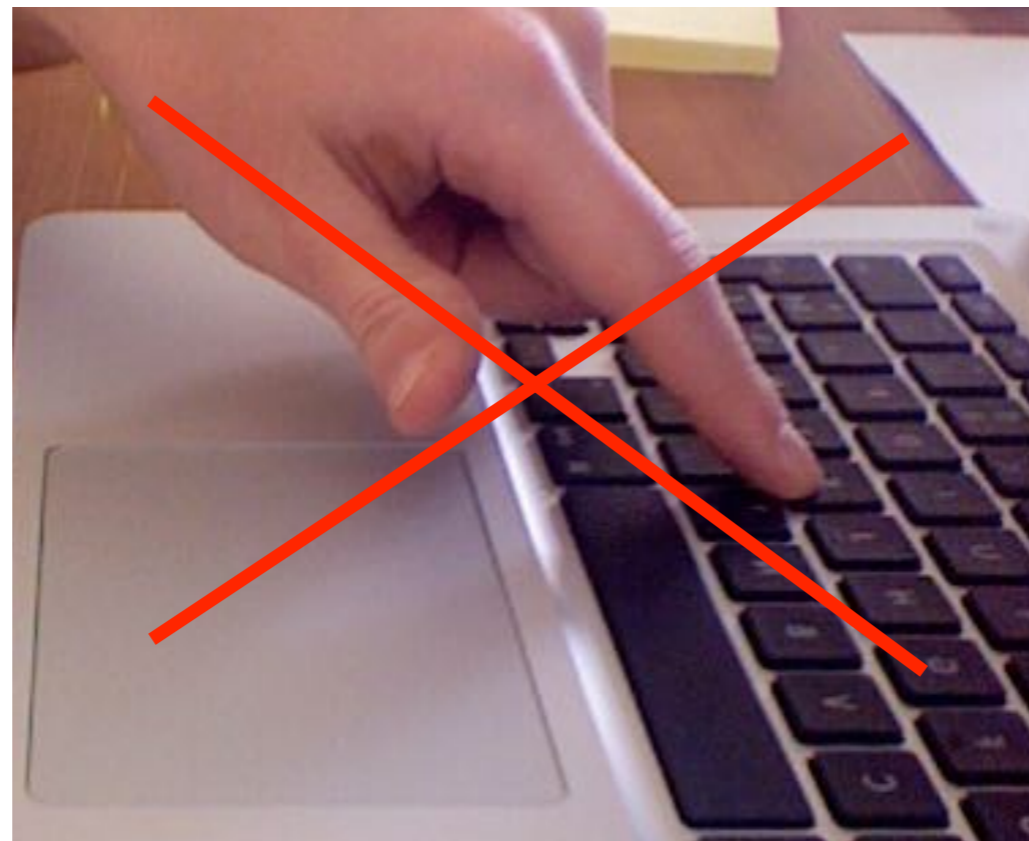
00:16:01

MANAGEMENT
SYSTEMS



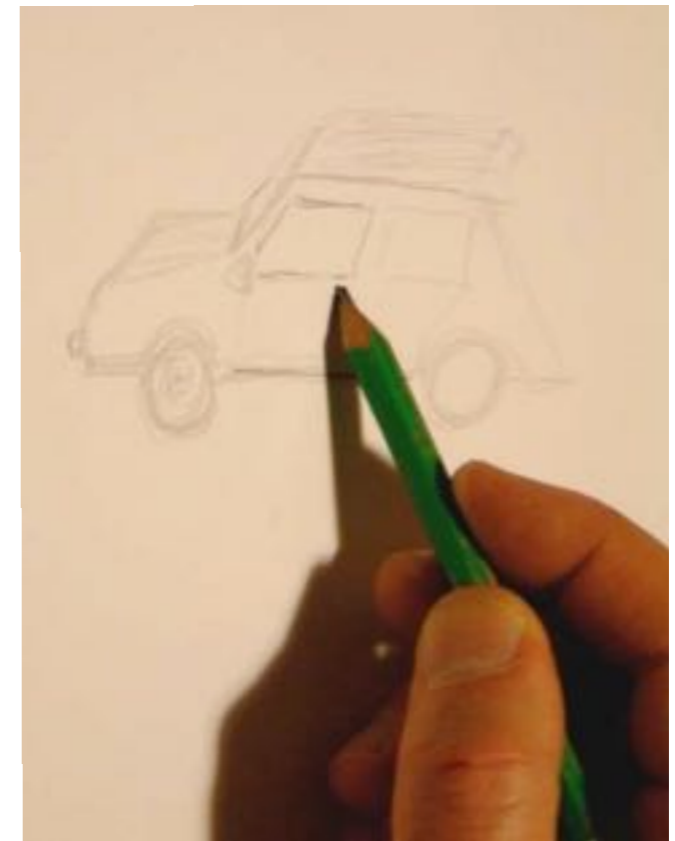
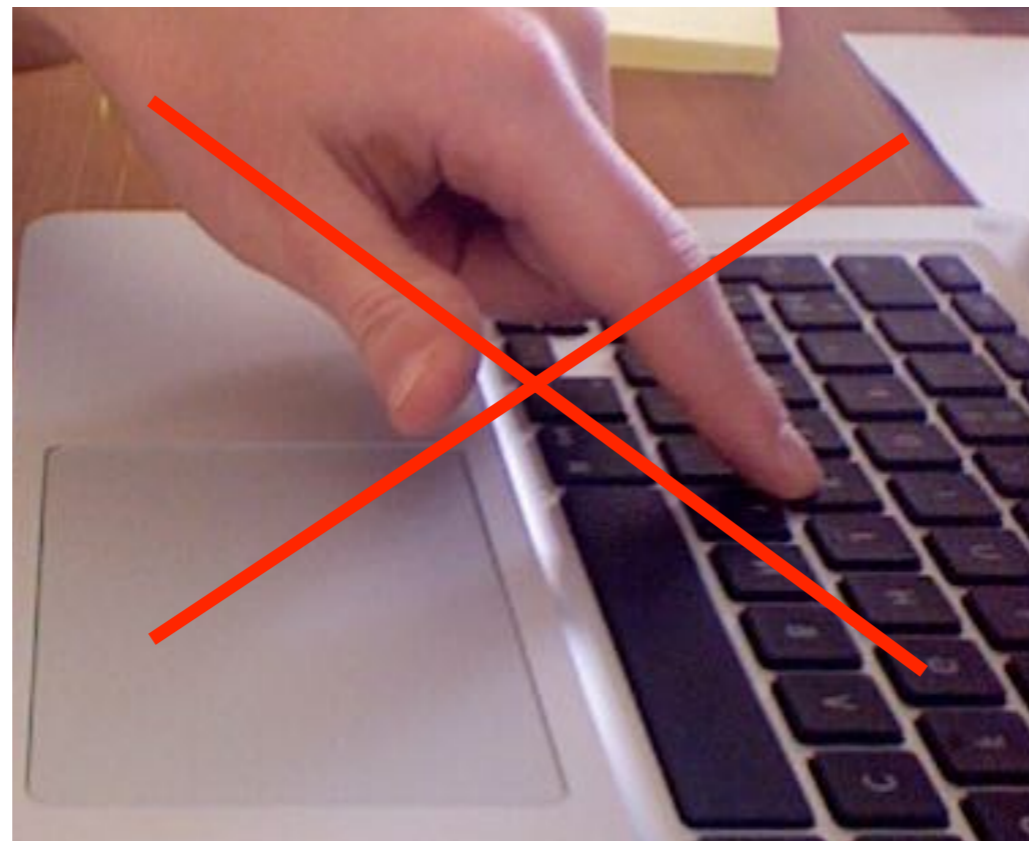
Definition

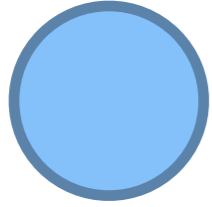
*A gesture is a motion of the body that contains information
[Kurtenbach]*

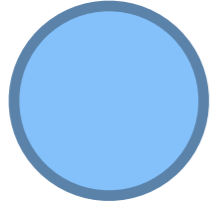


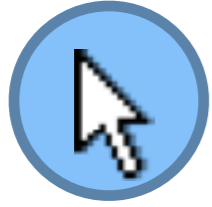
Definition

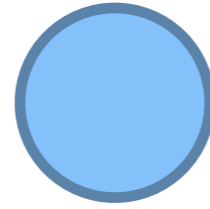
*A gesture is a motion of the body that contains information
[Kurtenbach]*

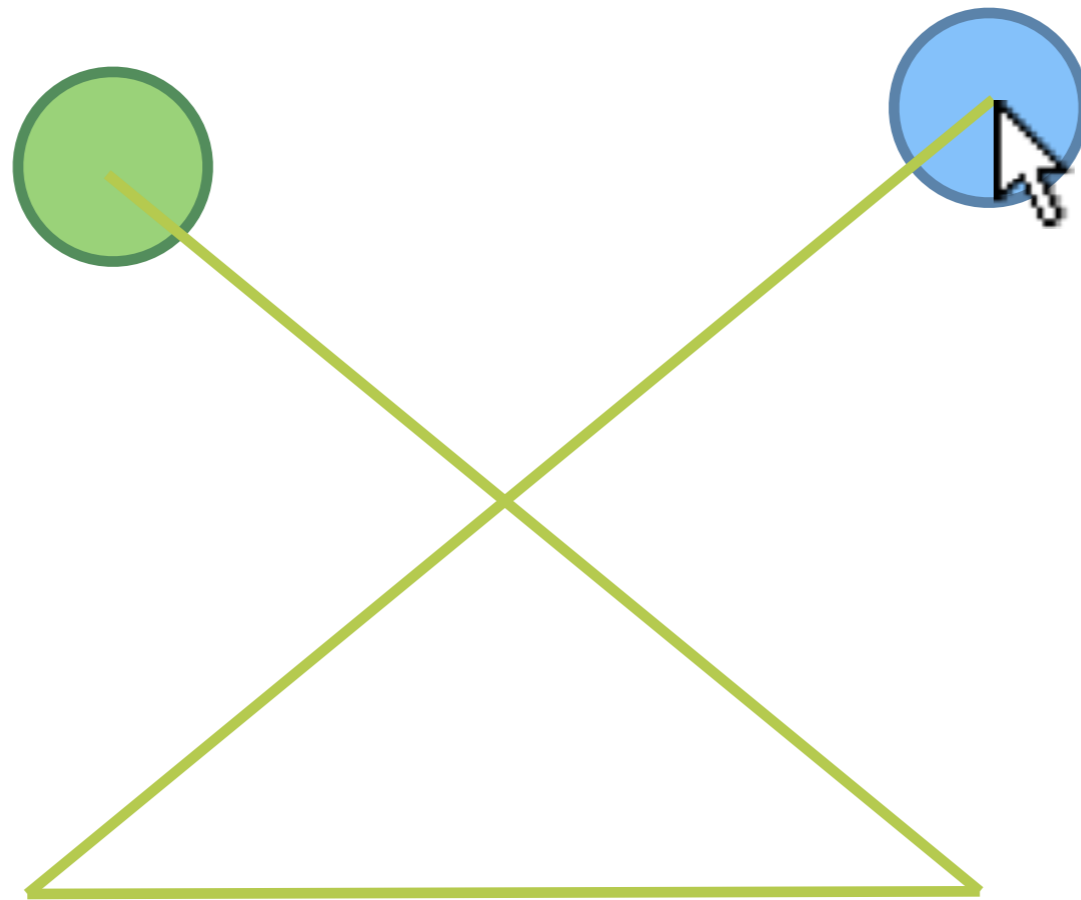




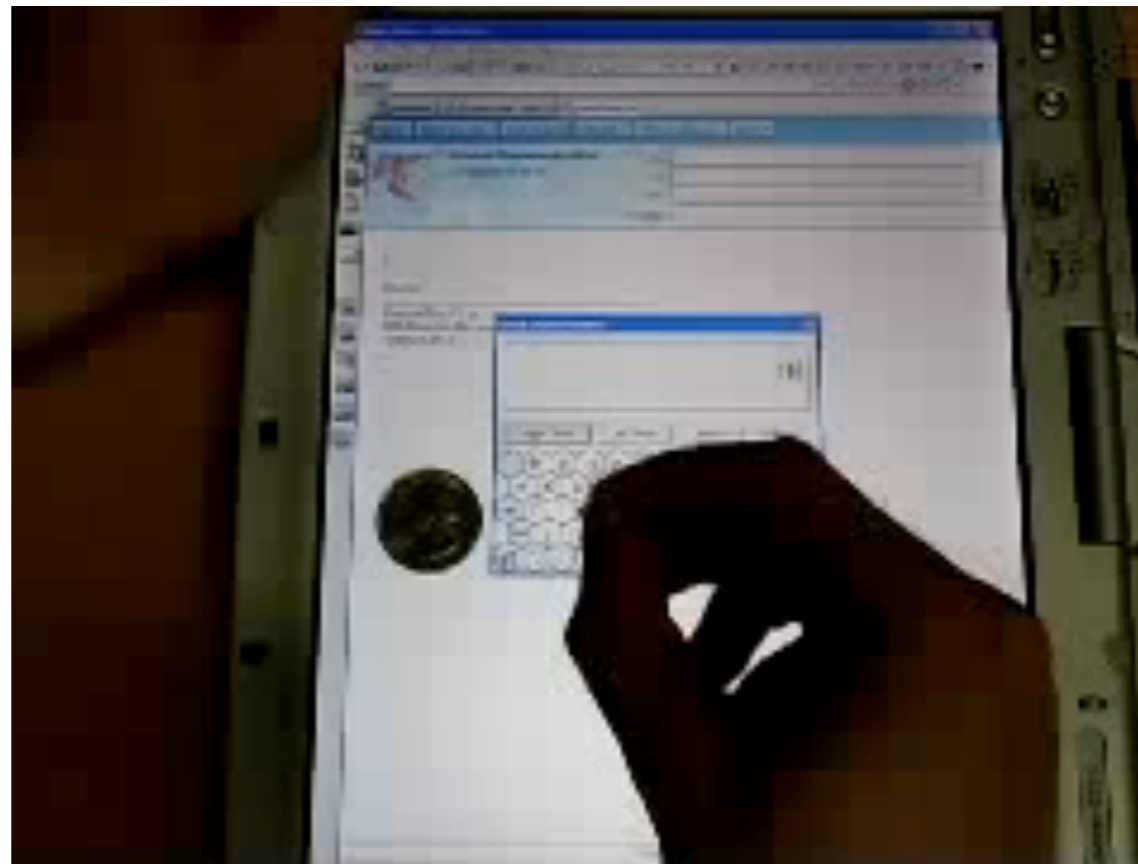




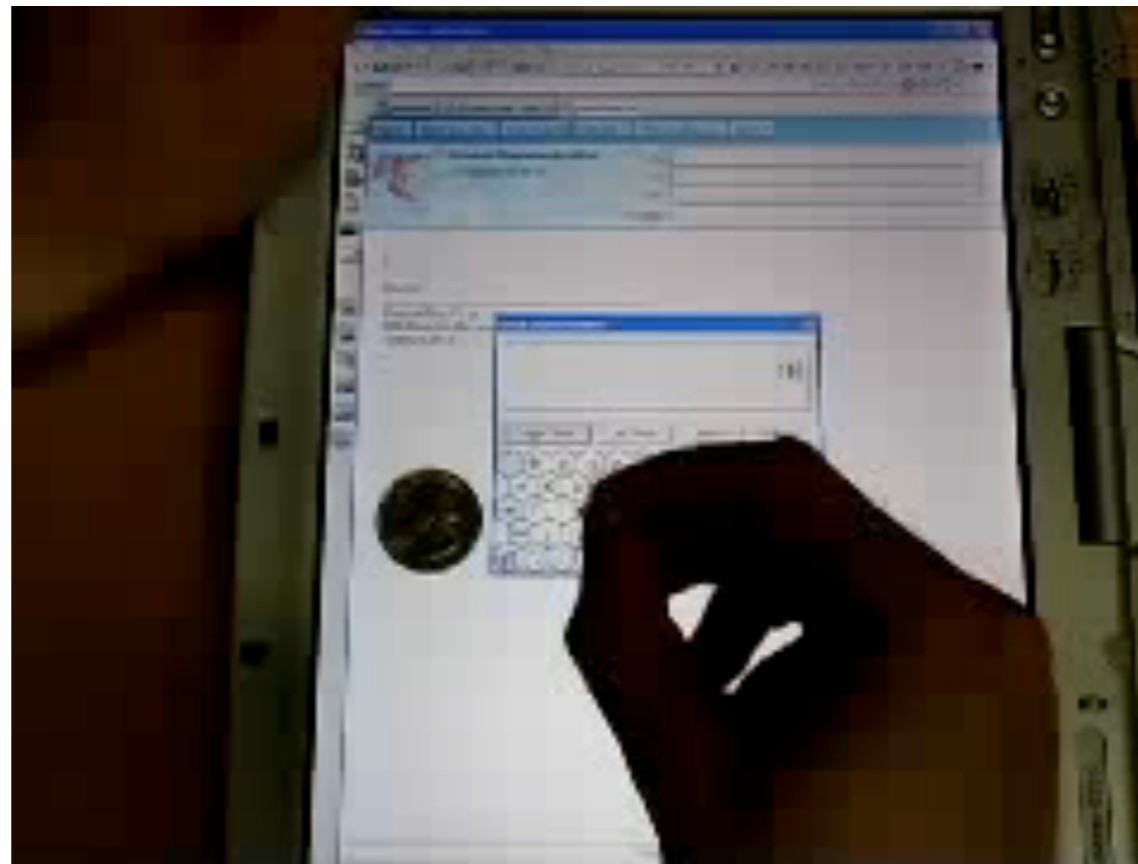




Shark 2004/Shapewriter 2007



Shark 2004/Shapewriter 2007



Pourquoi?

- Capitaliser sur les compétences de l'utilisateur [Kurtenbach et al. 94]
- Combiner la selection de commande et l'operation en une seule action [Buxton et al. 86]
- Apprentissage rapide et implicite
- Très étudié en IHM

Applications (Mouse Gestures)



The image shows a screenshot of the Mozilla Firefox browser window. The title bar reads "Mozilla Firefox Start Page - Mozilla Firefox". The address bar shows "http://www.google.com.my/firefox?client=firefox-a&rls=org.mozilla:en-US:official". The browser interface includes a menu bar (File, Edit, View, History, Bookmarks, Tools, Help), a toolbar with navigation buttons, and a bookmarks bar with folders like "TMP", "Useful", "BBL", "UTAR", "Tech", "News", "mySeng@", "email", "Friends", "BT", "Forums", "m0n0", "Blogs", "myBLOG", "localhost", "myfyp", "java", "JSEffect", and "MLP".

The main content area features a "Firefox Start" banner with the Firefox logo. Below the banner is a Google search interface with the text "Google™ Web Images Groups News". The search bar is empty, and there are radio buttons for "the web" (selected) and "pages from Malaysia". A "Google Search" button is present. To the right of the search bar are links for "Advanced Search" and "Preferences".

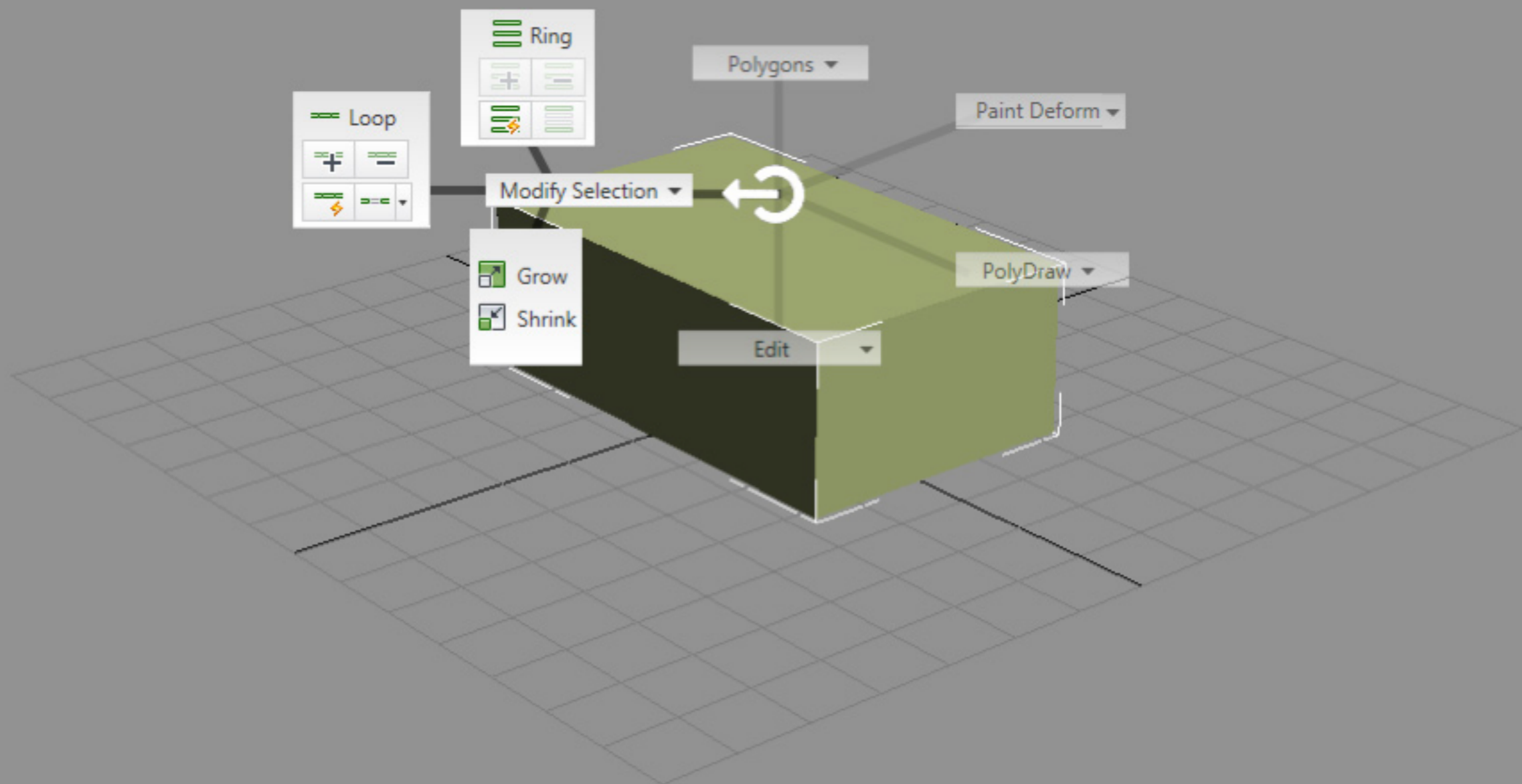
Below the search interface, there is a lightbulb icon and the text: "Thanks for choosing Firefox, the easy-to-use Web browser from Mozilla. [Find out more about us.](#)". At the bottom center, there is a link for "[About Mozilla](#)".

An orange hand-drawn line starts from the search bar area and points towards the "Find out more about us." link, indicating a mouse gesture.

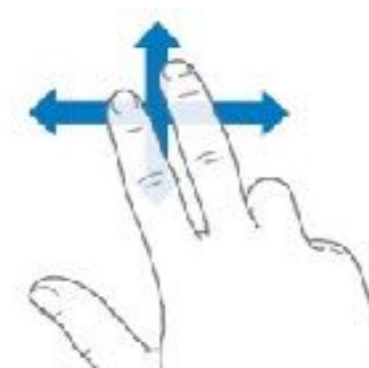
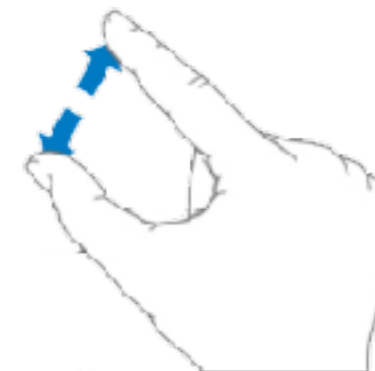


Applications (Mouse Gestures)

[+] [Perspective] [Smooth + Highlights]



Smartphone



Smartphone



Tabletop



Gaming



Gaming

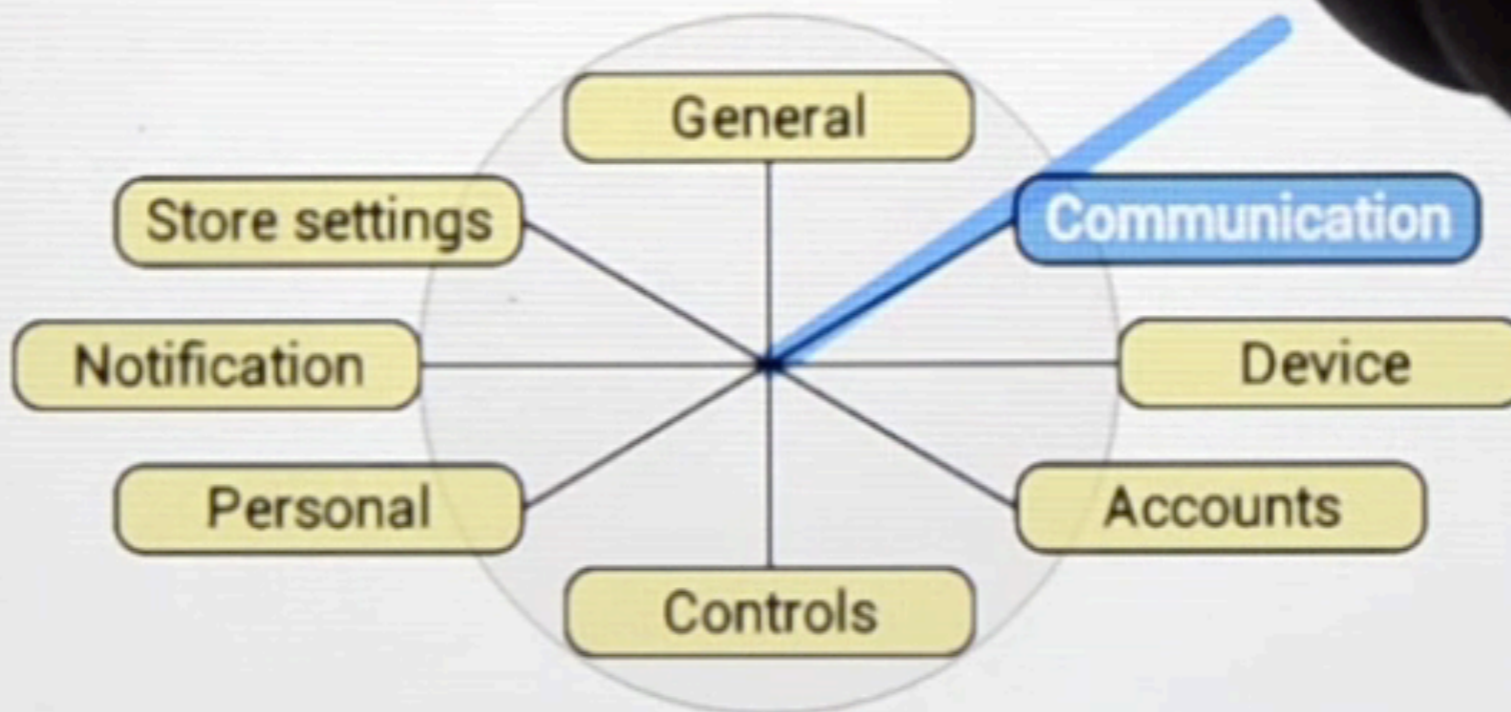


KINECT
SPORTS



Movies





Turn off

Turn off

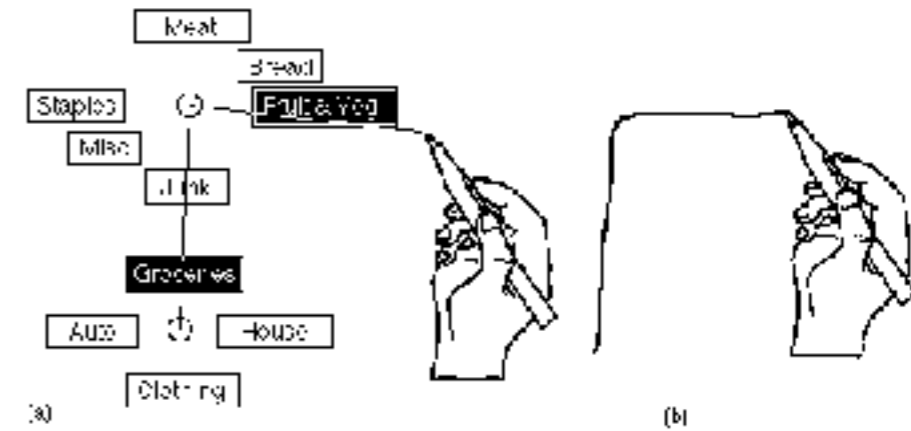
Brainstorming (2 minutes) Qualités & Défauts & Propriétés (pourquoi?)

Vitesse & Precision

Apprentissage & mémorisation

Satisfaction

Autres ?



Vitesse & Precision

- + Layout circulaire (réduit la distance moyenne)
- + Indépendant à la distance (tolérance à l'exécution)
- + Mode expert (Raccourcis)

Apprentissage & Memorisation

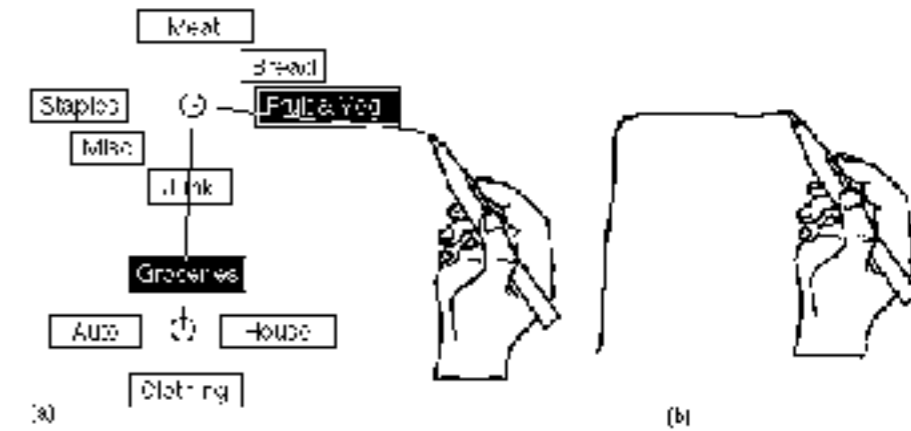
- + Mémoire spatiale (orientation)
- + Mémoire musculaire (transition fluide)
- + Association sémantique possible (ouvrir / fermer)

Satisfaction

- + Interaction gestuelle "appréciée"

Autres

- + En place
- + *Eyes-free* selection (selection possible sans regarder)
- **Nombre de commandes (profondeur :3 / largeur : 8)**
- **Espace requis**

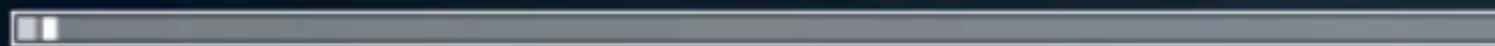


Demo of Marking Menus Versus Linear Menus

Autodesk Research



00:00

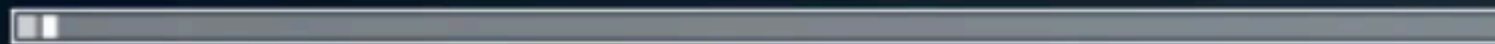


Demo of Marking Menus Versus Linear Menus

Autodesk Research



00:00

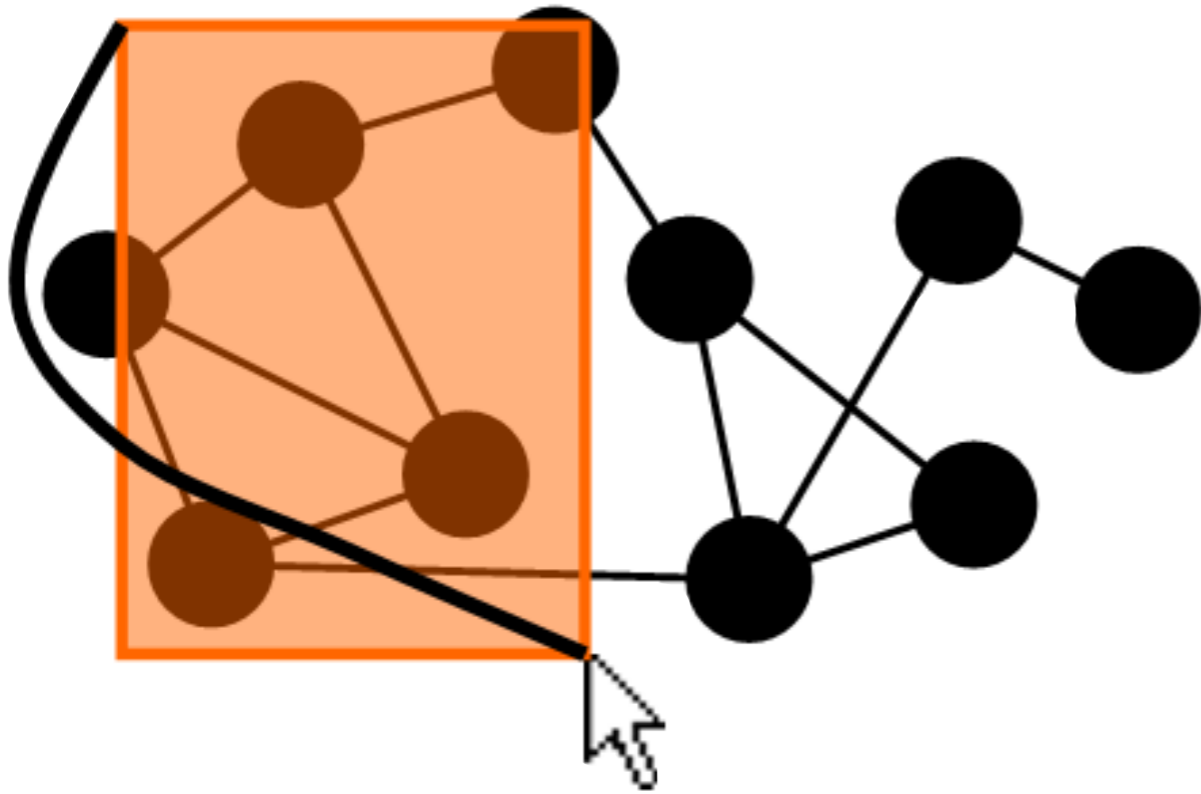




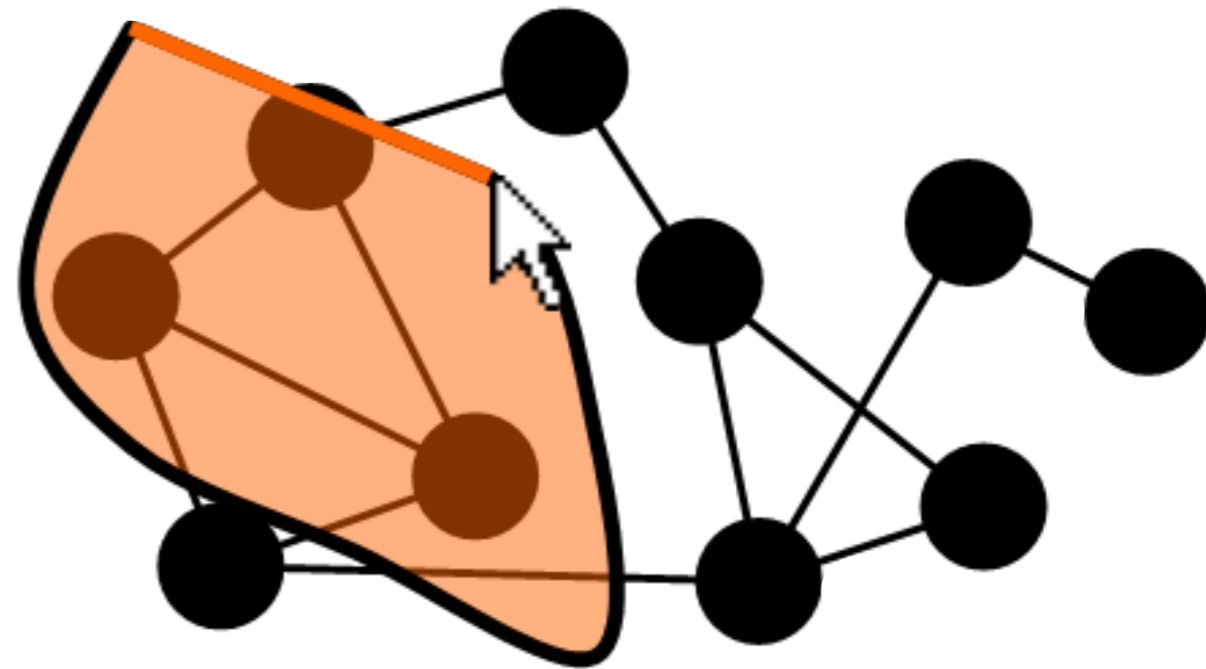
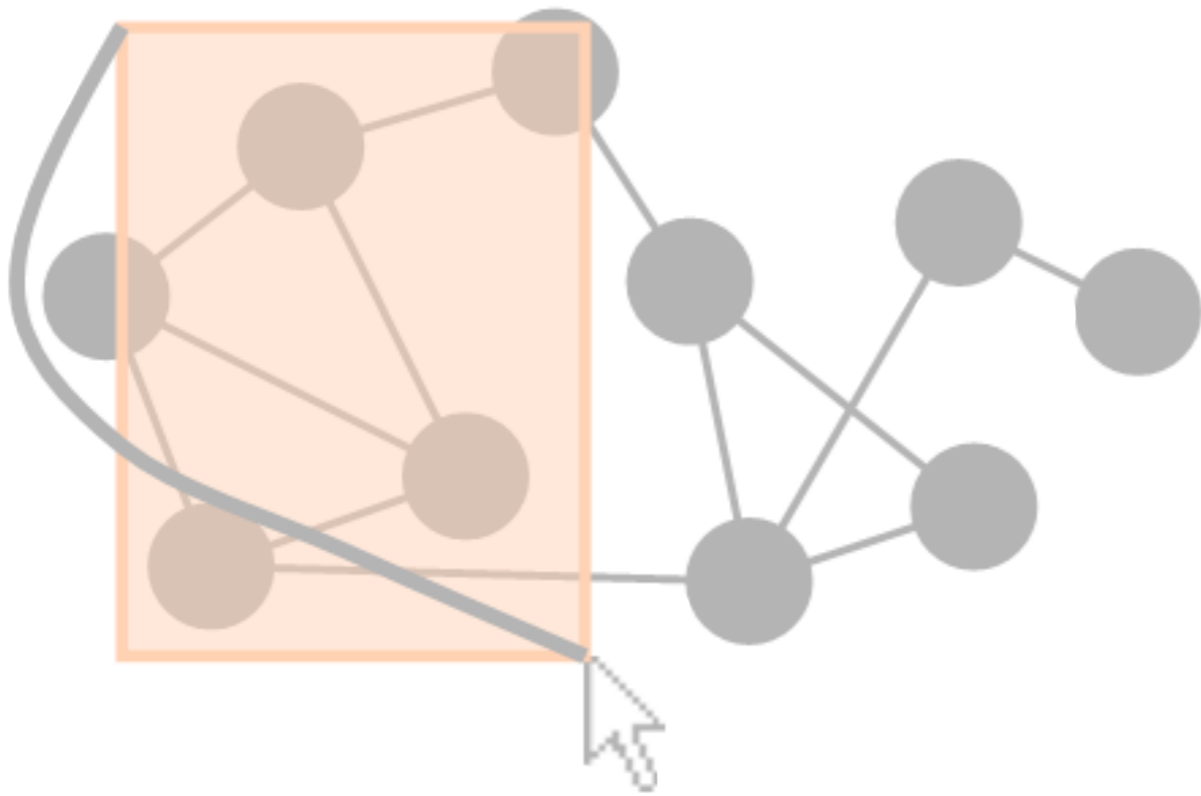
Entrée de texte



Selection de commande

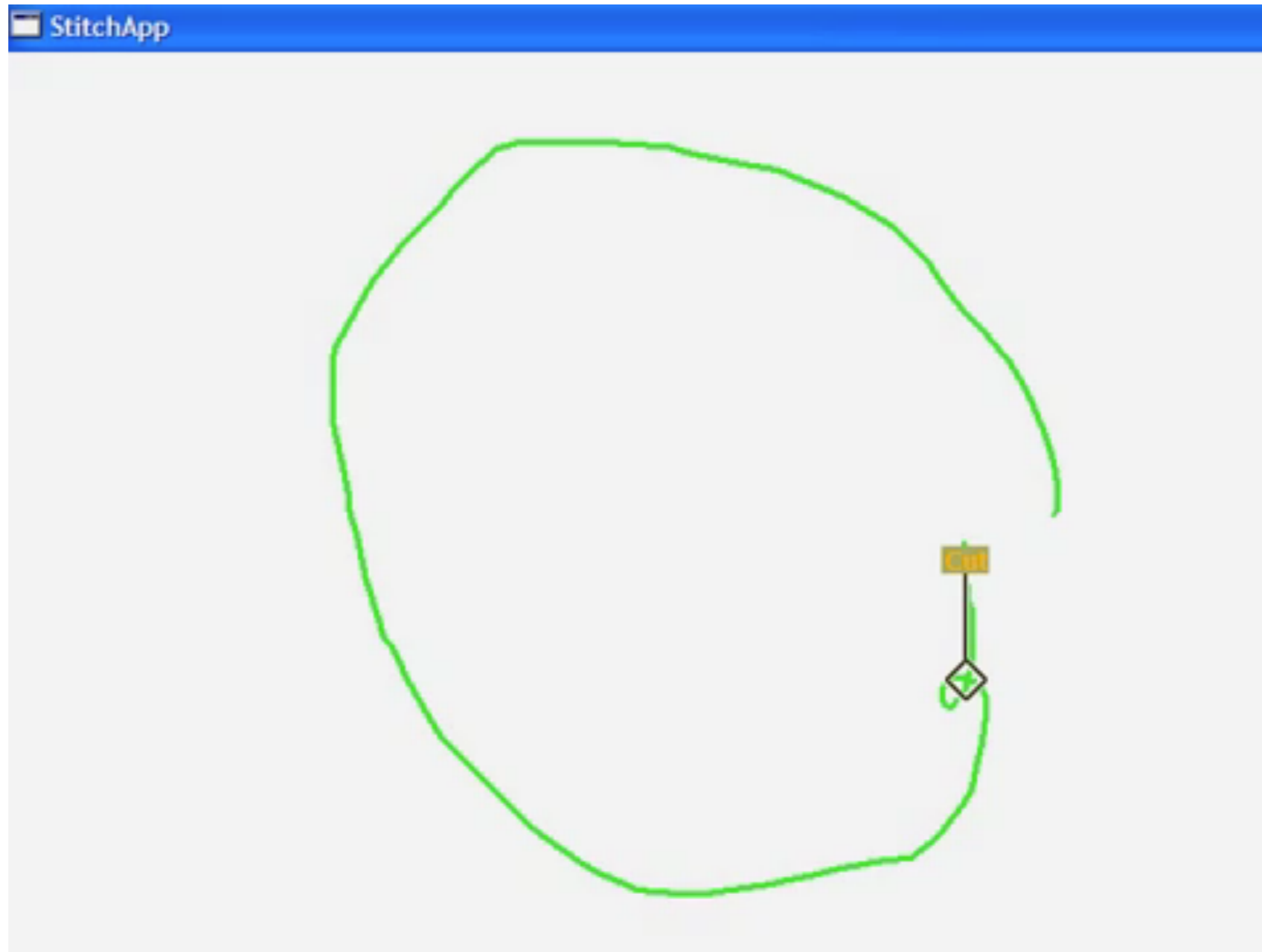


Sélection d'objets

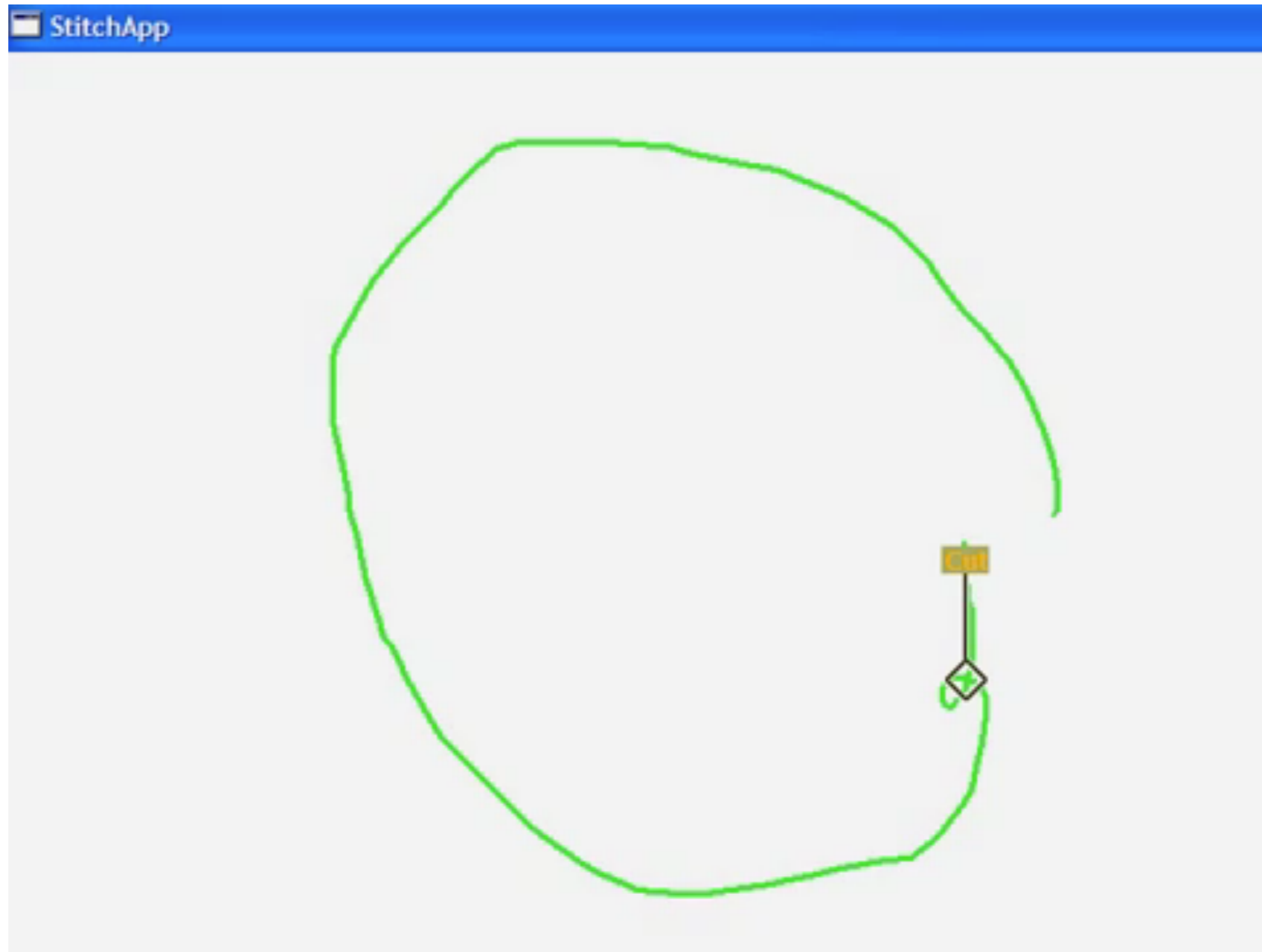


Sélection d'objets

Selection+commande en un seul geste

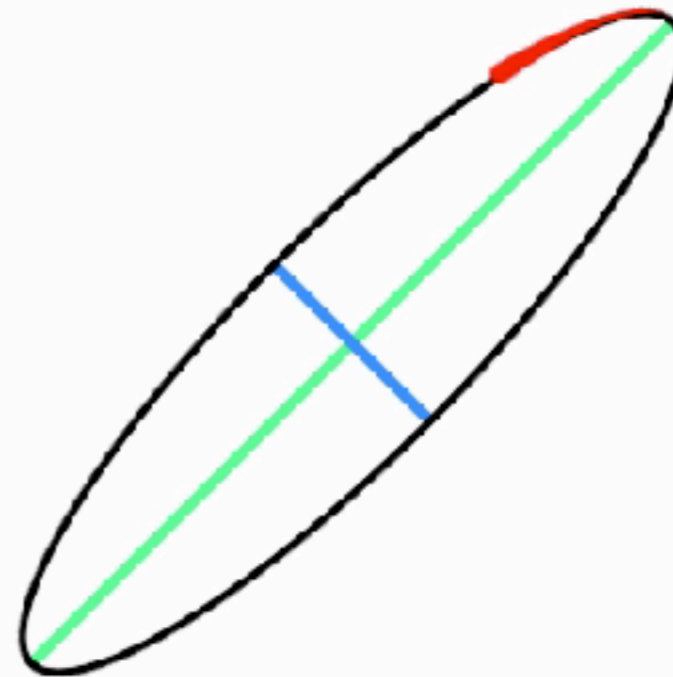


Selection+commande en un seul geste



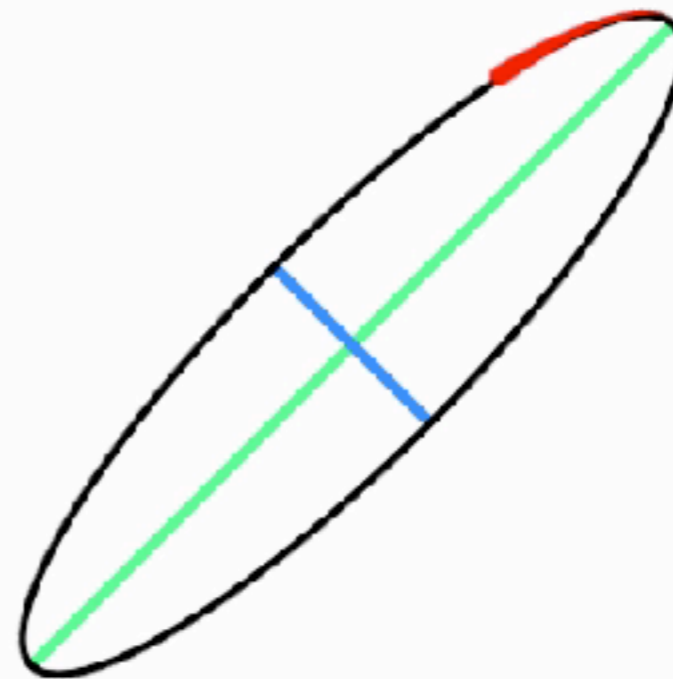
Oscillation elliptiques

6+1 dimensions



Oscillation elliptiques

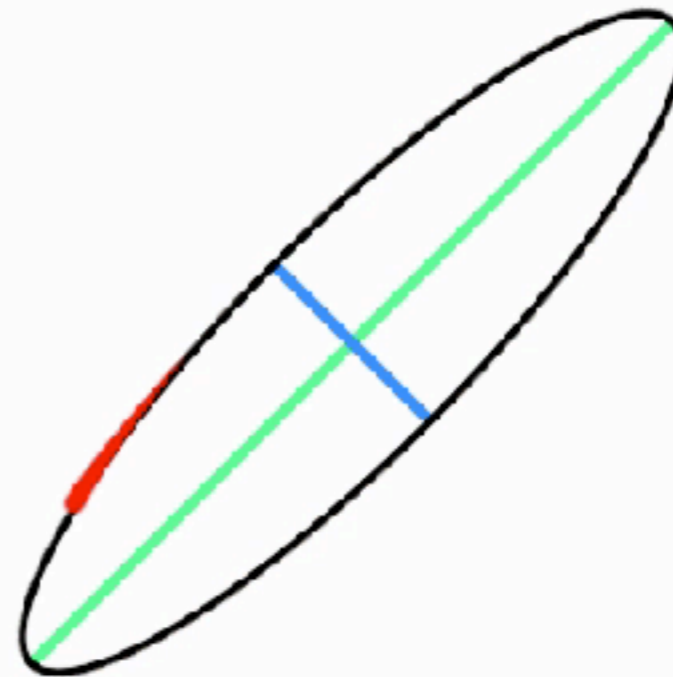
6+1 dimensions



Elliptic oscillatory gestures

6+1 dimensions

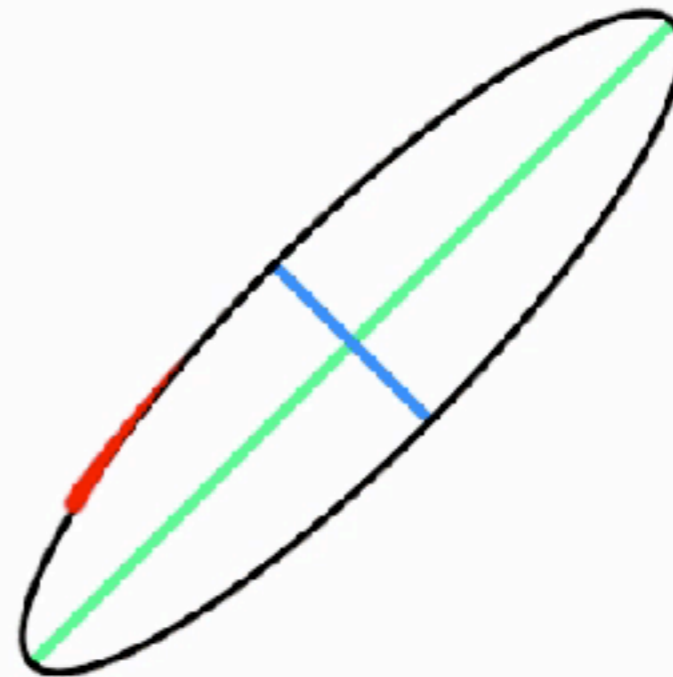
- Orientation



Elliptic oscillatory gestures

6+1 dimensions

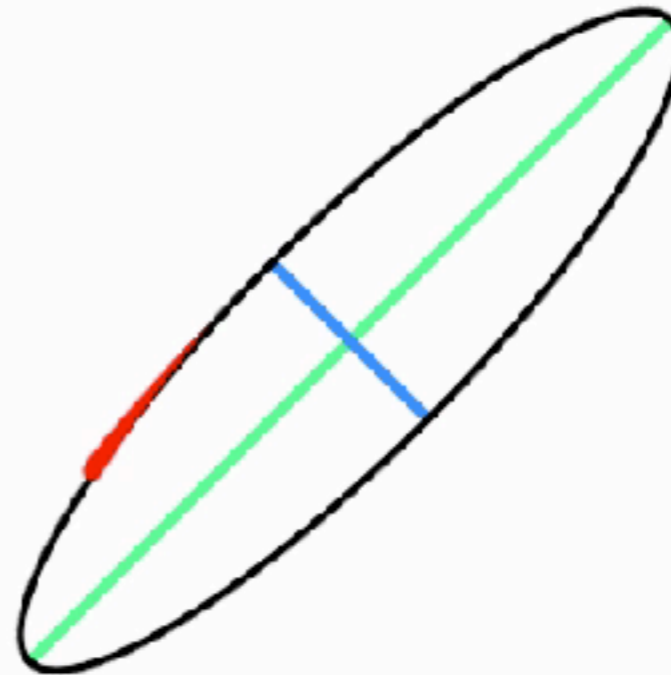
- Orientation



Elliptic oscillatory gestures

6+1 dimensions

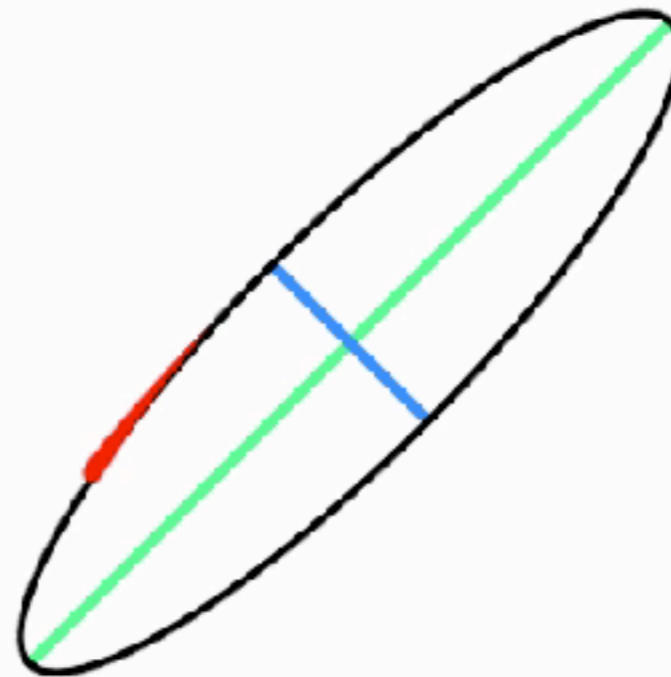
- Orientation
- Taille



Elliptic oscillatory gestures

6+1 dimensions

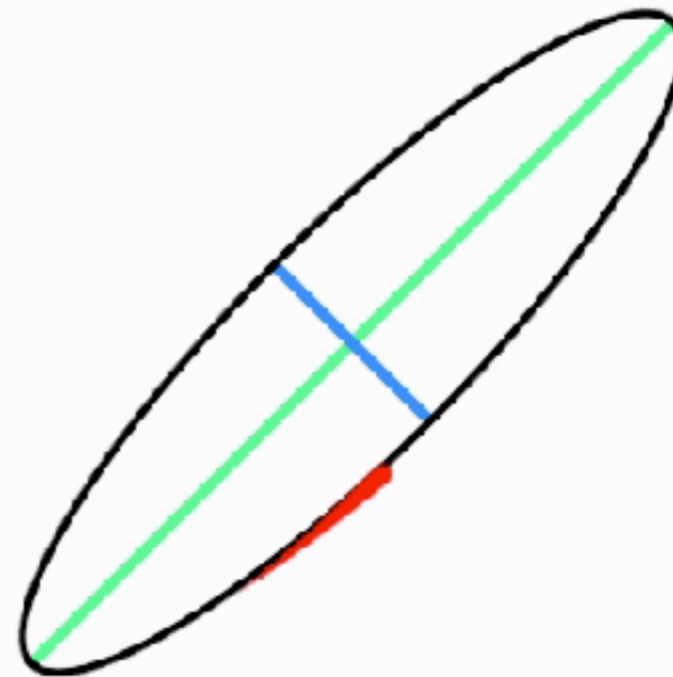
- Orientation
- Taille



Elliptic oscillatory gestures

6+1 dimensions

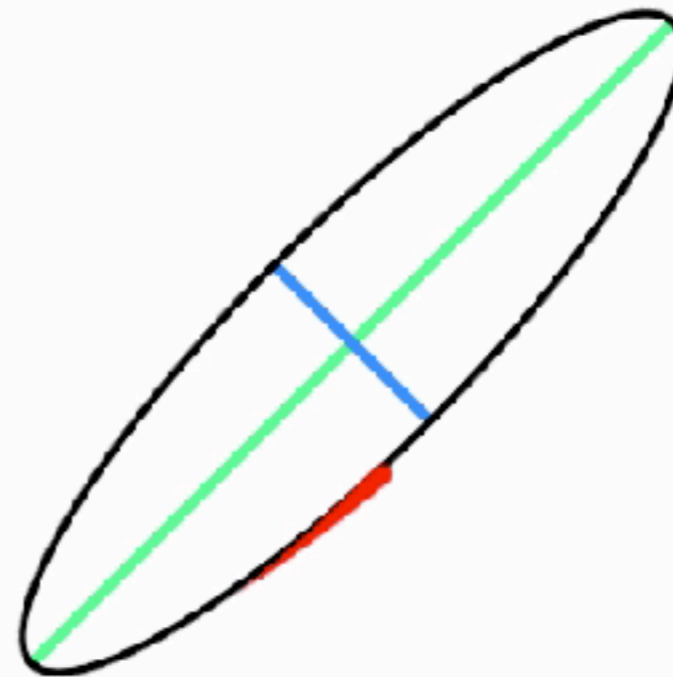
- Orientation
- Taille
- Excentricité



Elliptic oscillatory gestures

6+1 dimensions

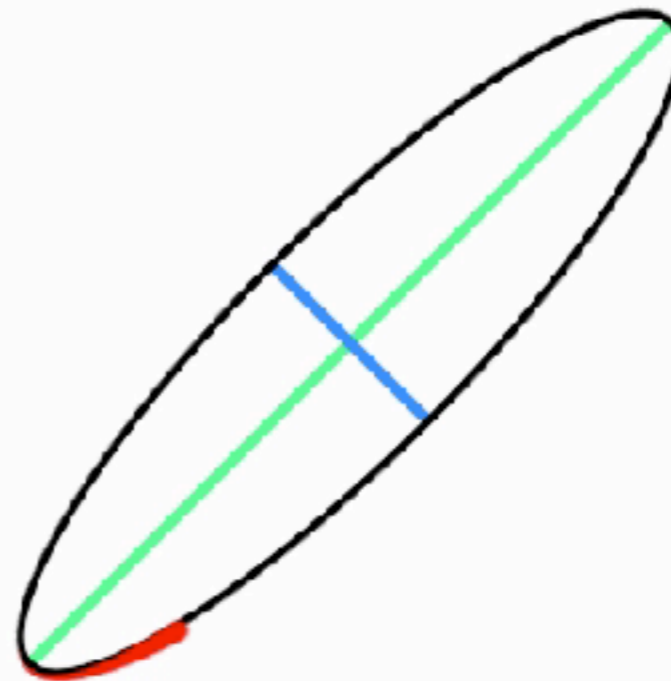
- Orientation
- Taille
- Excentricité



Elliptic oscillatory gestures

6+1 dimensions

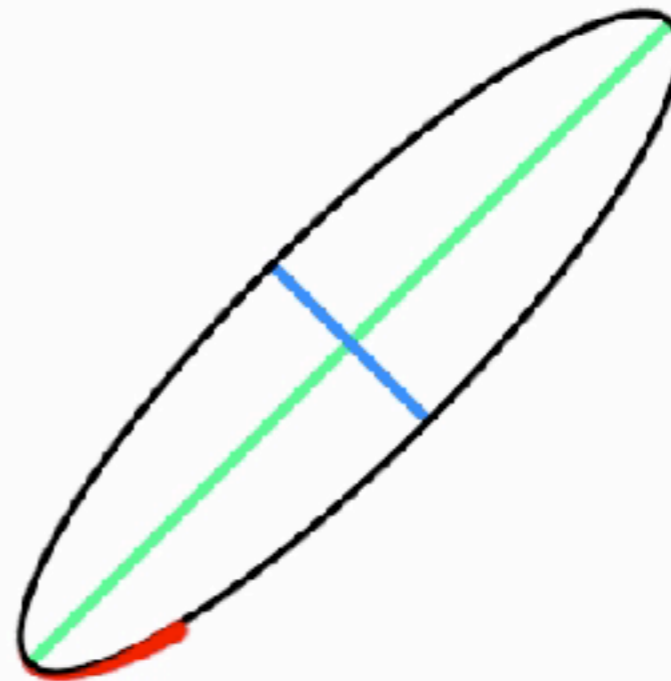
- Orientation
- Taille
- Excentricité
- **Position X**



Elliptic oscillatory gestures

6+1 dimensions

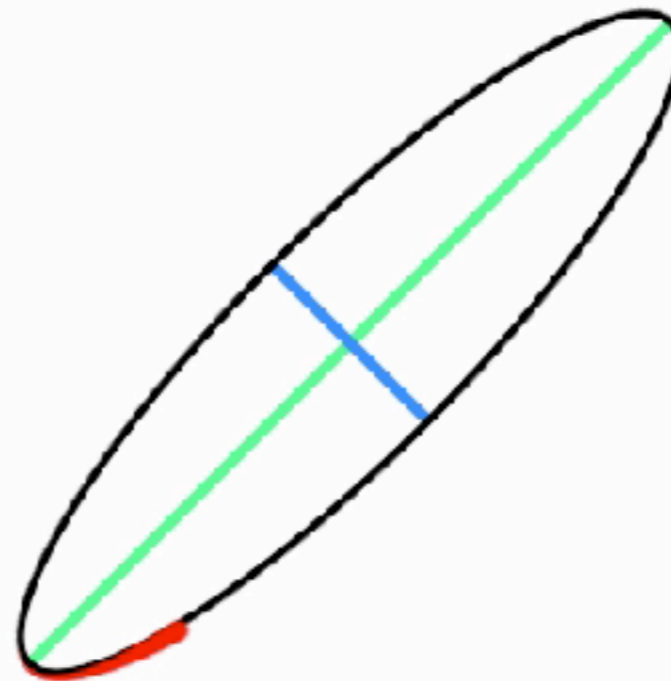
- Orientation
- Taille
- Excentricité
- **Position X**



Elliptic oscillatory gestures

6+1 dimensions

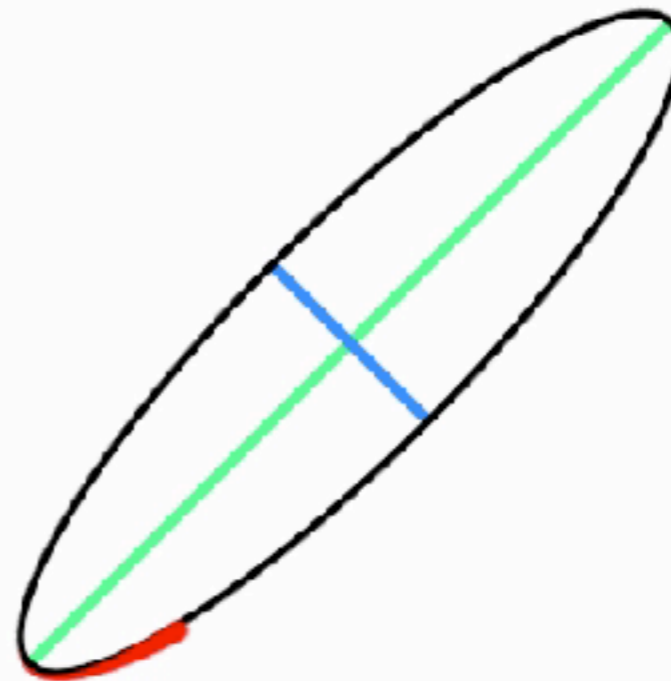
- Orientation
- Taille
- Excentricité
- Position X
- **Position Y**



Elliptic oscillatory gestures

6+1 dimensions

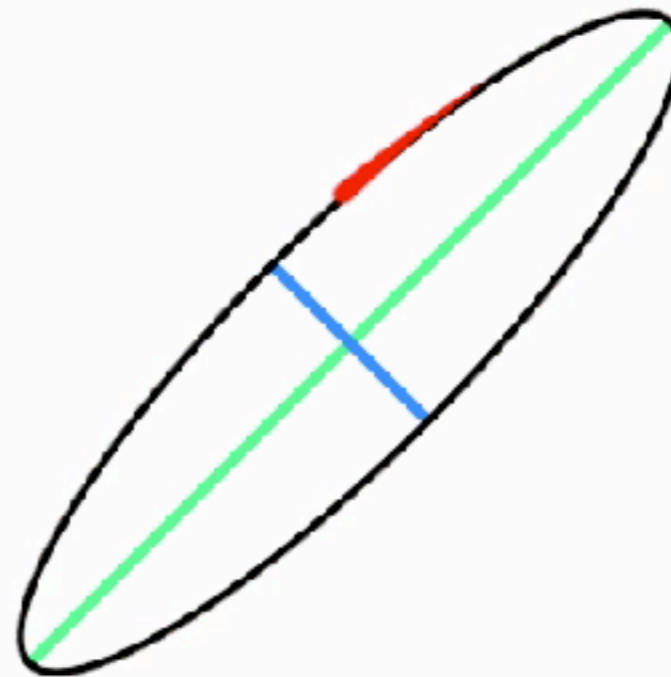
- Orientation
- Taille
- Excentricité
- Position X
- **Position Y**



Elliptic oscillatory gestures

6+1 dimensions

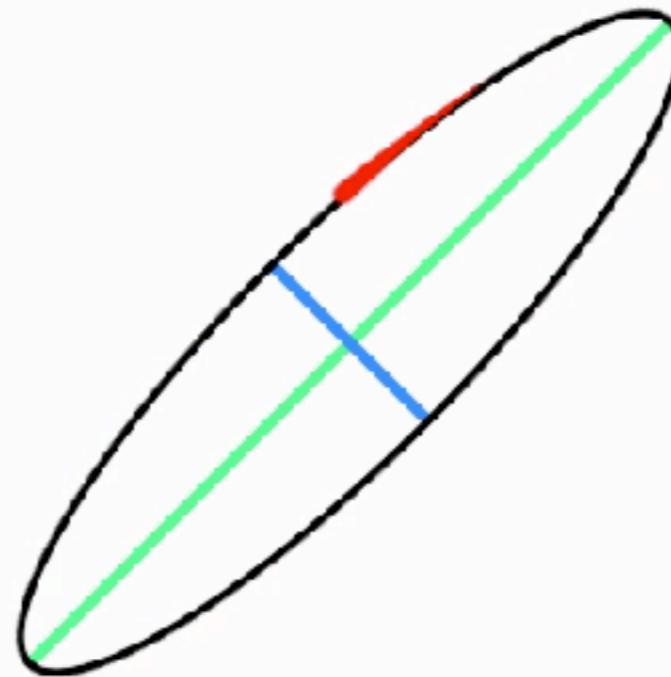
- Orientation
- Taille
- Excentricité
- Position X
- Position Y
- **Fréquence**



Elliptic oscillatory gestures

6+1 dimensions

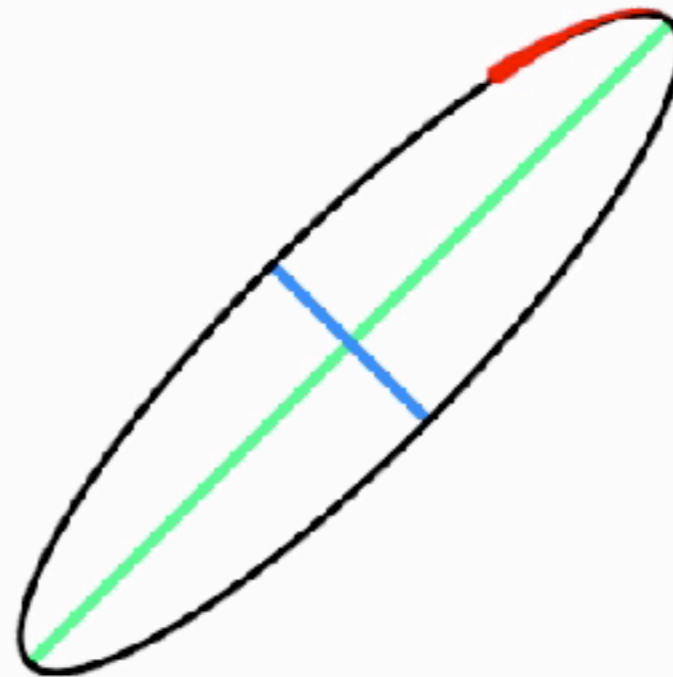
- Orientation
- Taille
- Excentricité
- Position X
- Position Y
- **Fréquence**



Elliptic oscillatory gestures

6+1 dimensions

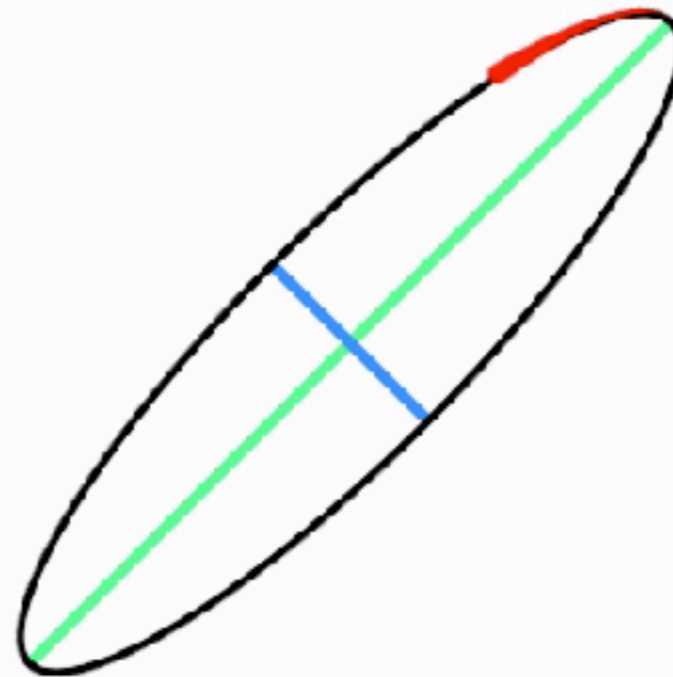
- Orientation
- Taille
- Excentricité
- Position X
- Position Y
- Fréquence
- **Rotation (CW or CCW)**



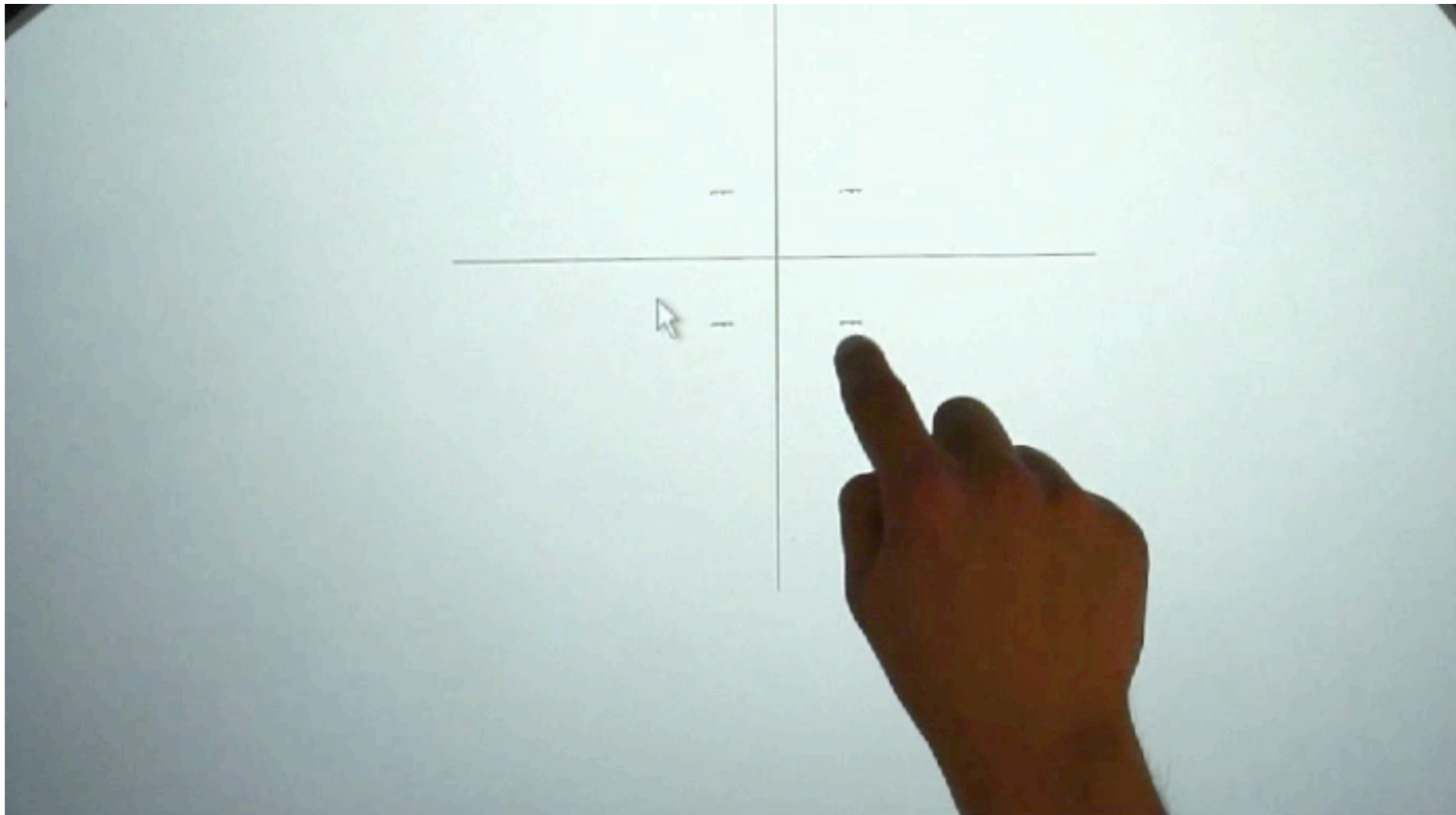
Elliptic oscillatory gestures

6+1 dimensions

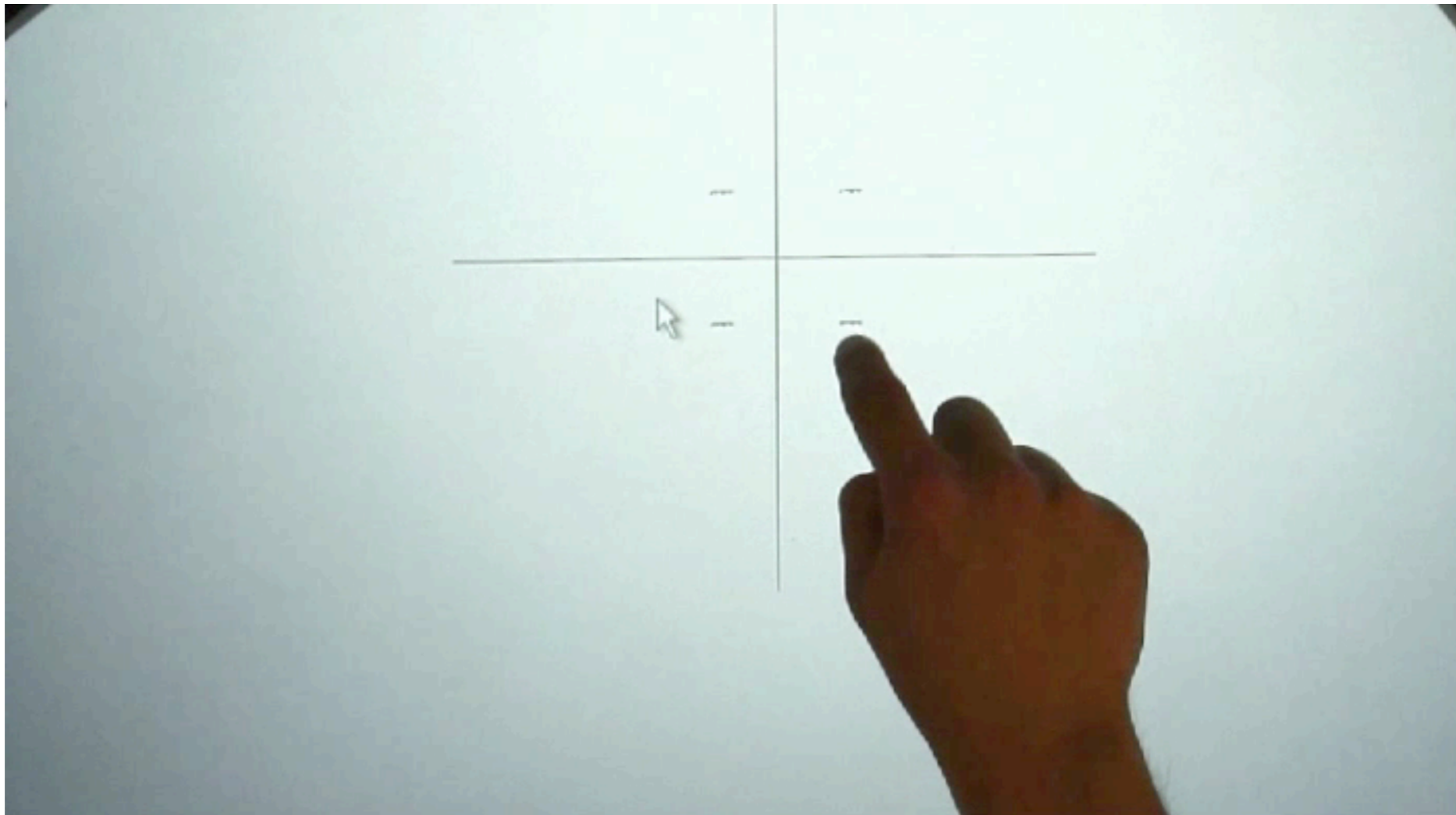
- Orientation
- Taille
- Excentricité
- Position X
- Position Y
- Fréquence
- **Rotation (CW or CCW)**



CycloZoom+



CycloZoom+



Typhoon Takes Aim at Northern Marianas



FWIW, about a decade ago a fishing boat offshore from my home town drew up a honking big WWII bomb. The Coast Guard decided that popping it was the safest solution, which they did in an empty prairie reachable by an inland waterway. Everyone for miles around felt their windows rattle, and no one knew what it was until the news carried the story later.

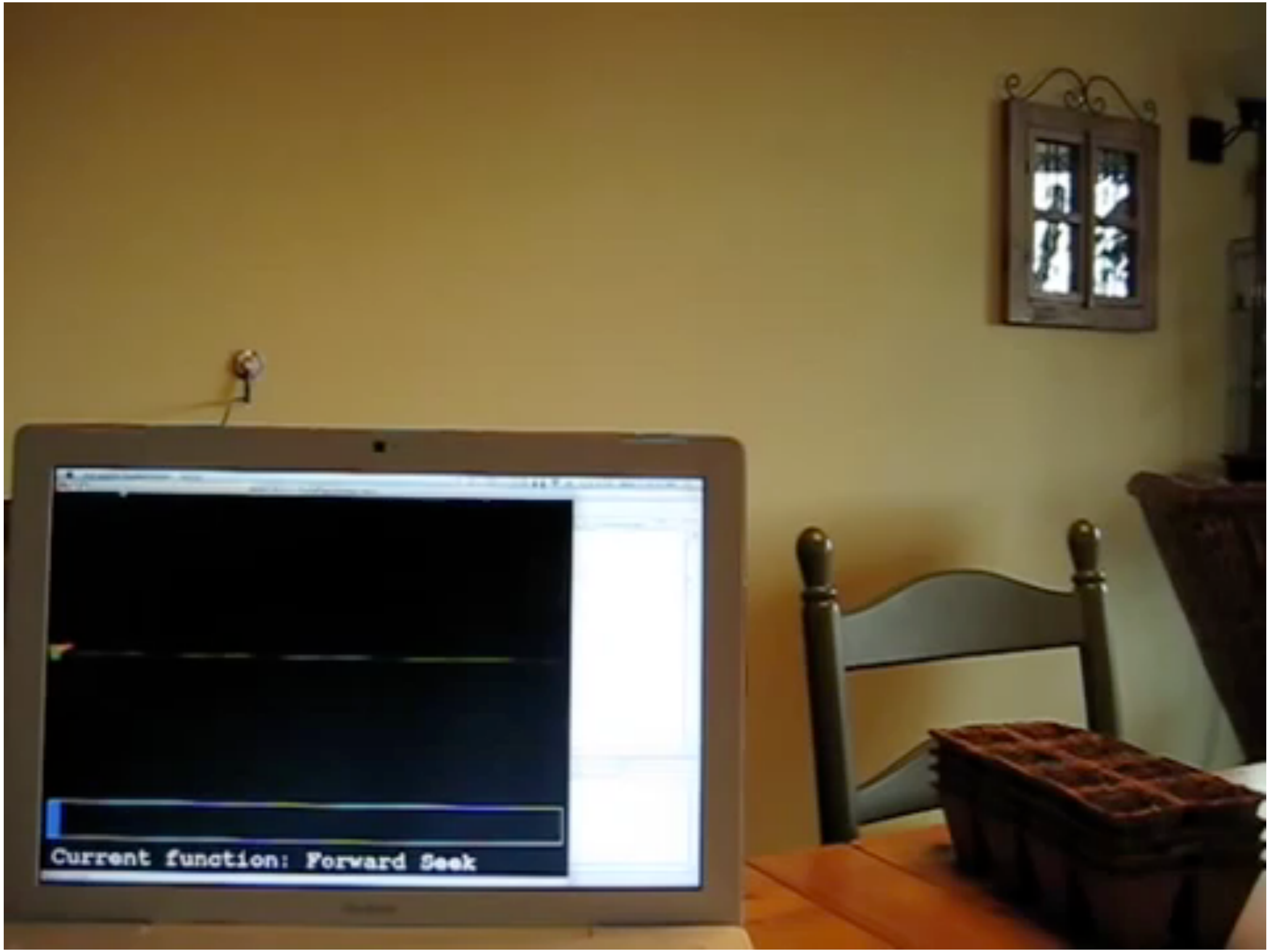
In the past, says Lee, countries had their own scientific facilities, and "could choose to work alone or together. But with ITER, or a new global linear collider, the world can only build one. The only way is to put our resources together." The deadlock, Lee adds, "is not a problem of ITER negotiations. It's generic."

Typhoon Takes Aim at Northern Marianas

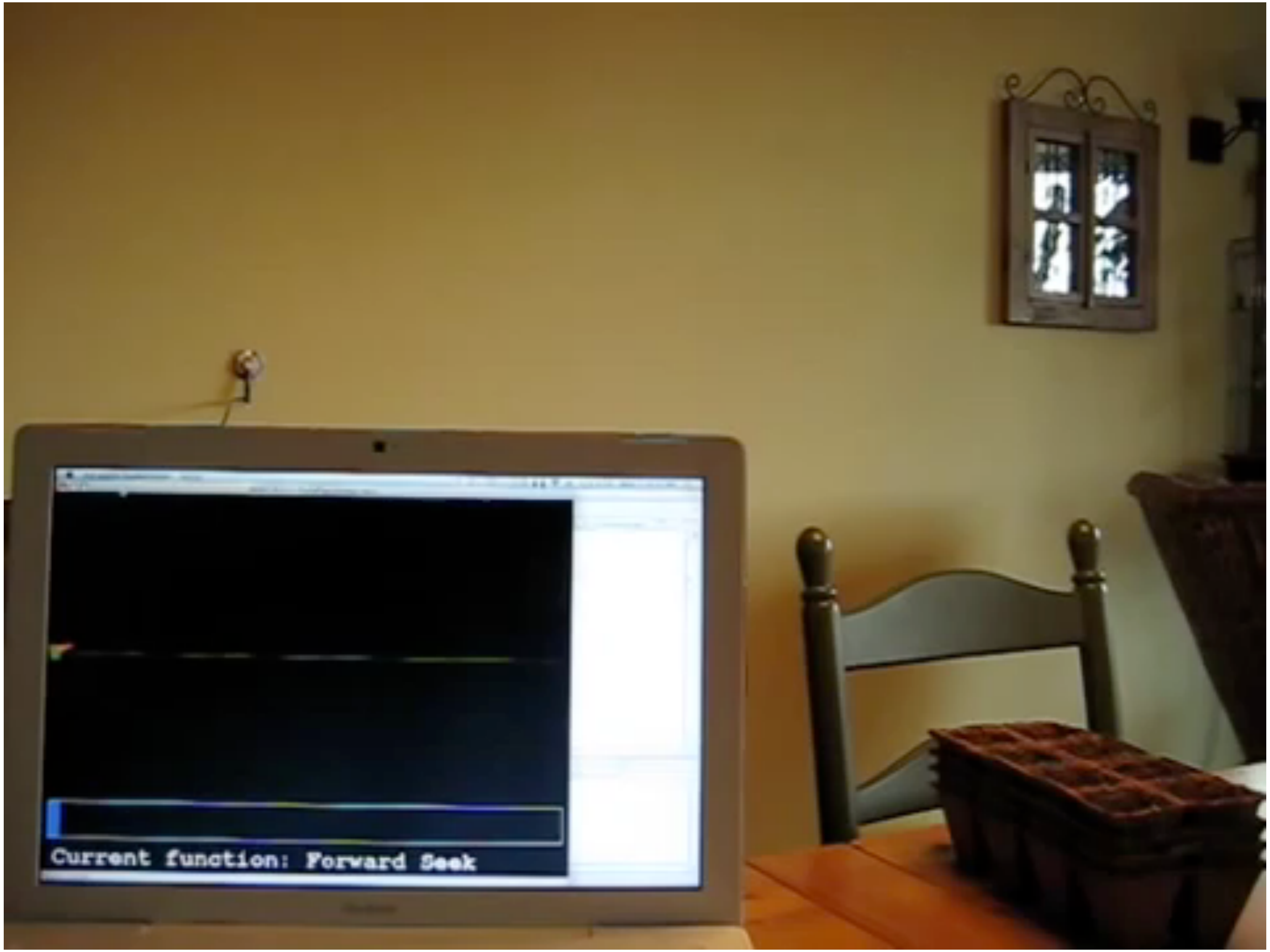


FWIW, about a decade ago a fishing boat offshore from my home town drew up a honking big WWII bomb. The Coast Guard decided that popping it was the safest solution, which they did in an empty prairie reachable by an inland waterway. Everyone for miles around felt their windows rattle, and no one knew what it was until the news carried the story later.

In the past, says Lee, countries had their own scientific facilities, and "could choose to work alone or together. But with ITER, or a new global linear collider, the world can only build one. The only way is to put our resources together." The deadlock, Lee adds, "is not a problem of ITER negotiations. It's generic."



Current function: Forward Seek



Current function: Forward Seek

Challenges

Choisir un set de gestes

Définir une association geste/commande

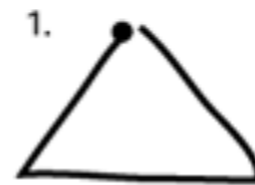
*Imagine that you are designer
You have a list of 24 commands and
You want to build a gestural interface*

Comment procéder ?

Create a gesture set

Symbolic gestures

self-mapping



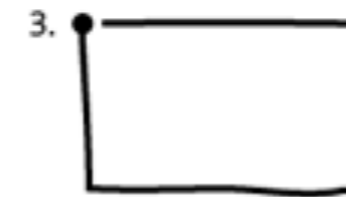
1.

triangle



2.

"x"



3.

rectangle



4.

circle



5.

check



6.

caret



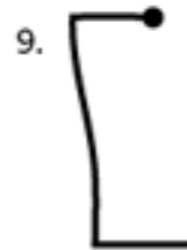
7.

zig-zag



8.

arrow



9.

left square bracket



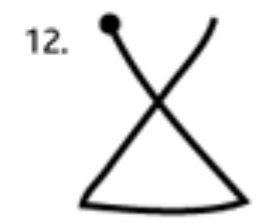
10.

right square bracket



11.

"v"



12.

delete



13.

left curly brace



14.

right curly brace



15.

star



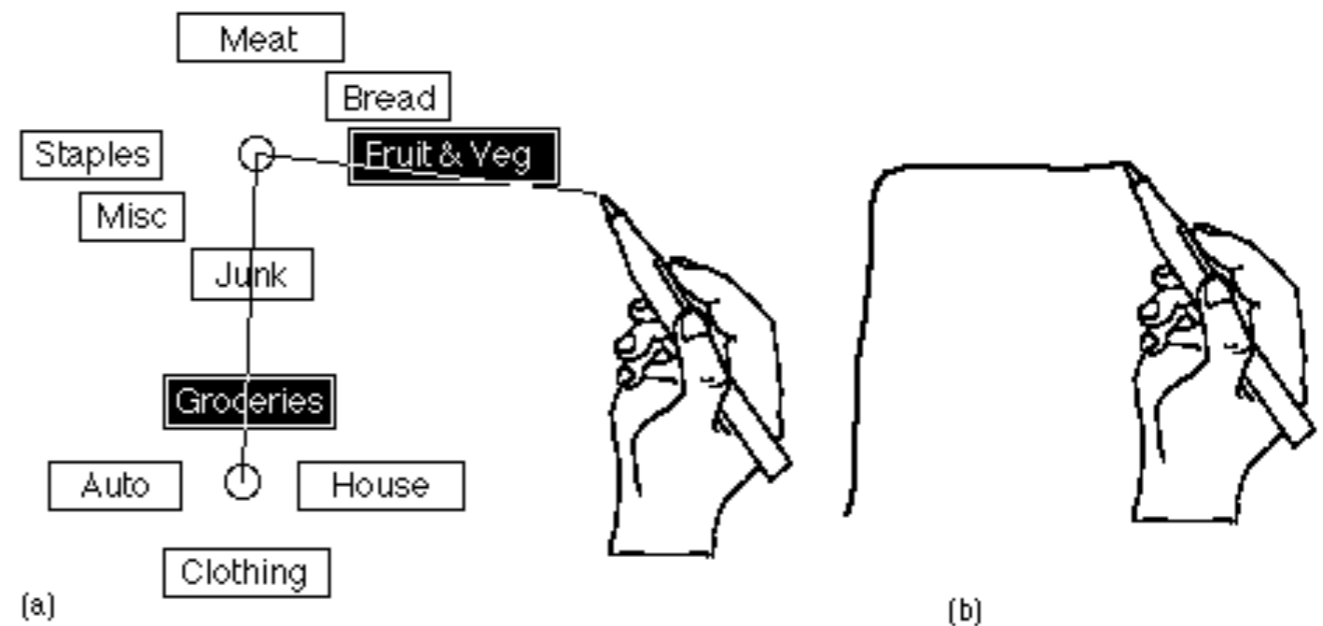
16.

pigtail

Create a gesture set

Abstract Gestures

- ▶ + Organization
- ▶ + Accuracy
- ▶ - Gesture shape suggests no meaning



Define a mapping

User defined gestures

Capture “natural” mappings

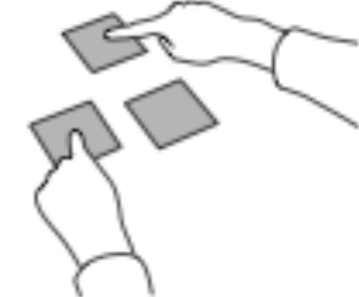
Select Single₁: tap



Select Single₂: lasso



Select Group₁: hold and tap



Select Group₂ and Select Group₃: Use Select Single₁ or Select Single₂ on all items in the group.

Move₁: drag



Move₂: jump

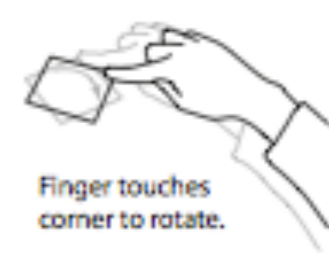


Object jumps to index finger location.

Pan: drag hand



Rotate: drag corner

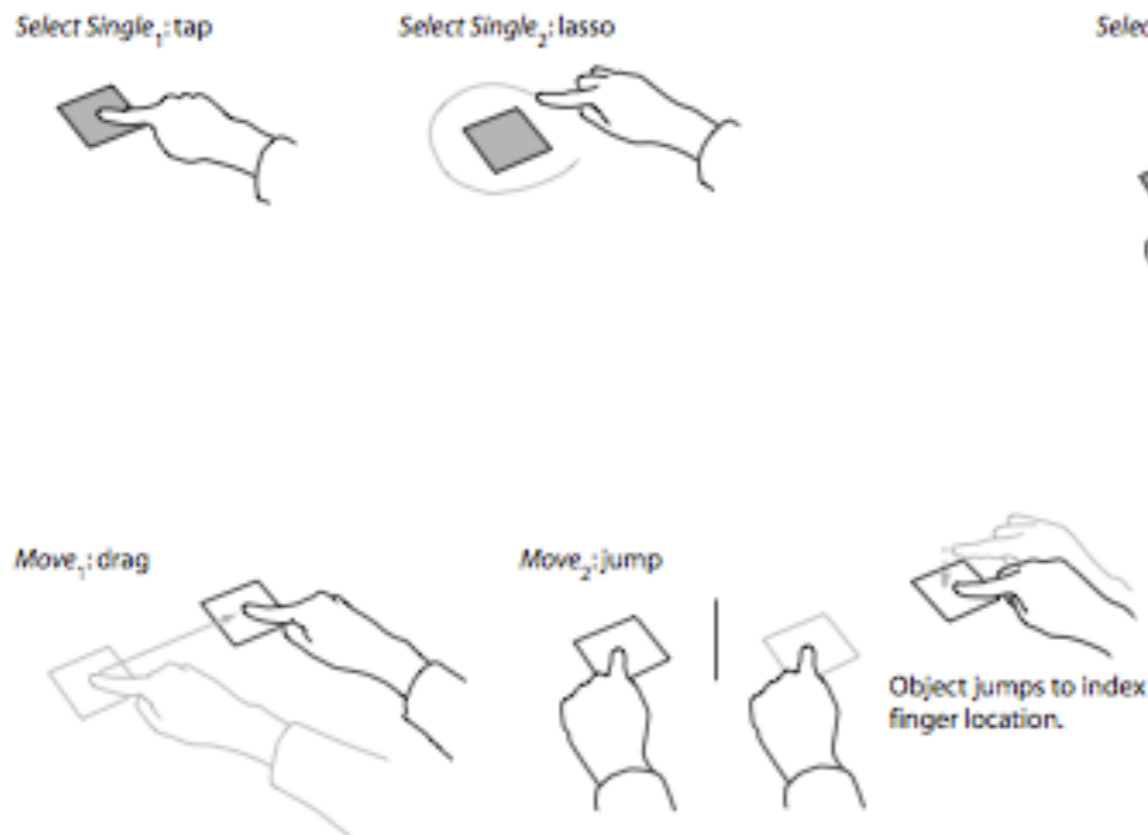


Finger touches corner to rotate.

Define a mapping

User defined gestures

Capture “natural” mappings



| command | number of gestures | % choosing “winner” |
|---------------|--------------------|---------------------|
| accept | 1 | 100% |
| minimize | 3 | 90.9% |
| previous | 2 | 90.9% |
| select single | 3 | 90.9% |
| help | 3 | 86.4% |
| next | 2 | 86.4% |
| open | 5 | 86.4% |
| move | 3 | 81.8% |
| cut | 2 | 77.3% |
| rotate | 4 | 68.2% |
| shrink | 5 | 68.2% |
| delete | 5 | 63.6% |
| pan | 2 | 63.6% |
| undo | 4 | 63.6% |
| select group | 3 | 59.1% |
| menu | 5 | 54.5% |
| paste | 4 | 54.5% |
| reject | 5 | 54.5% |
| enlarge | 5 | 45.5% |
| zoom in | 5 | 45.0% |
| duplicate | 4 | 36.4% |
| zoom out | 6 | 22.7% |

Define a mapping

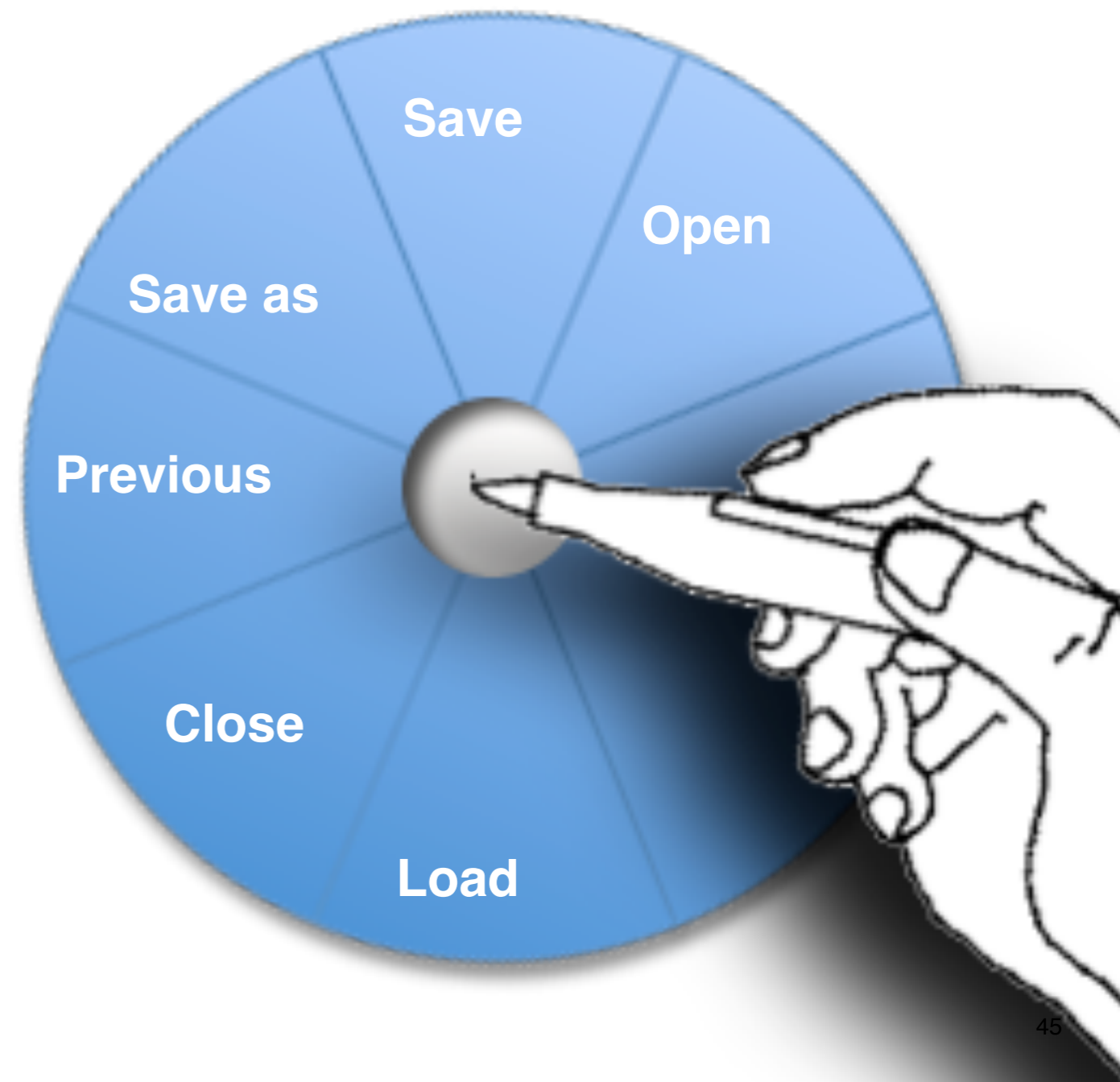
Semantic relationships

Focus on the relation between

- ▶ the different gestures
- ▶ the different commands

Highlight:

- ▶ Similarity
- ▶ Opposition
- ▶ etc.

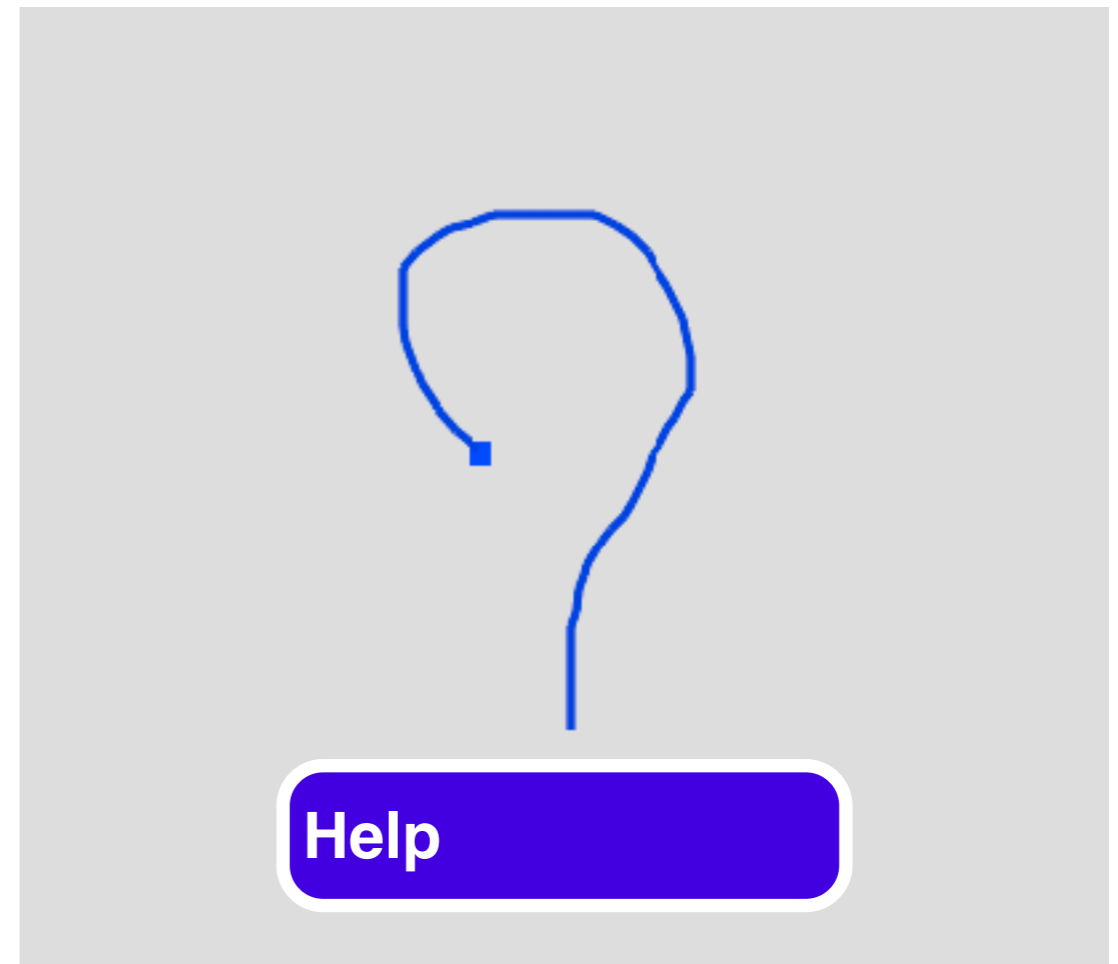
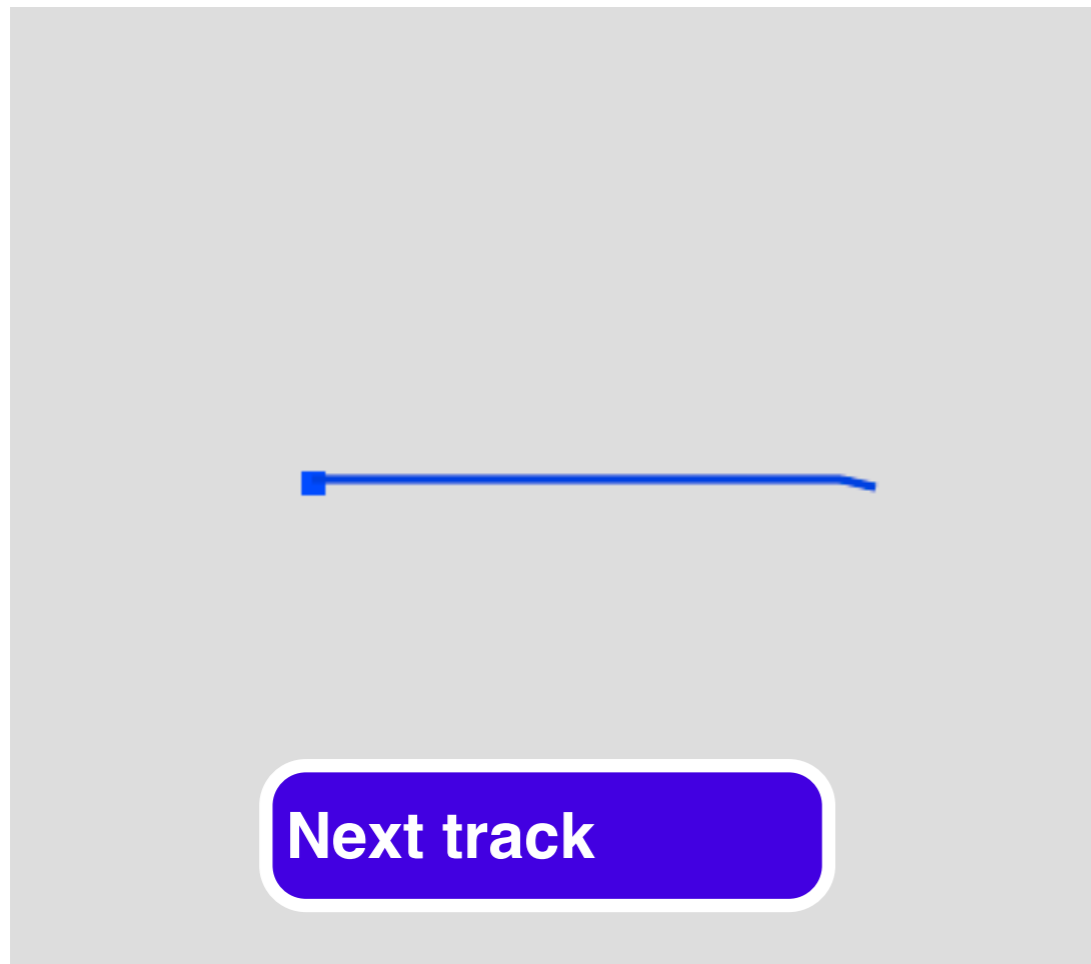


Touch-screen gestures design guidelines

Zhai, S., Kristensson, P. O., Appert, C., Andersen, T. H., & Cao, X. (2012). Foundational Issues in Touch-Screen Stroke Gesture Design-An Integrative Review. *Foundations and Trends in Human-Computer Interaction*, 5(2), 97-205.

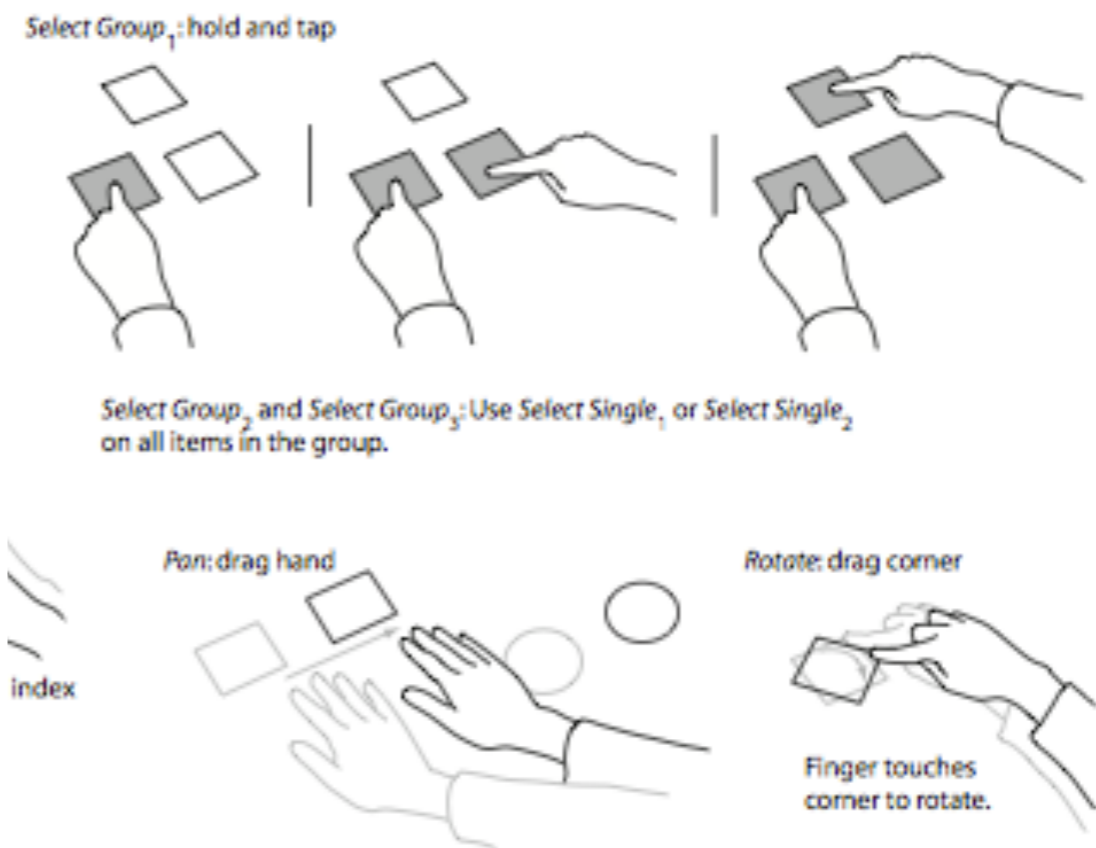
Gesture design guidelines [zhai et al. 2012]

Making Gestures Analogous to Physics or Convention



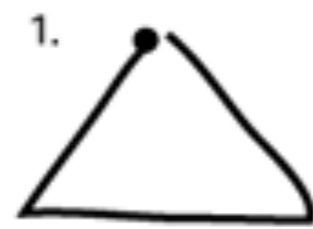
Gesture design guidelines [zhai et al. 2012]

Make gestures accessible to novices



Gesture design guidelines [zhai et al. 2012]

Making gestures as simple as possible



triangle



"X"



rectangle



circle



check



caret



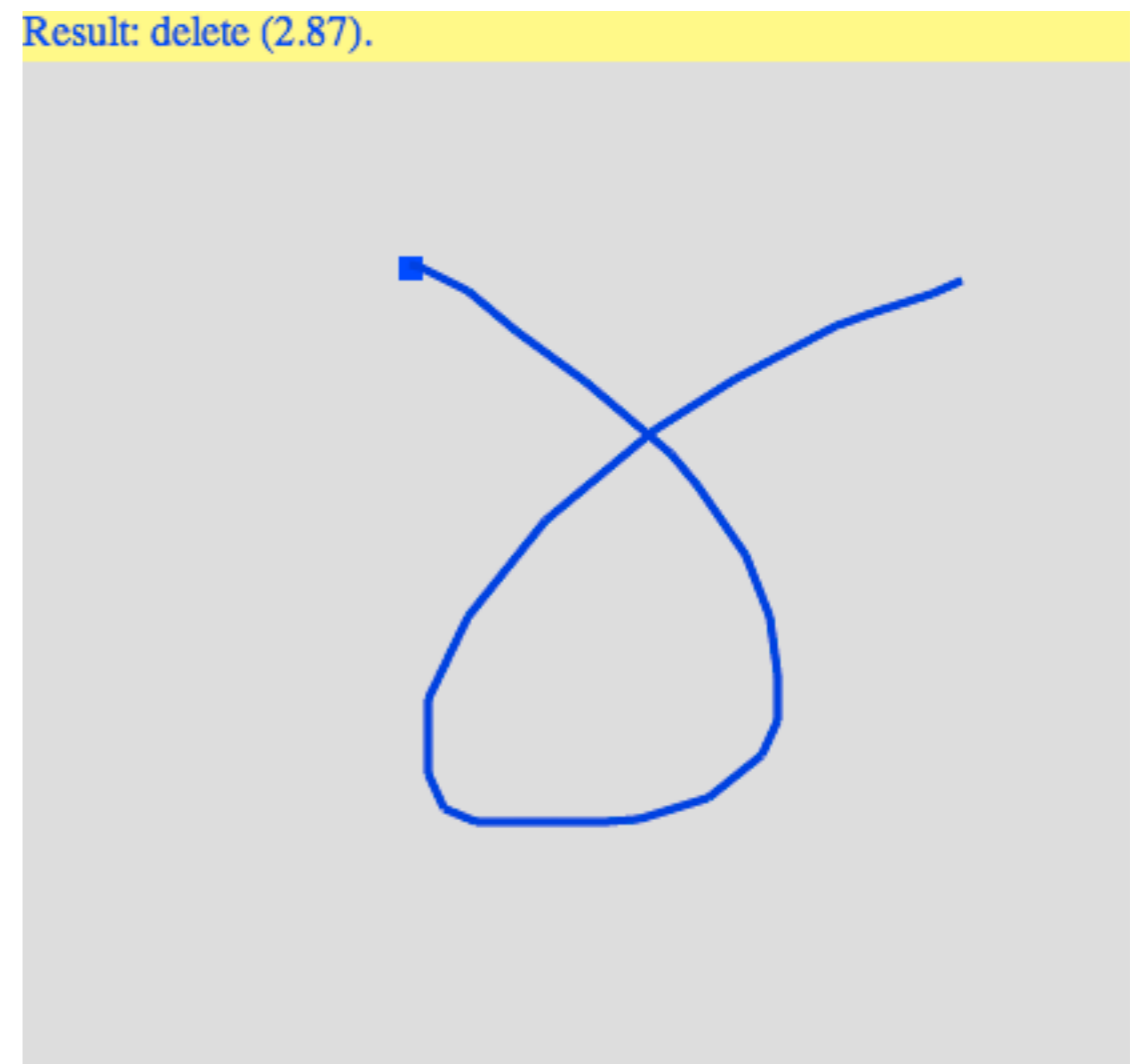
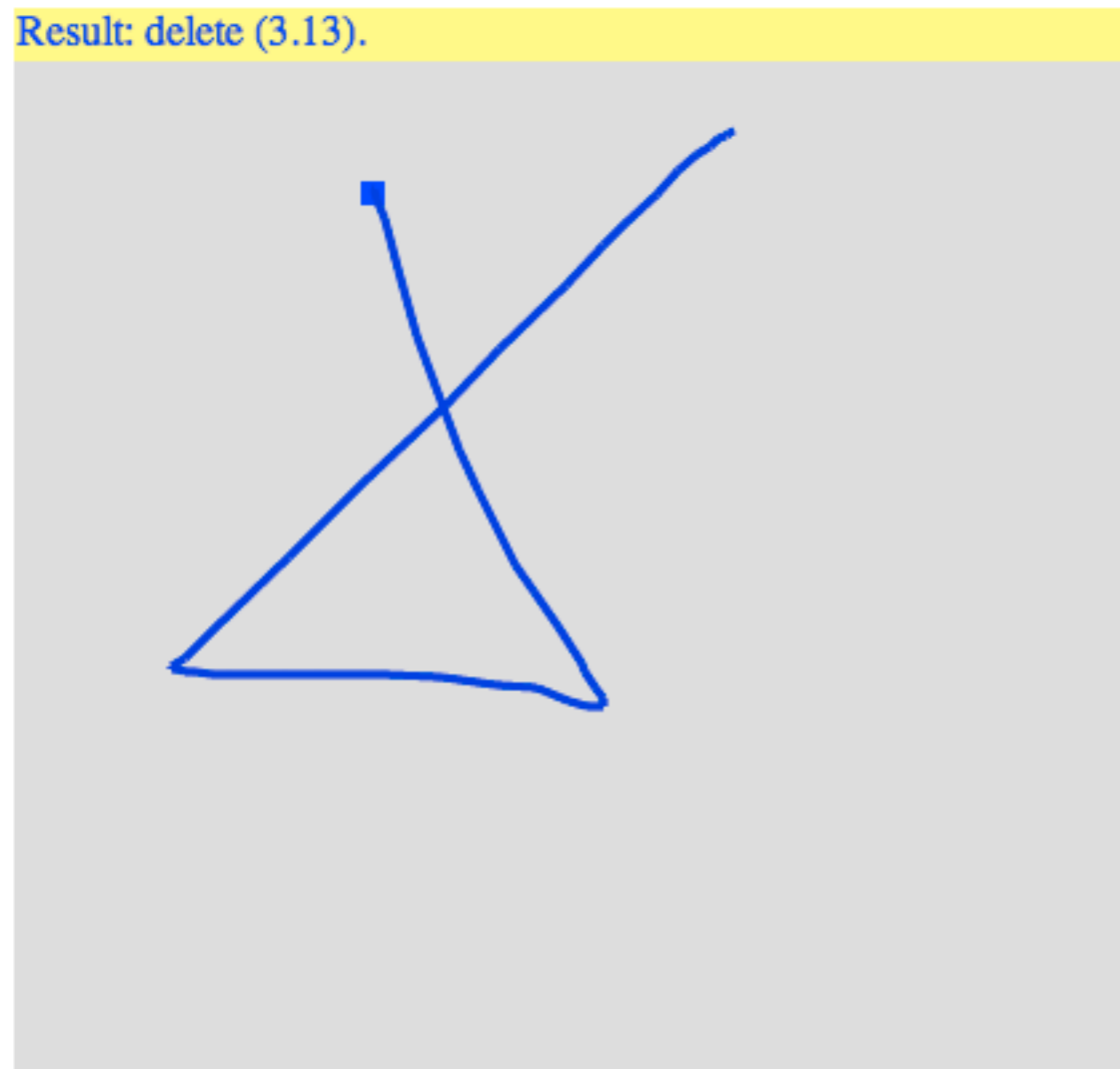
zig-zag



arrow

Gesture design guidelines [zhai et al. 2012]

Making Gestures Distinct



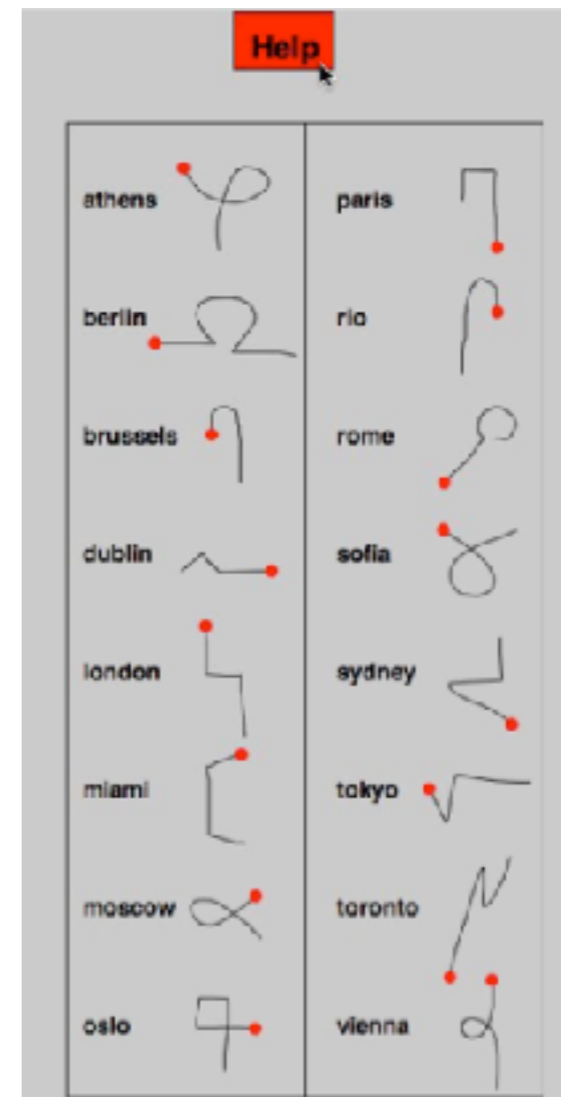
Gesture design guidelines [zhai et al. 2012]

Making gestures systematic



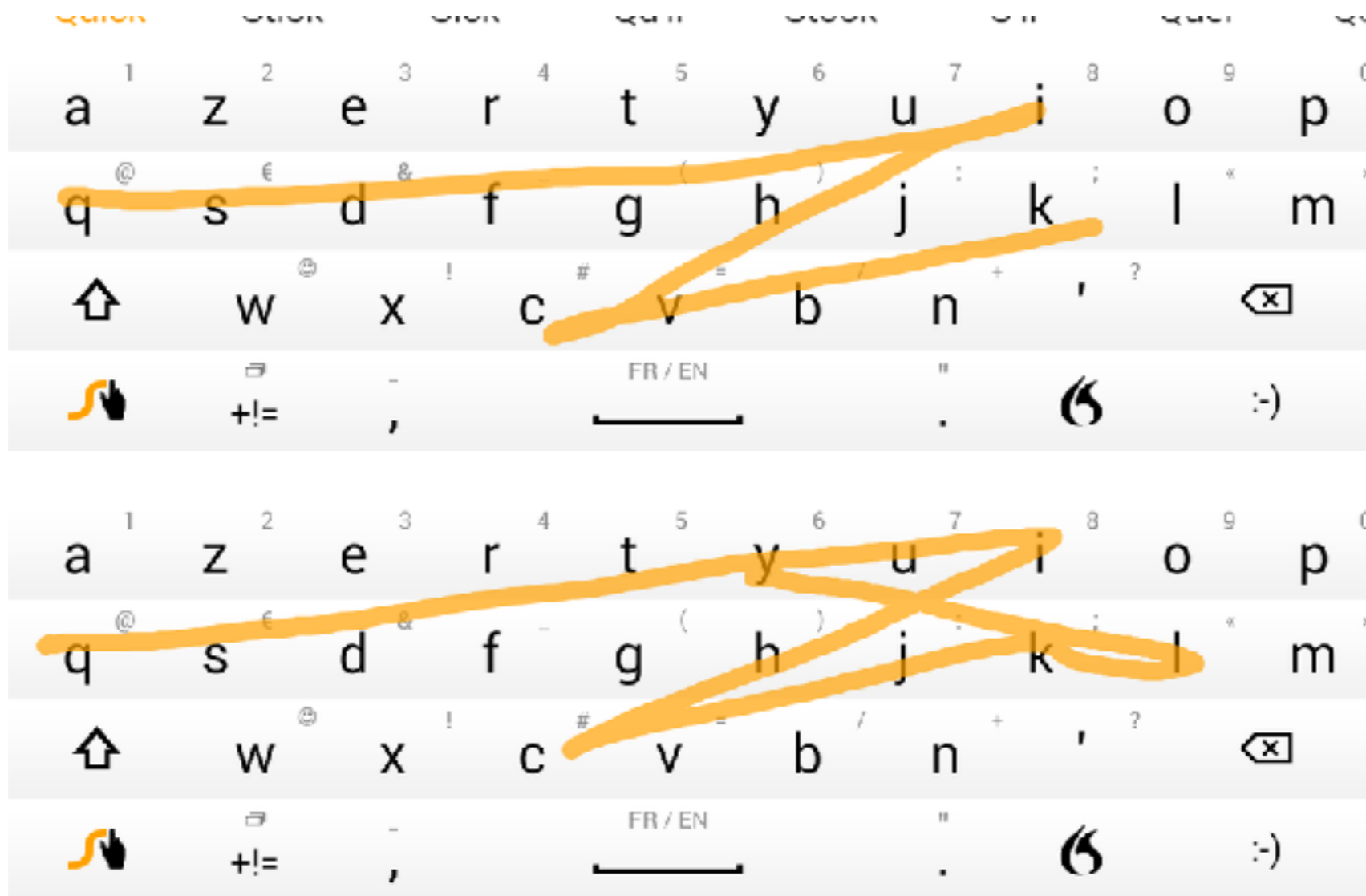
Gesture design guidelines [zhai et al. 2012]

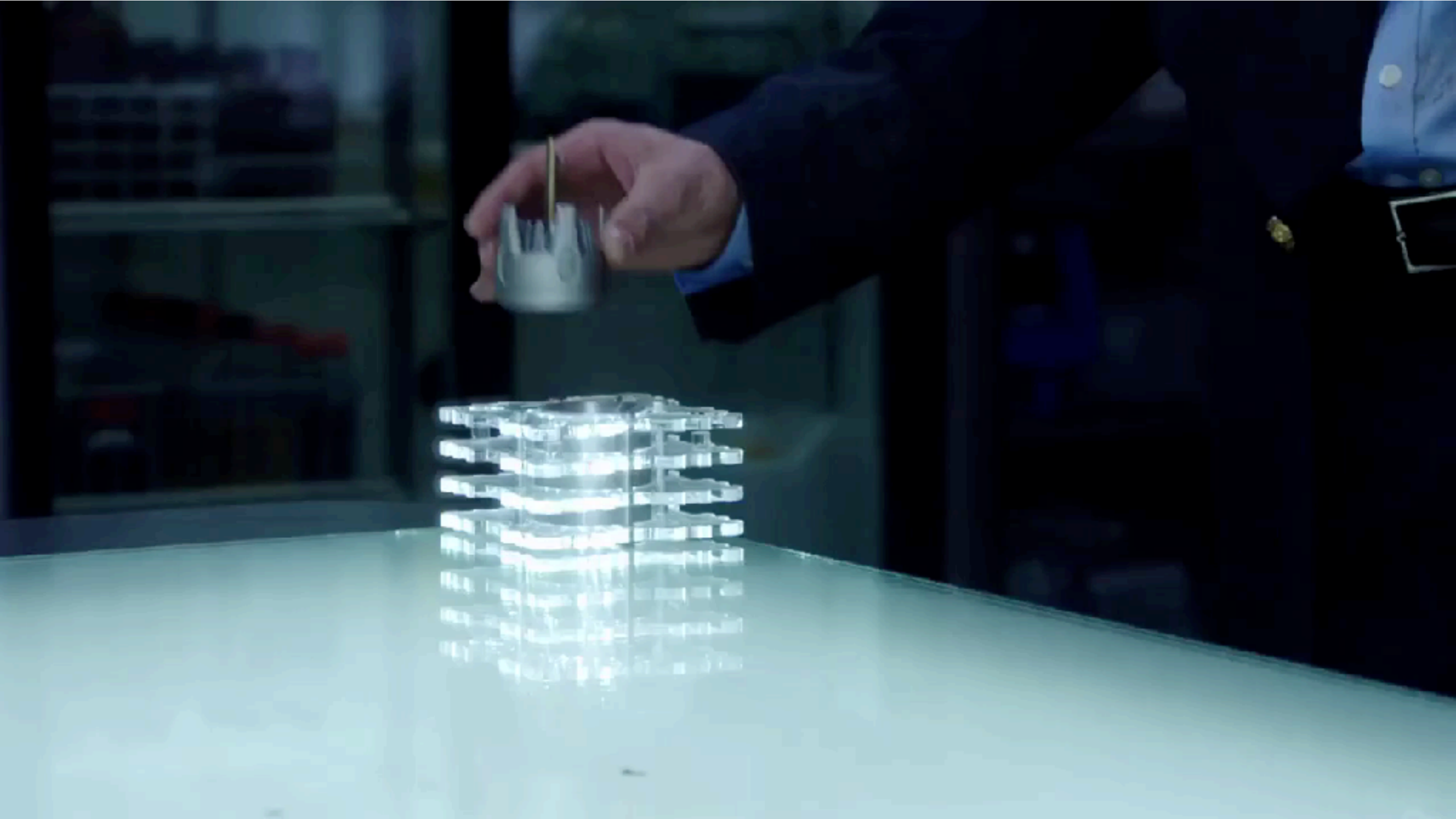
Making gestures self-revealing

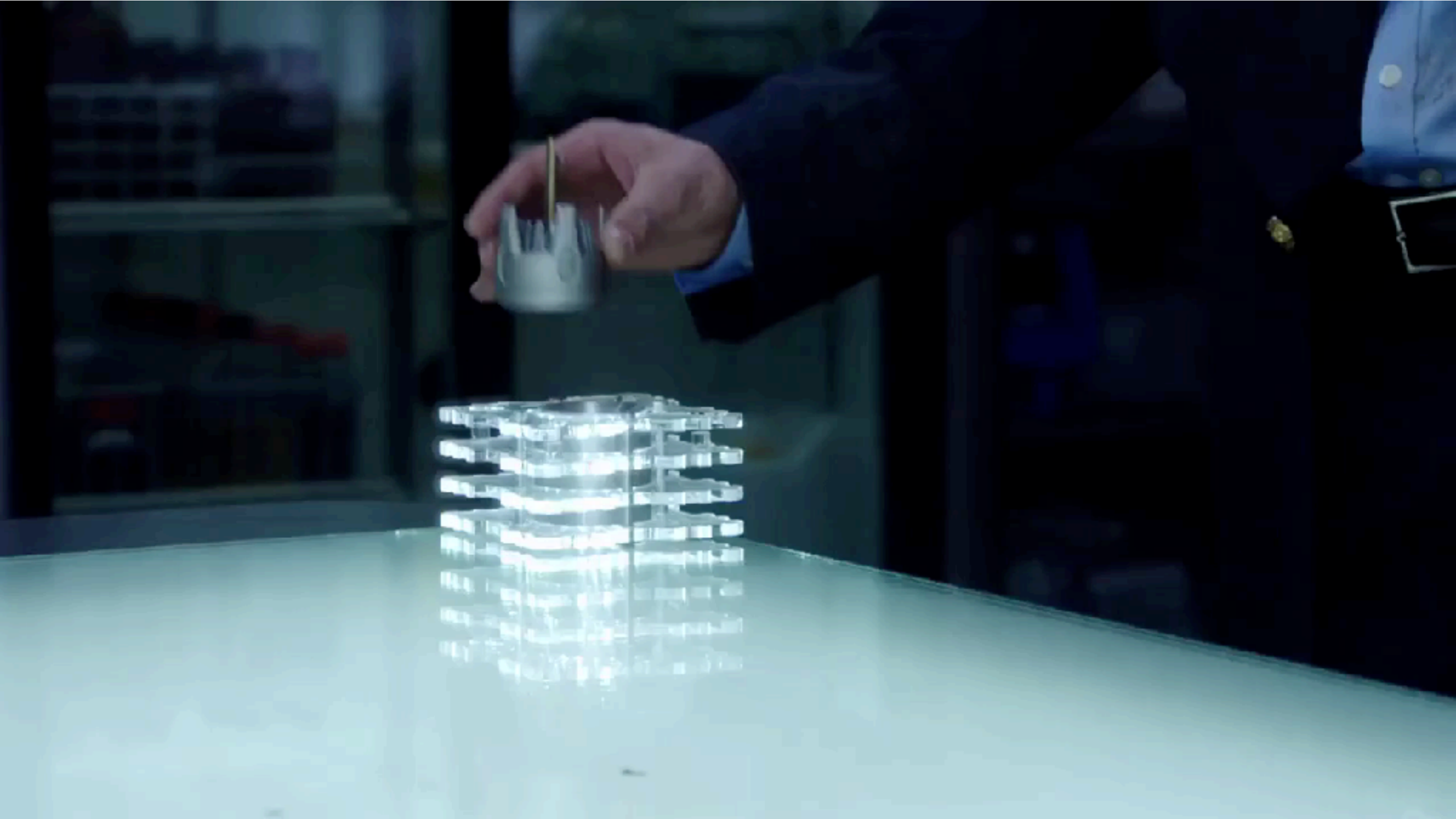


Gesture design guidelines [zhai et al. 2012]

Supporting the right level of chunking





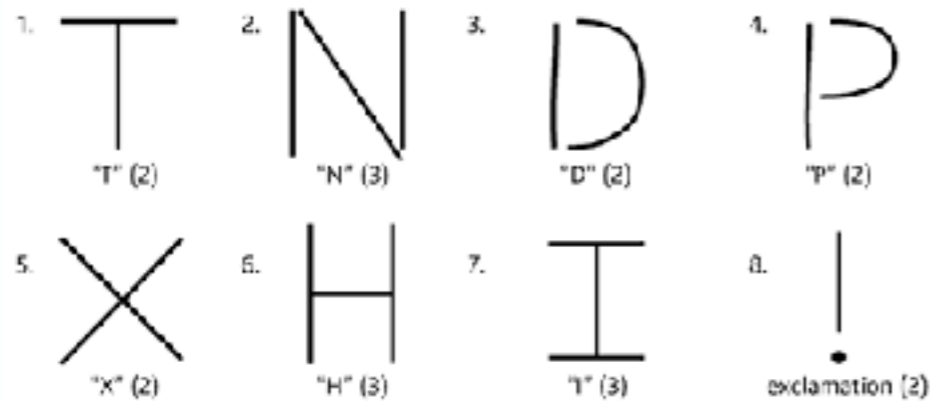


Focus de ce cours

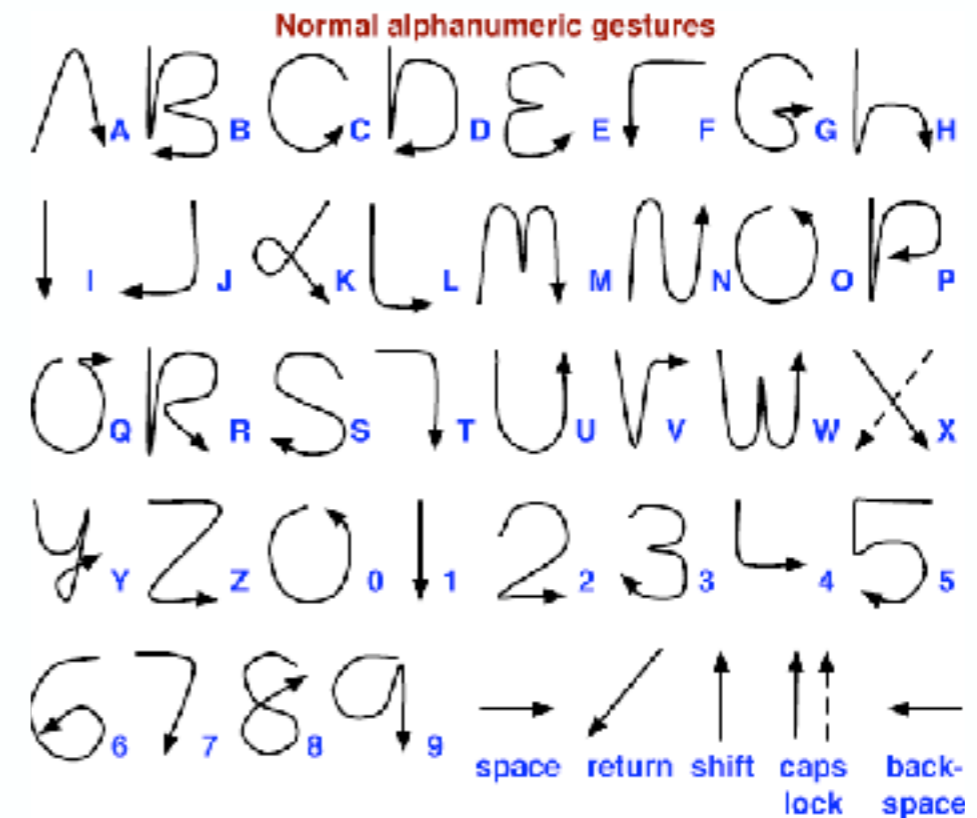
Reconnaître des gestes de l'utilisateur

Les associer à des commandes

On se limite aux gestes un seul trait (unistroke)



Multistroke



Unistroke - Graffiti (Palm OS)

Technique de reconnaissance de gestes élémentaires

libstroke (1997)

Librairie disponible dans FVWM <http://etla.net/libstroke/>

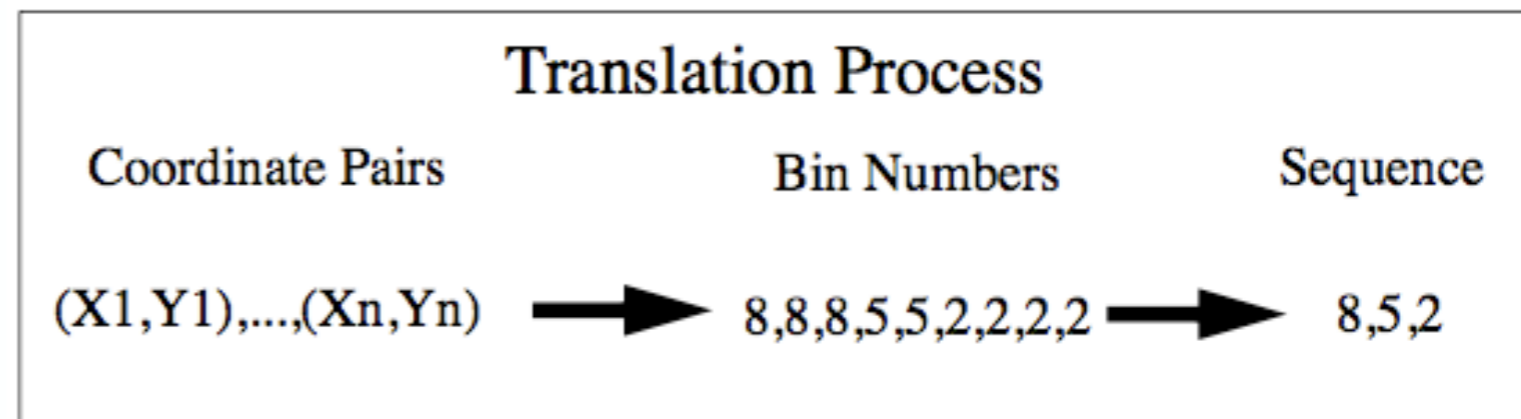
Les gestes sont décrits par des séquences de chiffres

```
# Strokes
#      num      button context mod. action
Stroke 14789    2      A      N      Exec exec xlock -mode blank&
Stroke 258      2      A      N      Exec exec xterm&
Stroke 563214789 2      A      N      Exec exec exmh&
Stroke 7415963  2      A      N      Exec exec netscape&
Stroke 741236987 2      A      N      Destroy
Stroke 1478963  2      A      N      Popup "Apps"
Stroke 74123    2      A      N      Module "winlist" FvwmWinList
Stroke 74159    2      A      N      Move
Stroke 852      2      A      N      Menu "Apps" Nop
```

| | | |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 5 | 6 |
| 7 | 8 | 9 |

Libstroke: algorithme

- A. Calcul de la bounding box à partir de minx, maxx, miny, maxy
- B. Division de la boite en 9 cellules
- C. Étiquetage des points
- D. Factorisation
- E. Comparaison aux motifs



Extension Firefox: Mouse Gestures Redox

Techniques de reconnaissance

Classifieurs statistiques

Modèles de Markov cachés

Réseaux de neurones

Méthodes ad-hoc

3 techniques fréquemment utilisées:

- ▶ Rubine
- ▶ Dynamic Time Warping (DTW)
- ▶ 1\$ recognizer

Techniques de reconnaissance

\$1 recognizer

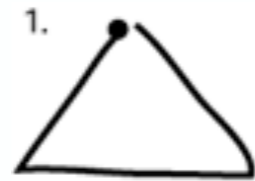
Simple à implémenter

Taux de reconnaissance comparables à Rubine et DTW

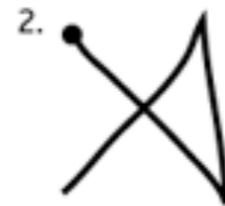
1 seul exemple suffit

*Jacob O. Wobbrock, Andrew D. Wilson, Yang Li.
Gestures without libraries, toolkits or training: a \$1 recognizer for user interface prototypes
UIST '07, 159-168.*

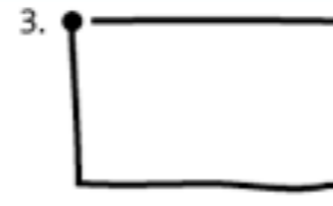
\$1 recognizer



triangle



"x"



rectangle



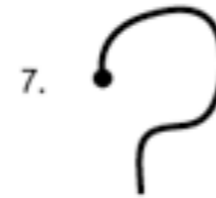
circle



check



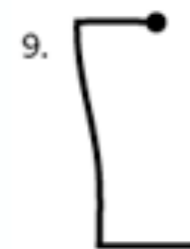
caret



question



arrow



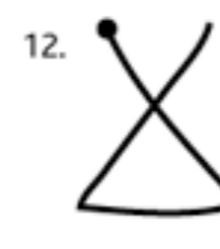
left square bracket



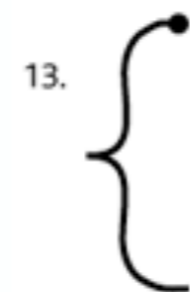
right square bracket



"v"



delete



left curly brace



right curly brace



star



pigtail

Algorithme

A. L'utilisateur réalise un geste

- ▶ Le geste est représenté par une liste ordonnée de points

B. Ce geste est comparé à un ensemble de gestes de référence

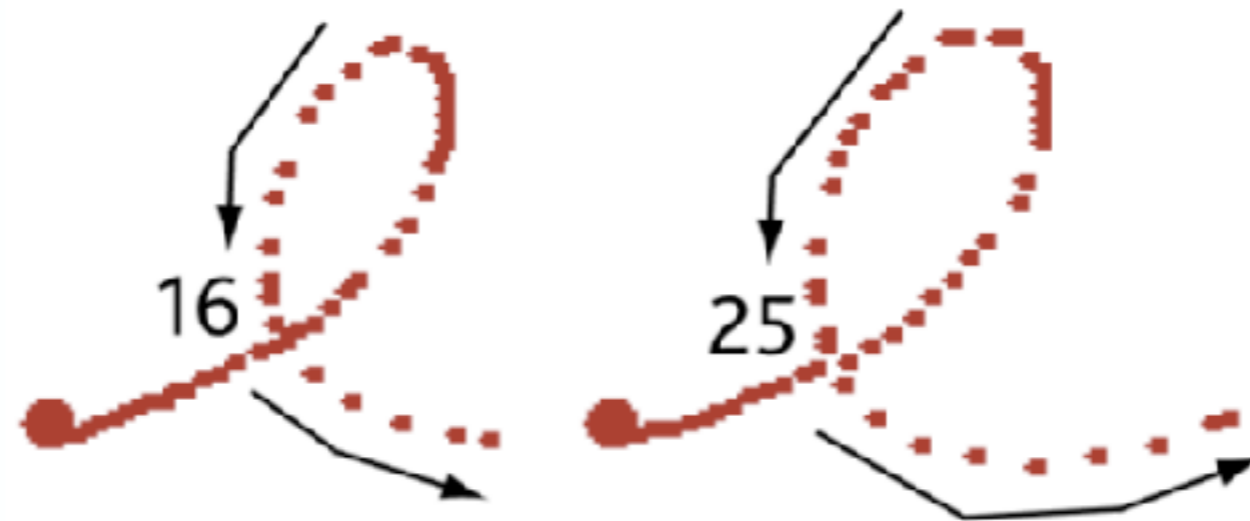
- Les références sont appelées « templates »
- On mesure la distance euclidienne

C. Le geste reconnu est celui pour lequel cette distance est minimale

Problèmes posés

Le nombre de points d'un geste dépend de :

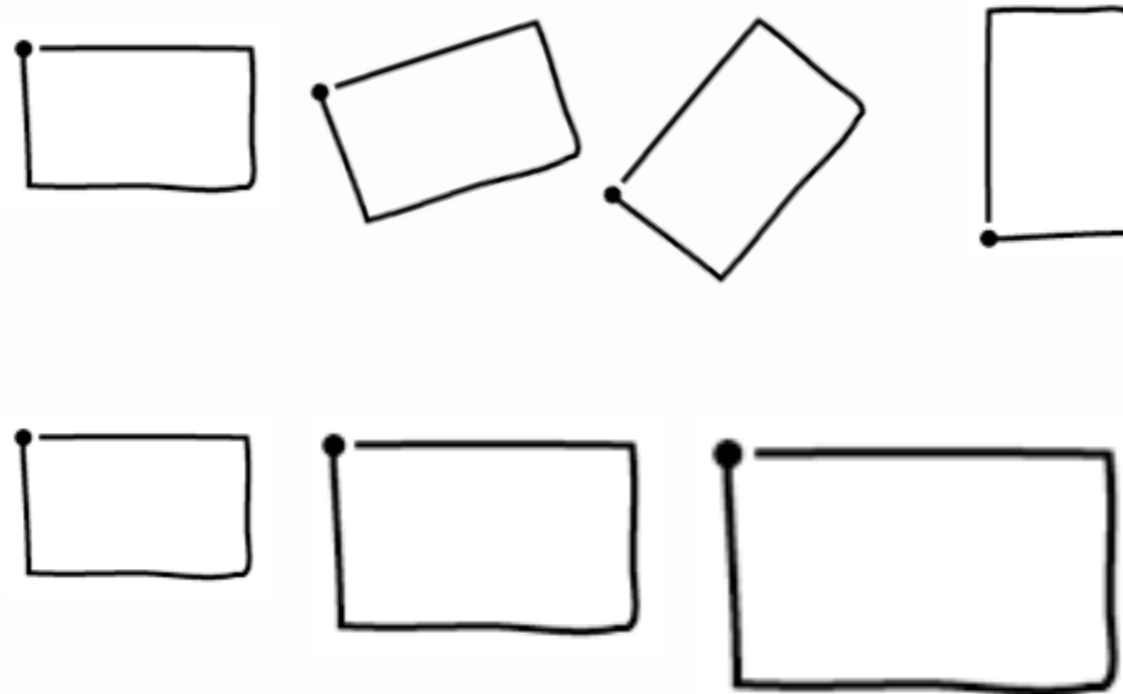
- La vitesse d'exécution
- La fréquence d'échantillonnage du périphérique
- ...



Problèmes posés

Un geste peut être réalisé à différentes positions sur l'interface.

Suivant différentes orientations et différentes échelles.



4 étapes

A. Ré-échantillonner le geste

→ Invariant à la fréquence d'acquisition

→ Invariant à la vitesse d'exécution

B. Ré-orientation du geste

→ Invariant à l'orientation

C. Mise à l'échelle et translation

→ Invariant à l'échelle

→ Invariant à la position

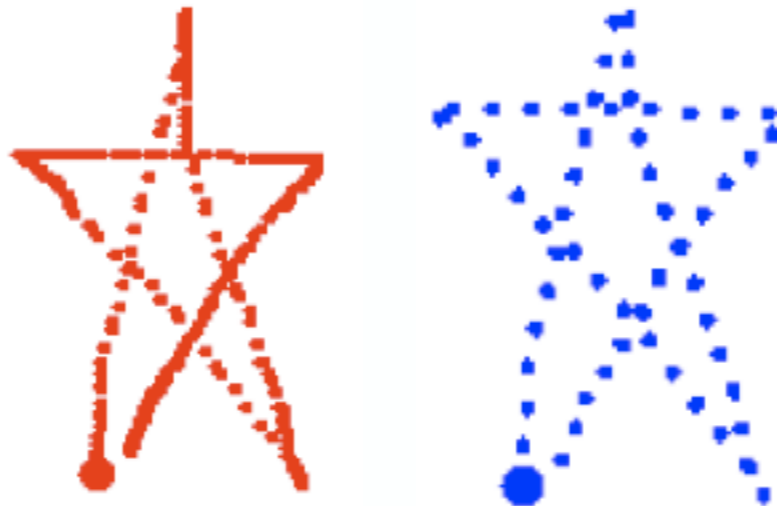
D. Reconnaissance du geste

1ère étape : Ré-échantillonnage

Le geste est défini par M points ordonnés.

On veut N points ordonnés équidistants les uns des autres.

$N = 64$



1ère étape : Ré-échantillonnage

A. Calcul de la distance l entre 2 points :

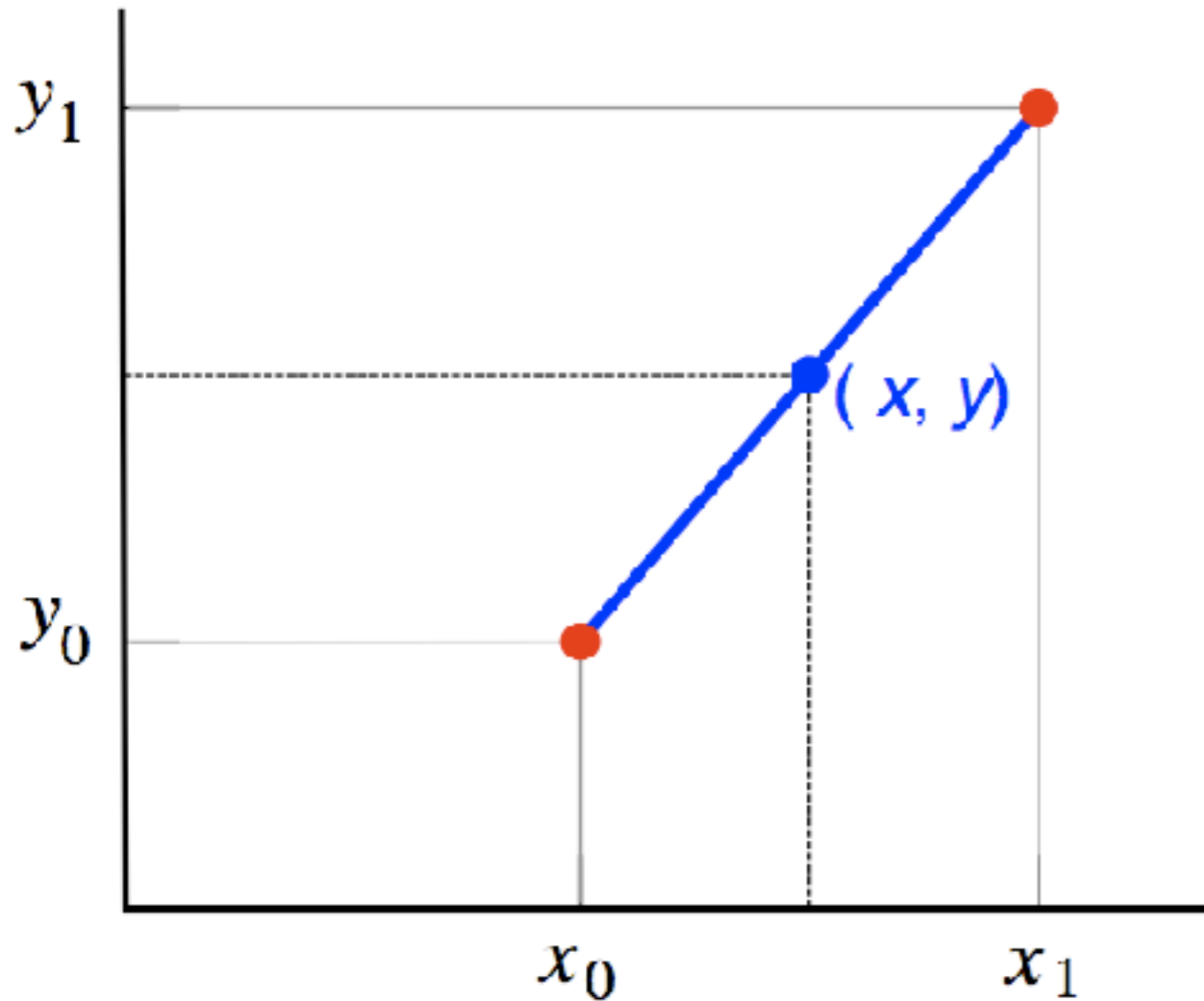
- ▶ Calcul de la longueur totale du geste.
- ▶ $l = \text{longueur} / (N-1)$

B. Interpolation linéaire sur les points du geste d'origine.

- ➔ Permet de calculer la distance en prenant les points 2 à 2.

1ère étape : Ré-échantillonnage

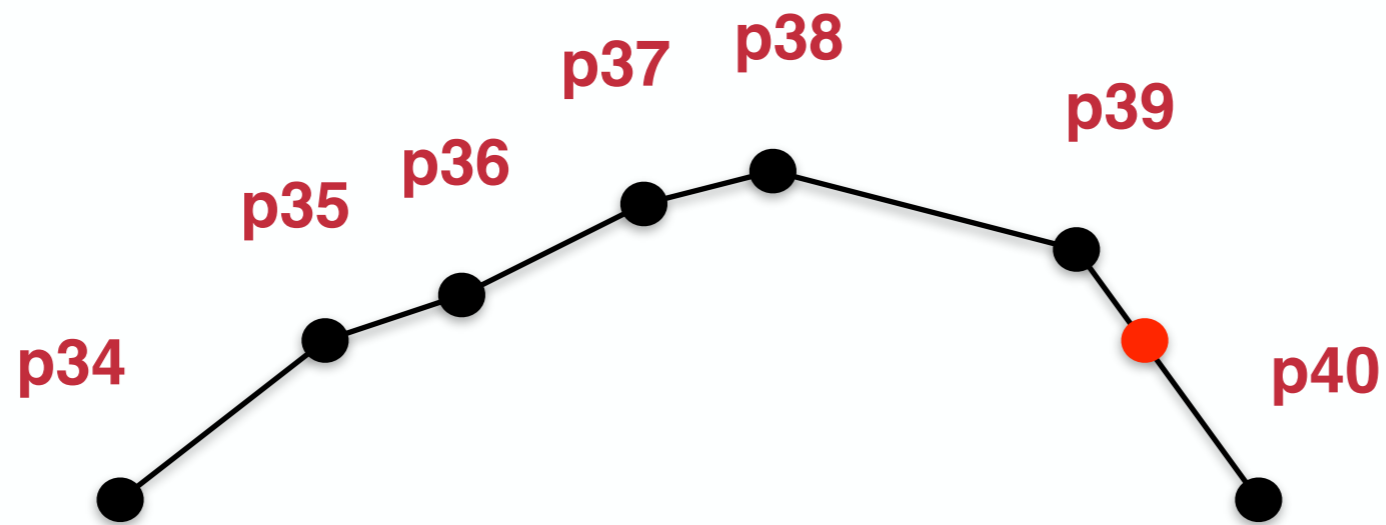
Interpolation linéaire



$$\frac{y - y_0}{x - x_0} = \frac{y_1 - y_0}{x_1 - x_0}$$

$$y = y_0 + (x - x_0) \frac{y_1 - y_0}{x_1 - x_0}$$

1ère étape : Ré-échantillonnage



$$\sum_{i=35}^{40} distance(p_{i-1}, p_i) > I$$

1ère étape : Ré-échantillonnage

Step 1. Resample a *points* path into *n* evenly spaced points.

RESAMPLE(*points*, *n*)

```
1   $I \leftarrow \text{PATH-LENGTH}(\textit{points}) / (n - 1)$ 
2   $D \leftarrow 0$ 
3   $\textit{newPoints} \leftarrow \textit{points}_0$ 
4  foreach point  $p_i$  for  $i \geq 1$  in points do
5     $d \leftarrow \text{DISTANCE}(p_{i-1}, p_i)$ 
6    if  $(D + d) \geq I$  then
7       $q_x \leftarrow p_{i-1}_x + ((I - D) / d) \times (p_{i_x} - p_{i-1}_x)$ 
8       $q_y \leftarrow p_{i-1}_y + ((I - D) / d) \times (p_{i_y} - p_{i-1}_y)$ 
9      APPEND(newPoints, q)
10     INSERT(points, i, q) // q will be the next  $p_i$ 
11      $D \leftarrow 0$ 
12   else  $D \leftarrow D + d$ 
13 return newPoints
```

PATH-LENGTH(*A*)

```
1   $d \leftarrow 0$ 
2  for i from 1 to  $|A|$  step 1 do
3     $d \leftarrow d + \text{DISTANCE}(A_{i-1}, A_i)$ 
4  return d
```

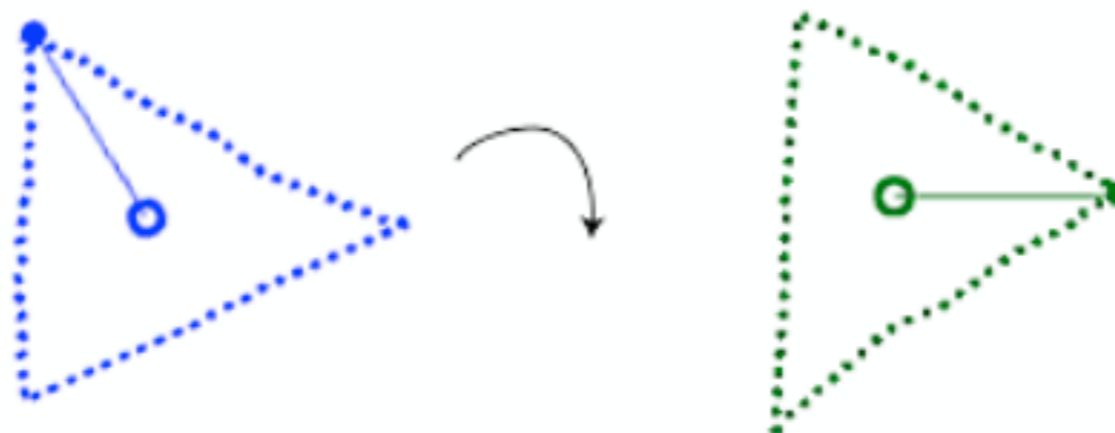
2e étape : Rotation « indicative »

A. Calcul du centre du geste (centroïde)

B. Calcul de l'angle entre :

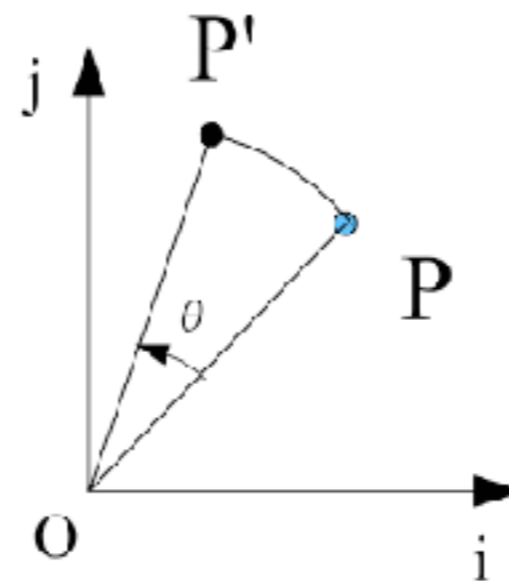
- Le centroïde,
- Le premier point
- L'horizontale

C. Rotation des points en utilisant cet angle



2e étape : Rotation « indicative »

Rotation d'un point



$$\begin{cases} x' &= x \cos(\theta) - y \sin(\theta) \\ y' &= x \sin(\theta) + y \cos \theta \end{cases}$$

Comment retrouver ? Soit α l'angle (i, OP) et $r = \|OP'\| = \|OP\|$. Alors

$$\begin{cases} x' &= r \cos(\alpha + \theta) \\ y' &= r \sin(\alpha + \theta) \end{cases}$$

$$\begin{cases} x' &= r \cos(\alpha) \cos(\theta) - r \sin(\alpha) \sin(\theta) \\ y' &= r \cos(\alpha) \sin(\theta) + r \sin(\alpha) \cos(\theta) \end{cases}$$

or $x = r \cos(\alpha)$ et $y = r \sin(\alpha)$

2e étape : Rotation « indicative »

Step 2. Rotate *points* so that their indicative angle is at 0° .

ROTATE-TO-ZERO(*points*)

- 1 $c \leftarrow \text{CENTROID}(\textit{points})$ // computes (\bar{x}, \bar{y})
- 2 $\theta \leftarrow \text{ATAN}(c_y - \textit{points}_{0_y}, c_x - \textit{points}_{0_x})$ // for $-\pi \leq \theta \leq \pi$
- 3 $\textit{newPoints} \leftarrow \text{ROTATE-BY}(\textit{points}, -\theta)$
- 4 **return** *newPoints*

ROTATE-BY(*points*, θ)

- 1 $c \leftarrow \text{CENTROID}(\textit{points})$
- 2 **foreach** point *p* in *points* **do**
- 3 $q_x \leftarrow (p_x - c_x) \cos \theta - (p_y - c_y) \sin \theta + c_x$
- 4 $q_y \leftarrow (p_x - c_x) \sin \theta + (p_y - c_y) \cos \theta + c_y$
- 5 APPEND(*newPoints*, *q*)
- 6 **return** *newPoints*

3e étape : mise à l'échelle et translation

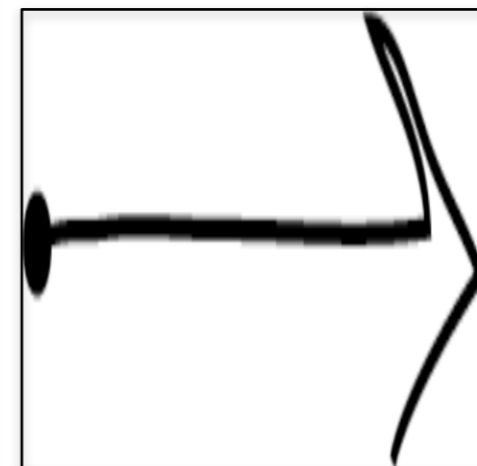
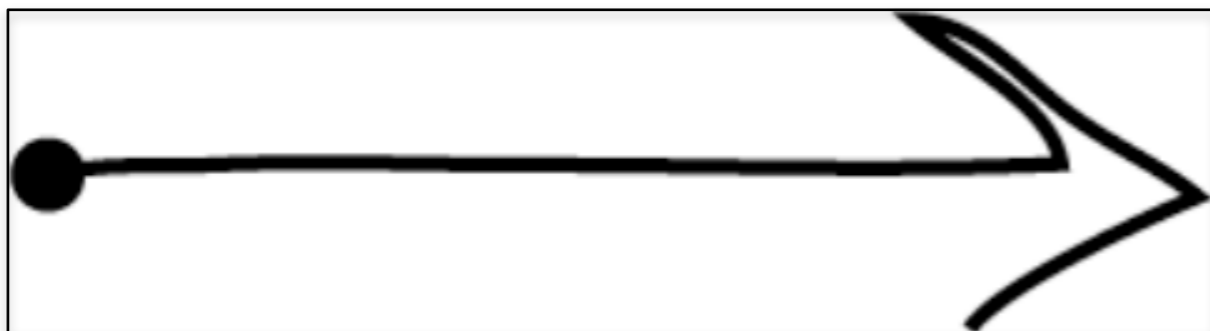
Mise à l'échelle non uniforme: on ramène le geste à un carré de référence

A. Calcul de la bounding box

- ▶ Calcul de minx, maxx, miny, maxy

B. Mise à l'échelle

C. Translation à l'origine



3e étape : mise à l'échelle et translation

Step 3. Scale *points* so that the resulting bounding box will be of $size^2$ dimension; then translate *points* to the origin. BOUNDING-BOX returns a rectangle according to (min_x, min_y) , (max_x, max_y) . For gestures serving as templates, Steps 1-3 should be carried out once on the raw input points. For candidates, Steps 1-4 should be used just after the candidate is articulated.

SCALE-TO-SQUARE(*points*, *size*)

```
1   $B \leftarrow \text{BOUNDING-BOX}(points)$ 
2  foreach point  $p$  in points do
3     $q_x \leftarrow p_x \times (size / B_{width})$ 
4     $q_y \leftarrow p_y \times (size / B_{height})$ 
5    APPEND(newPoints,  $q$ )
6  return newPoints
```

TRANSLATE-TO-ORIGIN(*points*)

```
1   $c \leftarrow \text{CENTROID}(points)$ 
2  foreach point  $p$  in points do
3     $q_x \leftarrow p_x - c_x$ 
4     $q_y \leftarrow p_y - c_y$ 
5    APPEND(newPoints,  $q$ )
6  return newPoints
```

4e étape : reconnaissance

Step 4. Match *points* against a set of *templates*. The *size* variable on line 7 of RECOGNIZE refers to the *size* passed to SCALE-TO-SQUARE in Step 3. The symbol φ equals $\frac{1}{2}(-1 + \sqrt{5})$. We use $\theta = \pm 45^\circ$ and $\theta_\Delta = 2^\circ$ on line 3 of RECOGNIZE. Due to using RESAMPLE, we can assume that A and B in PATH-DISTANCE contain the same number of points, i.e., $|A|=|B|$.

RECOGNIZE(*points*, *templates*)

```
1   $b \leftarrow +\infty$ 
2  foreach template  $T$  in templates do
3     $d \leftarrow$  DISTANCE-AT-BEST-ANGLE(points,  $T$ ,  $-0$ ,  $0$ ,  $\theta_\Delta$ )
4    if  $d < b$  then
5       $b \leftarrow d$ 
6       $T' \leftarrow T$ 
7   $score \leftarrow 1 - b / 0.5\sqrt{(size^2 + size^2)}$ 
8  return  $\langle T', score \rangle$ 
```

4e étape : reconnaissance

Un geste candidat C est comparé à chaque templates T_i

- Calcul de la distance moyenne d_i entre les points

Le tel

$$d_i = \frac{\sum_{k=1}^N \sqrt{(C[k]_x - T_i[k]_x)^2 + (C[k]_y - T_i[k]_y)^2}}{N}$$

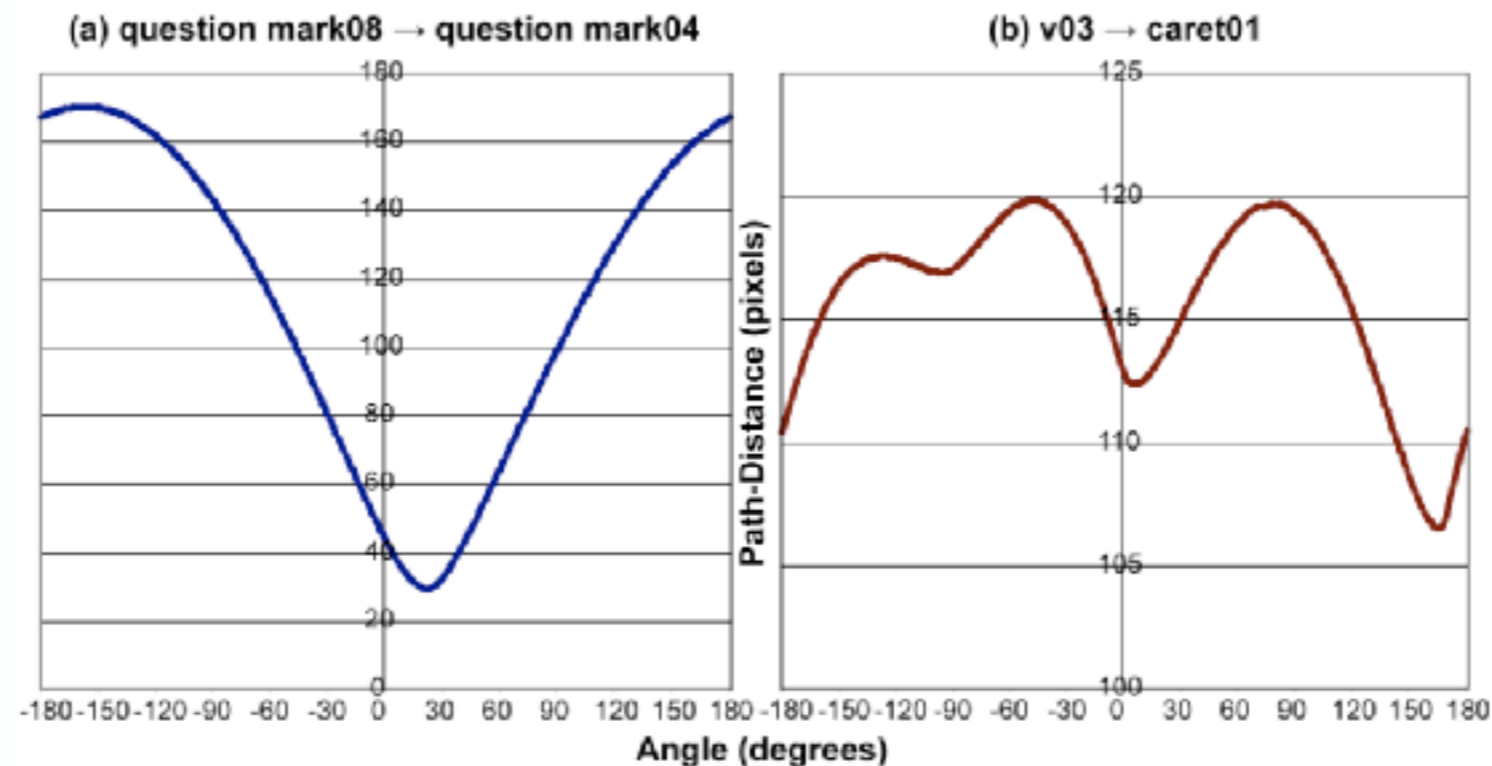
La dis

$$score = 1 - \frac{d_i^*}{\frac{1}{2} \sqrt{size^2 + size^2}}$$

4e étape : reconnaissance

L'angle indicatif ne garantit pas que le geste candidat C sera parfaitement aligné avec un template

On cherche à ajuster l'angle de rotation de C pour minimiser la distance entre C et T_i



4e étape : reconnaissance

Golden Section Search :

- Recherche d'une valeur minimale pour une fonction unimodale

Similaire à la recherche dichotomique

$$a + c = b$$

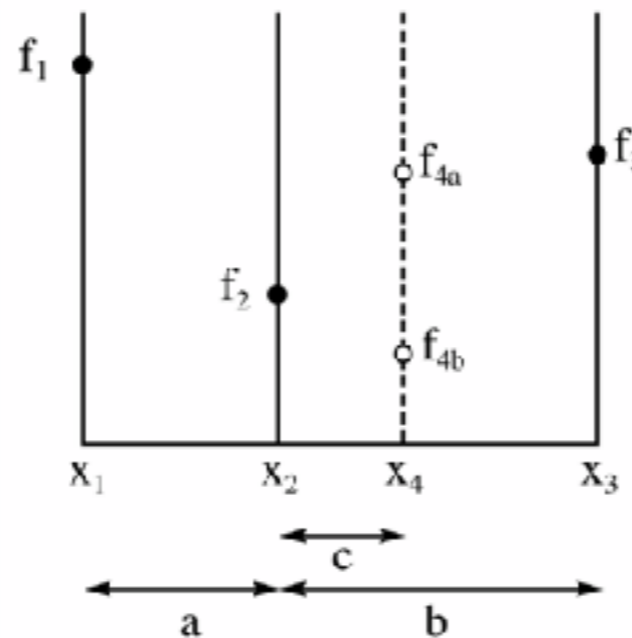
$$c/a = a/b$$

$$\varphi = b/a$$

$$\varphi^2 - \varphi - 1 = 0$$

$$\varphi_1 = (1 + \sqrt{5})/2$$

$$\varphi_2 = (-1 + \sqrt{5})/2$$



$$x_2 = (1 - \varphi)x_1 + \varphi x_4$$

$$x_3 = \varphi x_1 + (1 - \varphi)x_4$$

4e étape : reconnaissance

```
DISTANCE-AT-BEST-ANGLE(points, T,  $\theta_a$ ,  $\theta_b$ ,  $\theta_\Delta$ )
1   $x_1 \leftarrow \varphi\theta_a + (1 - \varphi)\theta_b$ 
2   $f_1 \leftarrow \text{DISTANCE-AT-ANGLE}(\textit{points}, T, x_1)$ 
3   $x_2 \leftarrow (1 - \varphi)\theta_a + \varphi\theta_b$ 
4   $f_2 \leftarrow \text{DISTANCE-AT-ANGLE}(\textit{points}, T, x_2)$ 
5  while  $|\theta_b - \theta_a| > \theta_\Delta$  do
6    if  $f_1 < f_2$  then
7       $\theta_b \leftarrow x_2$ 
8       $x_2 \leftarrow x_1$ 
9       $f_2 \leftarrow f_1$ 
10      $x_1 \leftarrow \varphi\theta_a + (1 - \varphi)\theta_b$ 
11      $f_1 \leftarrow \text{DISTANCE-AT-ANGLE}(\textit{points}, T, x_1)$ 
12   else
13      $\theta_a \leftarrow x_1$ 
14      $x_1 \leftarrow x_2$ 
15      $f_1 \leftarrow f_2$ 
16      $x_2 \leftarrow (1 - \varphi)\theta_a + \varphi\theta_b$ 
17      $f_2 \leftarrow \text{DISTANCE-AT-ANGLE}(\textit{points}, T, x_2)$ 
18 return  $\text{MIN}(f_1, f_2)$ 
```


4e étape : reconnaissance

```
DISTANCE-AT-ANGLE(points, T,  $\theta$ )
1  newPoints  $\leftarrow$  ROTATE-BY(points,  $\theta$ )
2  d  $\leftarrow$  PATH-DISTANCE(newPoints, Tpoints)
3  return d

PATH-DISTANCE(A, B)
1  d  $\leftarrow$  0
2  for i from 0 to |A| step 1 do
3    d  $\leftarrow$  d + DISTANCE(Ai, Bi)
4  return d / |A|
```

Limitations

Pas possible de distinguer un carré d'un rectangle

Pas possible de distinguer une ellipse d'un cercle

Pas possible de distinguer une flèche vers le bas/haut

Pas possible de reconnaître des gestes « 1D »

Classifieur statistique

Rubine

Reconnaissance de gestes à un seul tracé (unistroke)

Traitement statistique de caractéristiques

Dean Rubine
Specifying gestures by example
SIGGRAPH '91, 329-337

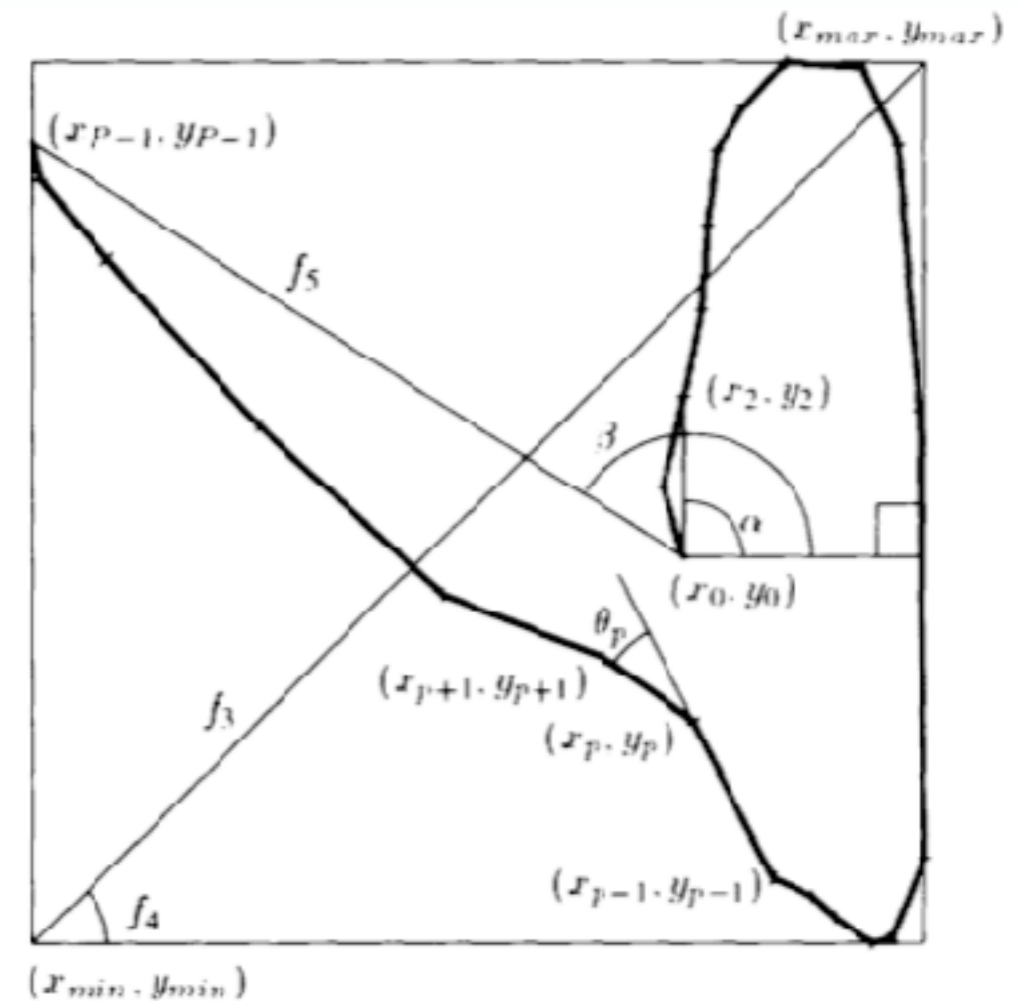
Rubine

Liste de points en entrée

- ▶ Élimination des points trop proches.
 - $d < 3$ pixels du point précédent
- ▶ Calcul d'un vecteur de caractéristiques statistiques.
- ▶ Comparaison aux gestes de référence.
 - Le geste avec le score le plus grand est renvoyé
- ▶ Détection des gestes ambigus

Caractéristiques

- A. cos de l'angle d'origine
- B. sin de l'angle d'origine
- C. longueur diagonale bounding box
- D. angle diagonale bounding box
- E. distance entre premier et dernier point
- F. cos de l'angle entre le premier et dernier point
- G. sin de l'angle entre le premier et dernier point
- H. longueur totale du tracé
- I. angle total traversé
- J. somme des angles absolus à chaque point
- K. somme des carrés de ces angles
- L. vitesse² maximum du geste
- M. durée du geste



Initialisation

Calcul des statistiques moyennes pour chaque caractéristique de chaque geste

Calcul de la matrice de covariance moyenne de tous les gestes

Objectif : trouver les coefficients pondérateurs des caractéristiques statistiques qui permettent de maximiser le score des gestes de chaque classe

Reconnaissance

Calcul des caractéristiques du geste candidat

Calcul d'une probabilité de correspondance pour chaque geste de référence

Le geste avec la probabilité la plus importante est choisi

Dynamic Time Warping

DTW

Déformation temporelle dynamique

Déterminer pour chaque élément d'une séquence, le meilleur élément correspondant dans l'autre séquence relativement à un certain voisinage et à une certaine métrique

Complexité polynomiale

Applications

Vidéo, audio, graphique...

Toutes données qui peuvent être transformées en représentation linéaire en fonction du temps (séries temporelles)

Echantillons ordonnés par une étiquette de temps

Reconnaissance vocale

Reconnaissance de gestes off-line et on-line

Alignement de protéines...

Principe de base

Séquence de référence $R = [r_1, r_2, \dots, r_n]$

Séquence de test $T = [t_1, t_2, \dots, t_m]$

Si $m = n$ alors on peut calculer la distance deux à deux:

$$D = \sum_{i=1}^n \text{distance}(r_i, t_i)$$

Possibilité de calculer la distance euclidienne ou d'utiliser une autre métrique (fonction qui donne un réel)

Plus possible d'utiliser cette méthode dès que $n \neq m$

Attention : les échantillons doivent être équidistants en temps

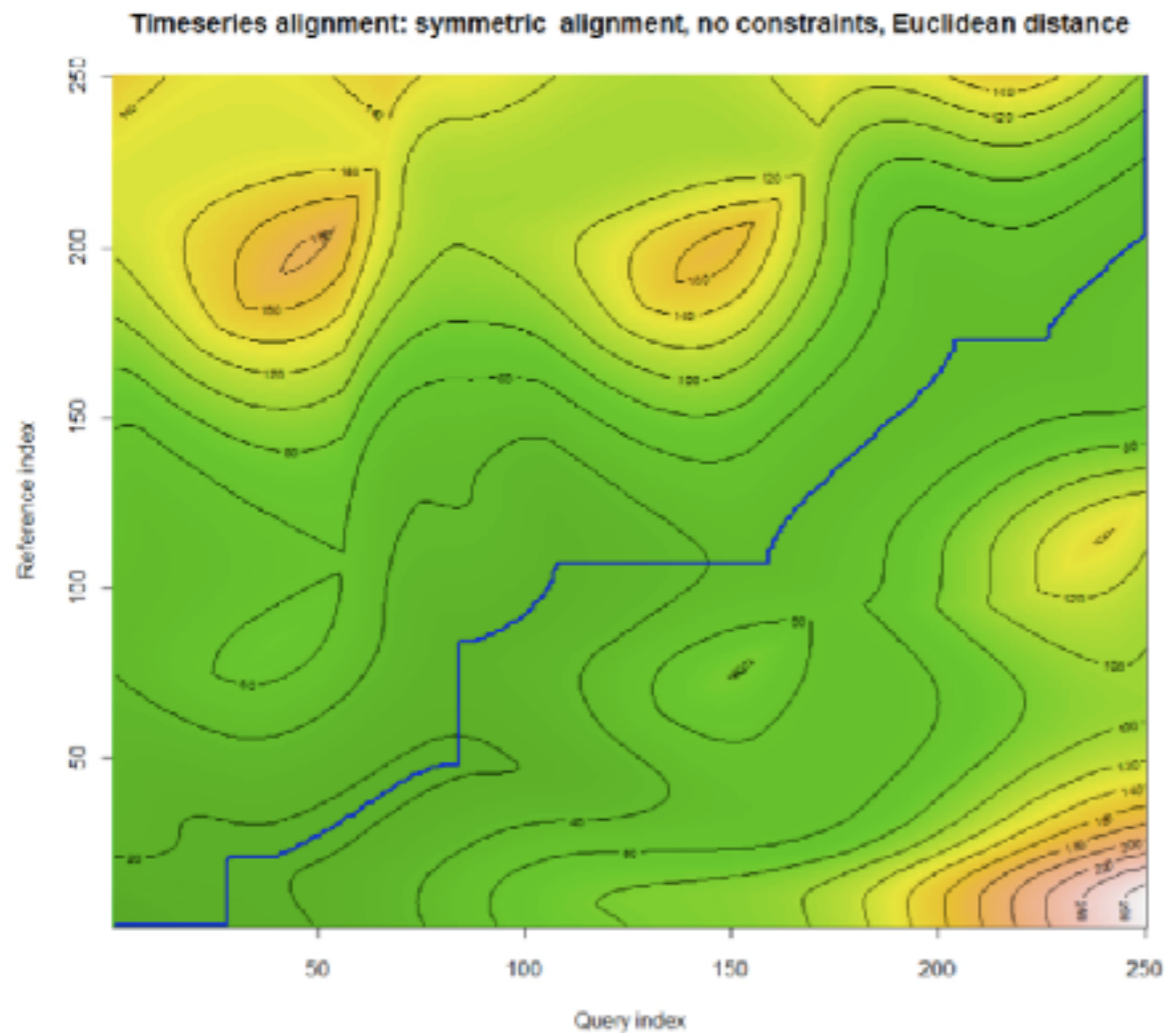
Principe de base

DTW réalise d'abord un alignement non linéaire en recherchant parmi tous les alignements possibles, celui qui minimise une fonction de coût cumulé

« Time Warping » : Dilation ou compression des séquence pour obtenir le meilleur alignement possible

Principe de base

Calcul du chemin $W = [w_1, w_2, \dots, w_k]$ de longueur minimale



$$\sum_{i=1}^k distance(w_i)$$

Principe de base

Conditions aux frontières

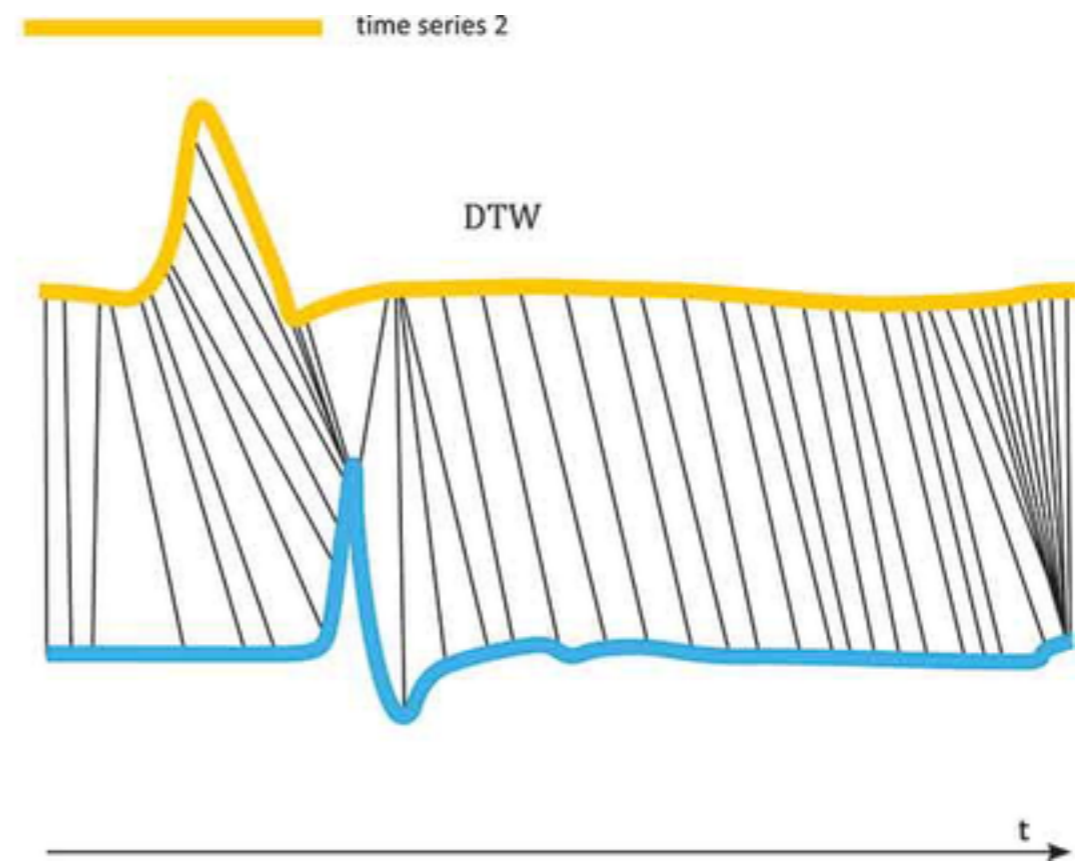
- ▶ $w_1 = (r_1, t_1)$
- ▶ $w_k = (r_n, t_m)$

Contraintes locales

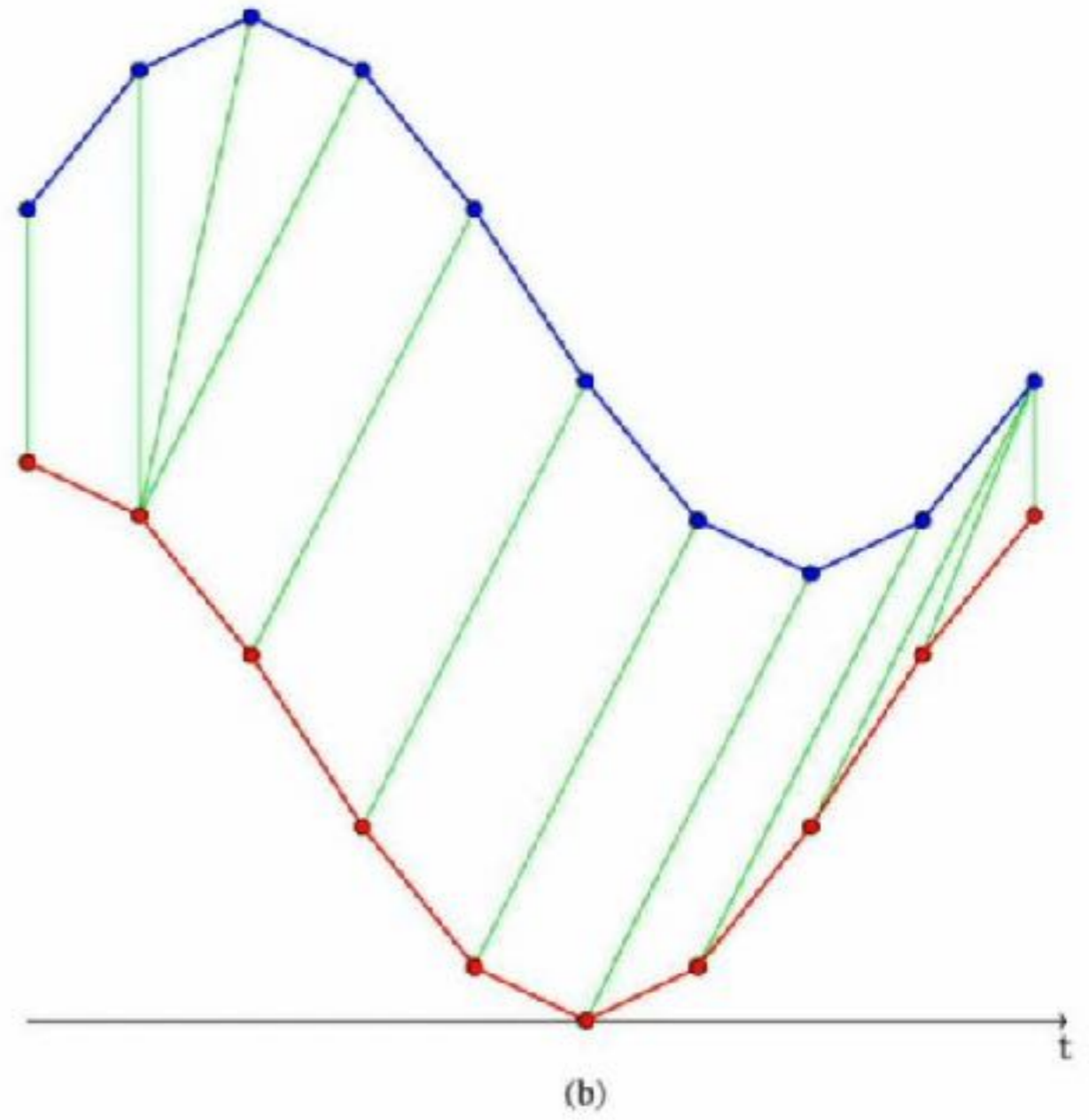
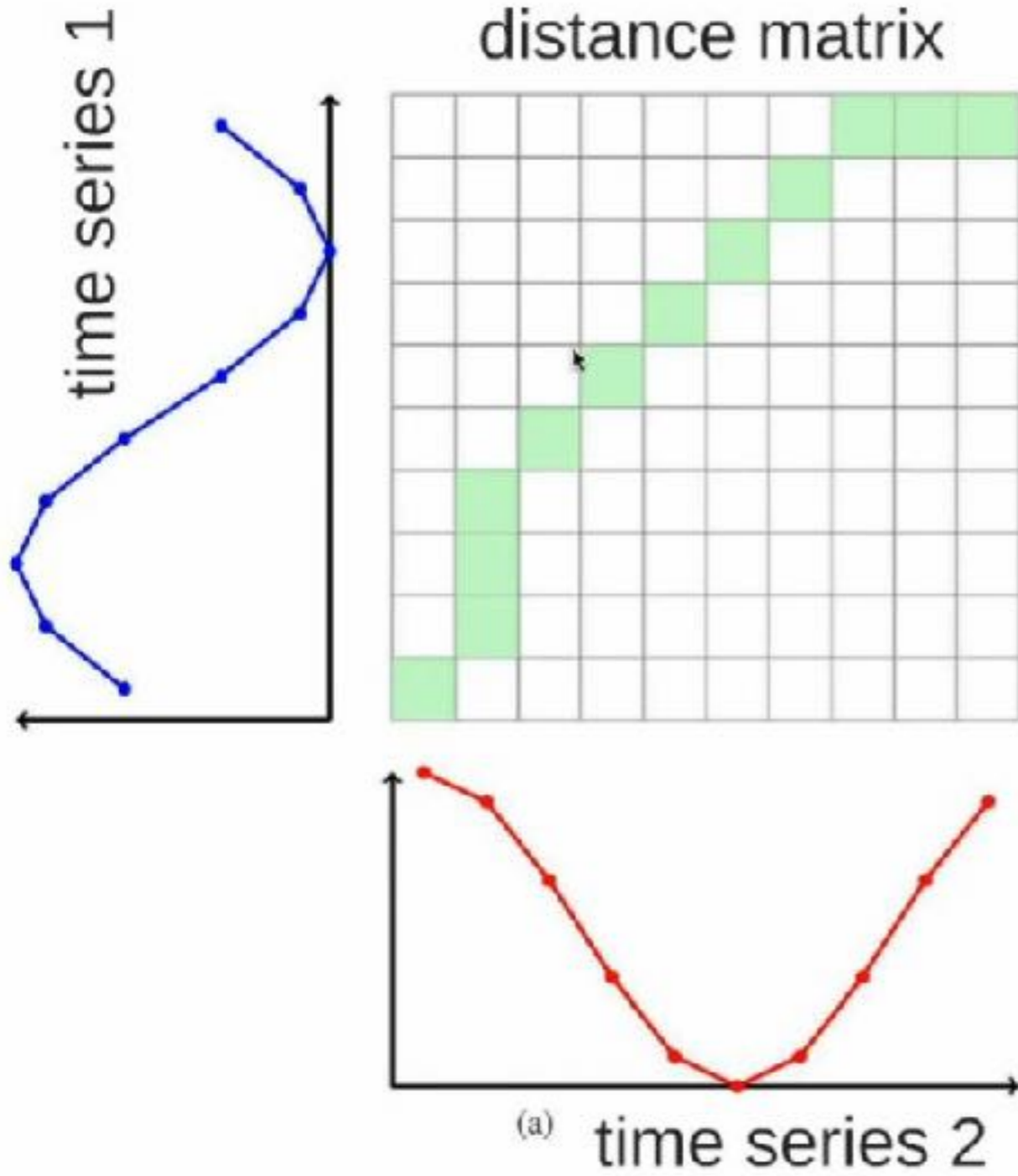
- ▶ Monotonie pour respecter le séquençement des points
- ▶ Eviter les sauts dans le temps
- ▶ Pour tout couple (r_i, t_j) , le choix des prédécesseurs est limité à (r_{i-1}, t_j) , (r_i, t_{j-1}) , (r_{i-1}, t_{j-1})

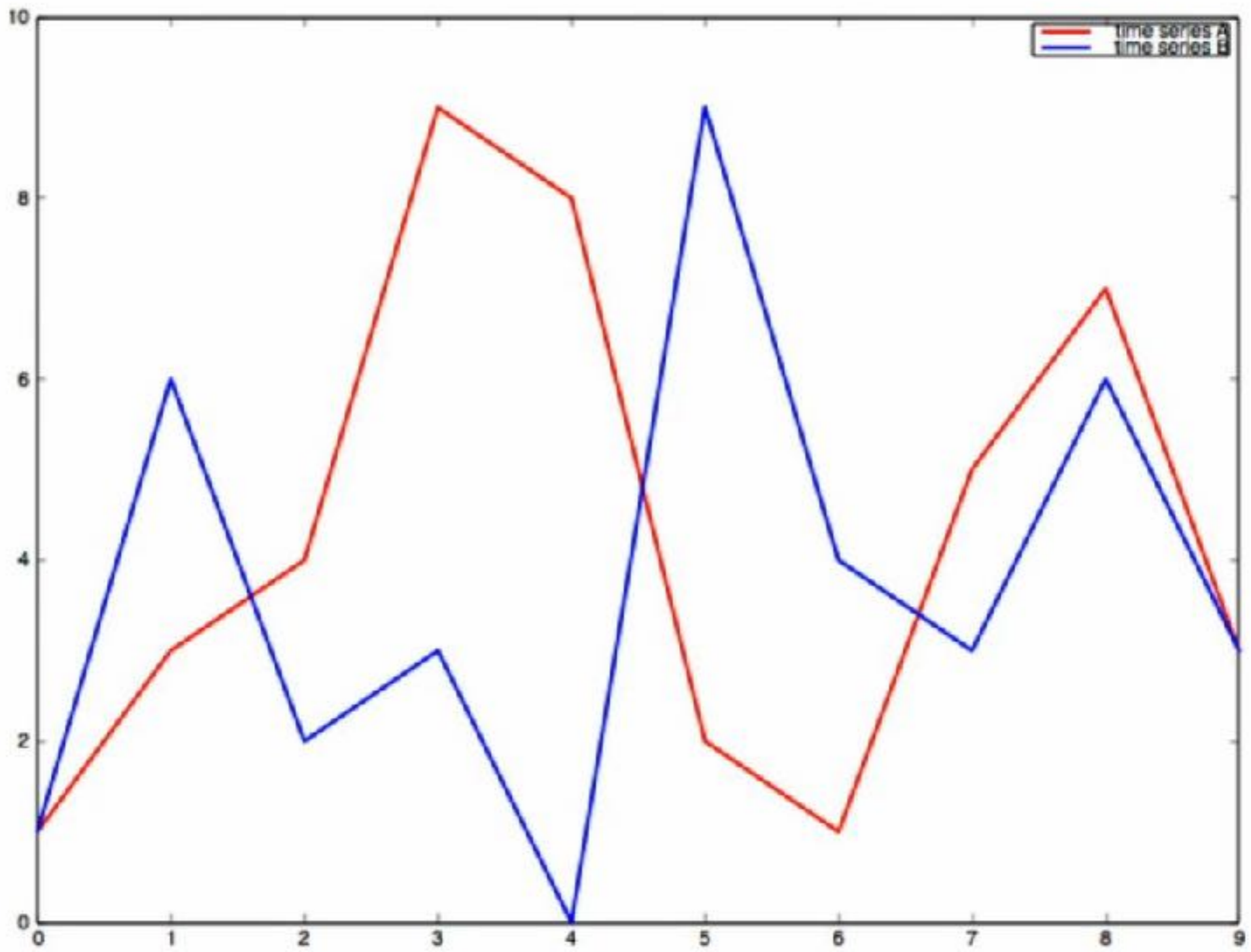
Exhaustivité

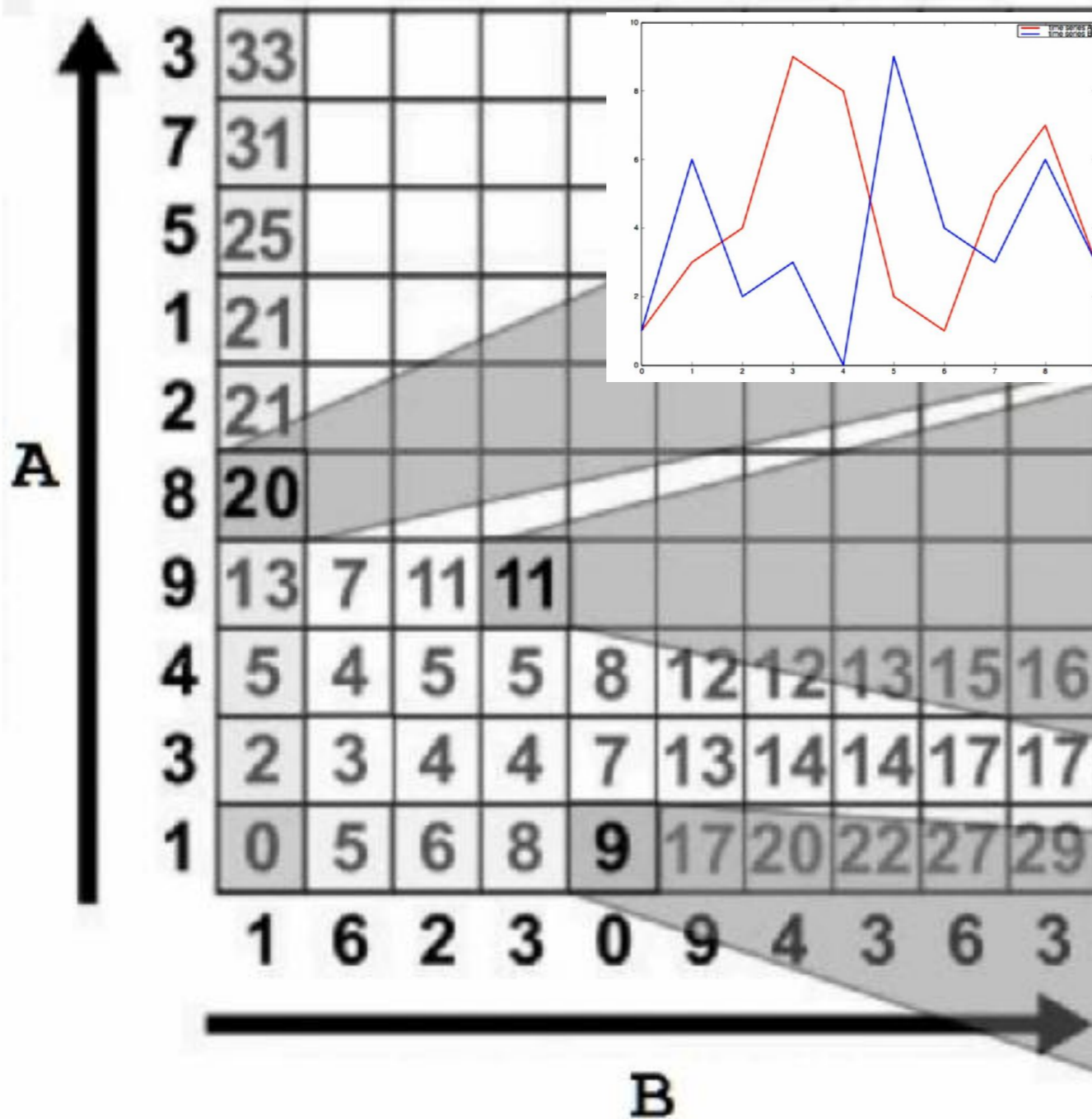
- ▶ Chaque élément de R doit être mis en relation avec au moins un élément de T et vice-versa
- ▶ $\max(m, n) \leq k \leq m + n - 1$



Principe de base



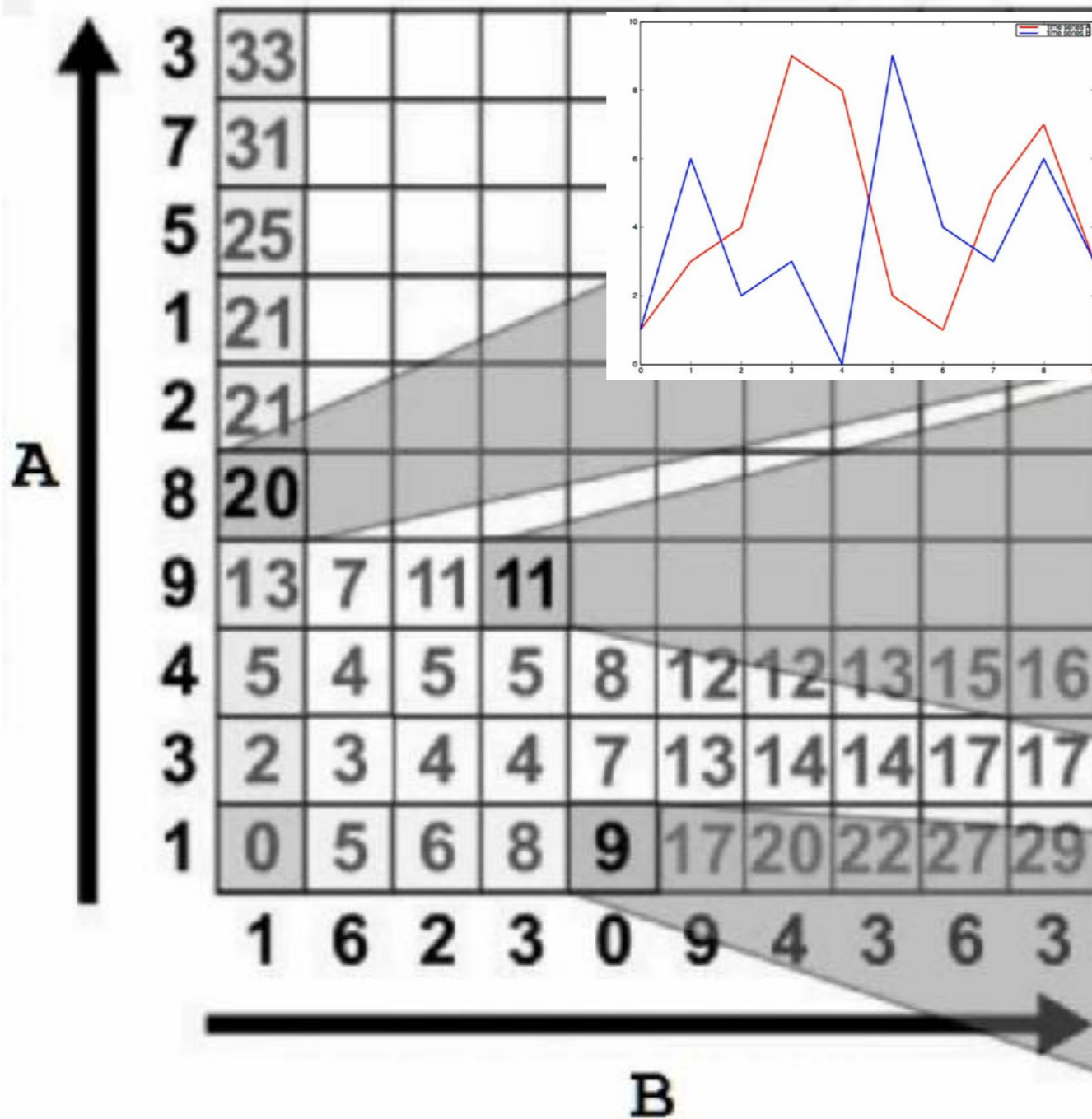




$$\begin{aligned}
 &= |A_i - B_j| + D[i-1, 0] \\
 &= |8 - 1| + 13 \\
 &= 20
 \end{aligned}$$

$$\begin{aligned}
 &= |A_i - B_j| + \min(D[i-1, j-1], \\
 &\quad D[i-1, j], \\
 &\quad D[i, j-1]) \\
 &= |9 - 3| + \min(5, 5, 11) \\
 &= 6 + 5 \\
 &= 11
 \end{aligned}$$

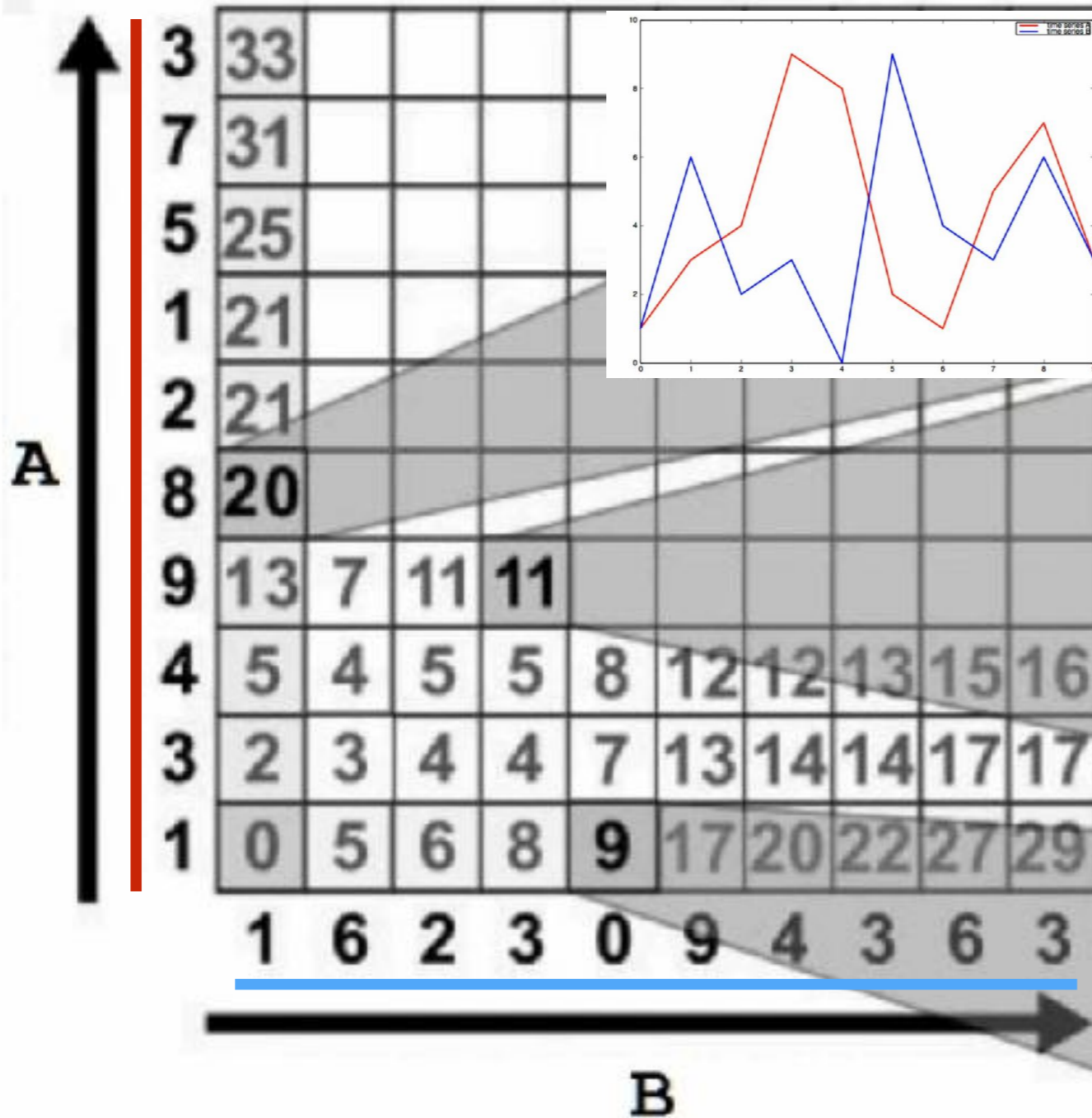
$$\begin{aligned}
 &= |A_i - B_j| + D[0, j-1] \\
 &= |1 - 0| + 8 \\
 &= 9
 \end{aligned}$$



$$\begin{aligned}
 &= |A_i - B_j| + D[i-1, 0] \\
 &= |8 - 1| + 13 \\
 &= 20
 \end{aligned}$$

$$\begin{aligned}
 &= |A_i - B_j| + \min(D[i-1, j-1], \\
 &\quad D[i-1, j], \\
 &\quad D[i, j-1]) \\
 &= |9 - 3| + \min(5, 5, 11) \\
 &= 6 + 5 \\
 &= 11
 \end{aligned}$$

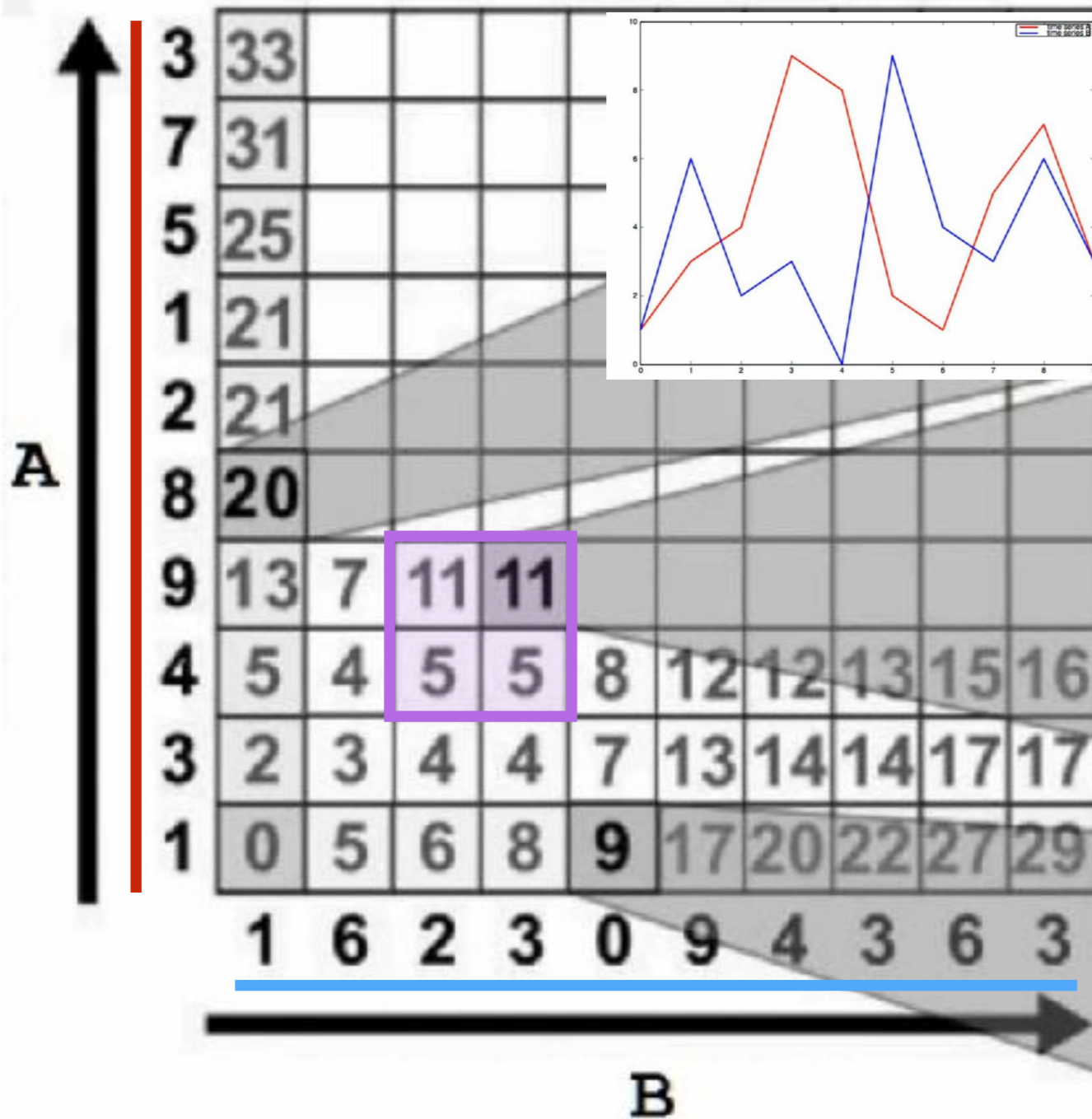
$$\begin{aligned}
 &= |A_i - B_j| + D[0, j-1] \\
 &= |1 - 0| + 8 \\
 &= 9
 \end{aligned}$$



$$\begin{aligned}
 &= |A_i - B_j| + D[i-1, 0] \\
 &= |8 - 1| + 13 \\
 &= 20
 \end{aligned}$$

$$\begin{aligned}
 &= |A_i - B_j| + \min(D[i-1, j-1], \\
 &\hspace{10em} D[i-1, j], \\
 &\hspace{10em} D[i, j-1]) \\
 &= |9 - 3| + \min(5, 5, 11) \\
 &= 6 + 5 \\
 &= 11
 \end{aligned}$$

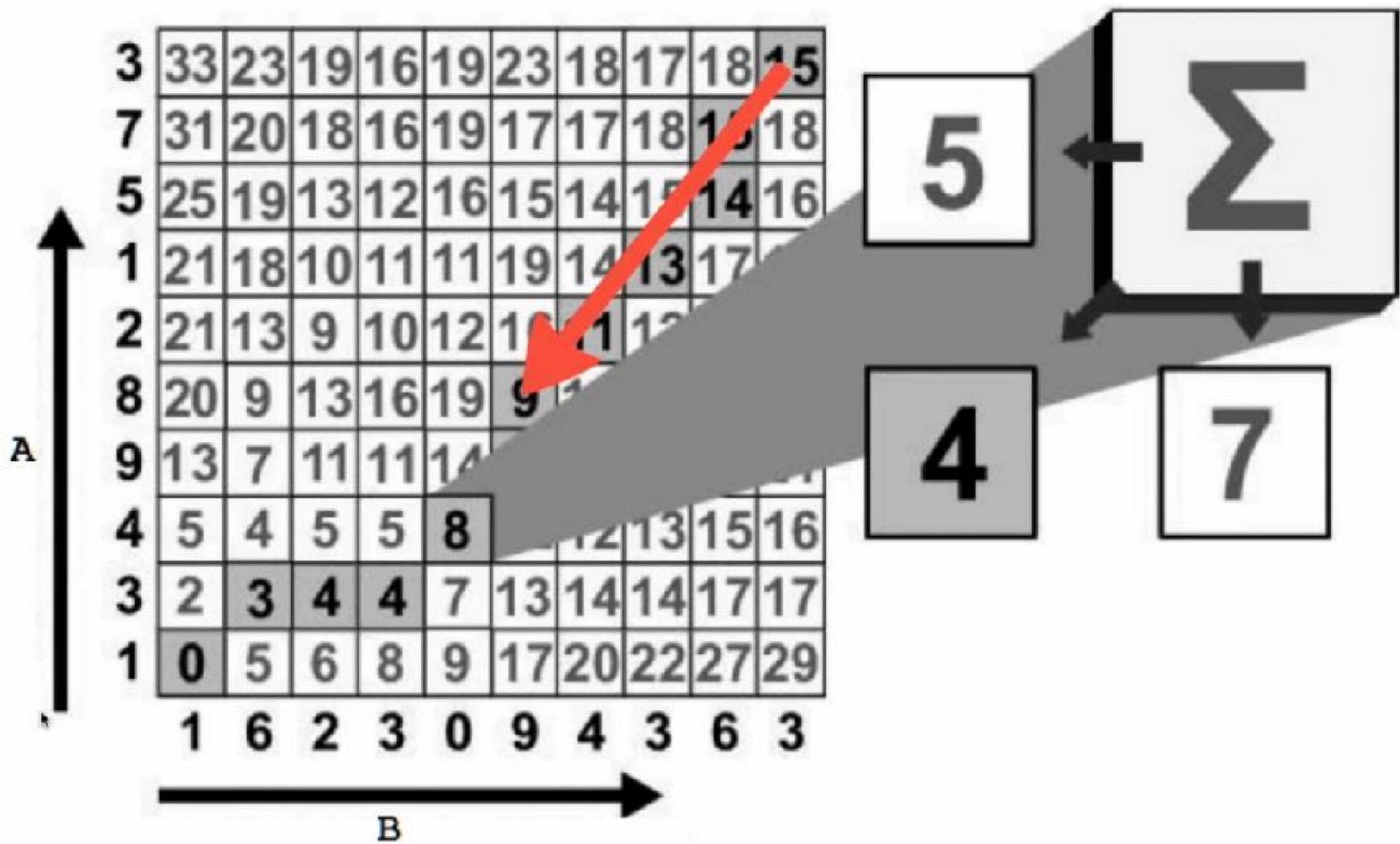
$$\begin{aligned}
 &= |A_i - B_j| + D[0, j-1] \\
 &= |1 - 0| + 8 \\
 &= 9
 \end{aligned}$$



$$\begin{aligned}
 &= |A_i - B_j| + D [i-1, 0] \\
 &= |8 - 1| + 13 \\
 &= 20
 \end{aligned}$$

$$\begin{aligned}
 &= |A_i - B_j| + \min(D [i-1, j-1], \\
 &\quad D [i-1, j], \\
 &\quad D [i, j-1]) \\
 &= |9 - 3| + \min(5, 5, 11) \\
 &= 6 + 5 \\
 &= 11
 \end{aligned}$$

$$\begin{aligned}
 &= |A_i - B_j| + D [0, j-1] \\
 &= |1 - 0| + 8 \\
 &= 9
 \end{aligned}$$



Reconnaissance de gestes

Le geste fait généralement l'objet de variance dans sa reproduction

Un seul geste exemple n'est souvent pas suffisant

Définition d'un modèle statistique tenant compte de ces variations

Utilisation de plusieurs gestes exemples

➔ Calcul de la moyenne et de la variance en fonction du temps

Reconnaissance de gestes

Chaque exemple est comparé à l'exemple le plus long avec DTW, avant de calculer les caractéristiques statistiques

Calcul d'un modèle de geste g pour chaque classe d'exemples

- ▶ Calcul de la moyenne et de la variance à chaque pas de temps :

$$\tilde{g}_m[t] \text{ et } \sigma^2(g_m[t])$$

Reconnaissance de gestes

Un geste candidat est aligné avec chaque modèle en utilisant DTW

A chaque nouvel événement, la séquence de test T est comparée à chaque modèle en utilisant la métrique suivante :

$$D_{i,j} = \sum \frac{1}{\sigma^2(g_m[t])} (\tilde{g}_m[t] - T[j])^2$$

Reconnaissance de gestes

Alignement de la séquence T avec les modèles

Retour dans le temps, présent = 0

Relaxation des contraintes

Longueur minimale pour la séquence T

Le score du geste est l'inverse de la distance calculée par DTW

Le geste avec le score le plus élevé est le geste reconnu

URL TP

<http://www.malacria.com/teaching/telecomlille/tpgestures>