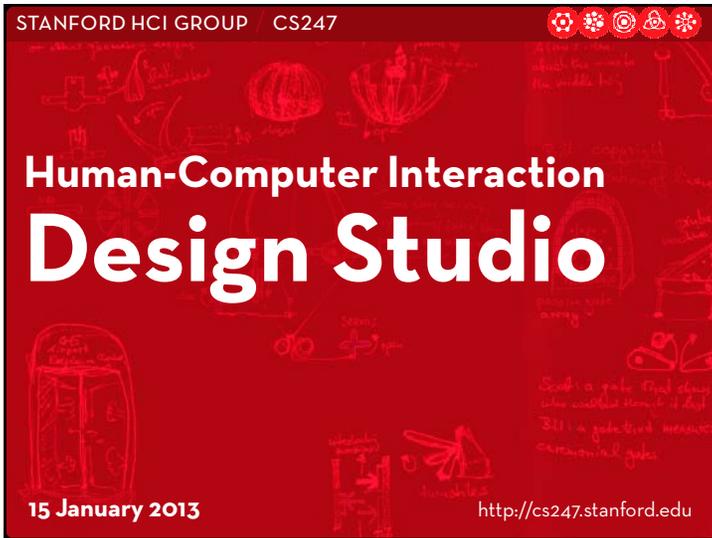


STANFORD HCI GROUP / CS247

Human-Computer Interaction
Design Studio

15 January 2013

<http://cs247.stanford.edu>



What was the first
“gestural interface”?



Myron Krueger

There were things I resented about computers.
I resented that I had to sit down to use them.
... that it was denying that I had a body.
... that it wasn't perceptual – it was all symbolic.

I started thinking that artists and musicians had
the best relationships to their tools.

As early as '74, the computer could see you.

[Krueger 1988](#)

P2: Shadow Boxing

“Experience a computer; don’t learn to use it”

Draw inspiration from prior work:

- Manipulate physical environment to enhance experience or sensing (bright lights, audio, ...).
- Manipulate virtual environment (add virtual objects).
- Explore potentially ambiguous input/output relationships - *without “deep” recognition*. For example: optical flow, regions of activity, etc.
- Add sensor channels: depth camera, microphone, ...

TOPICS

“Natural” User Interfaces

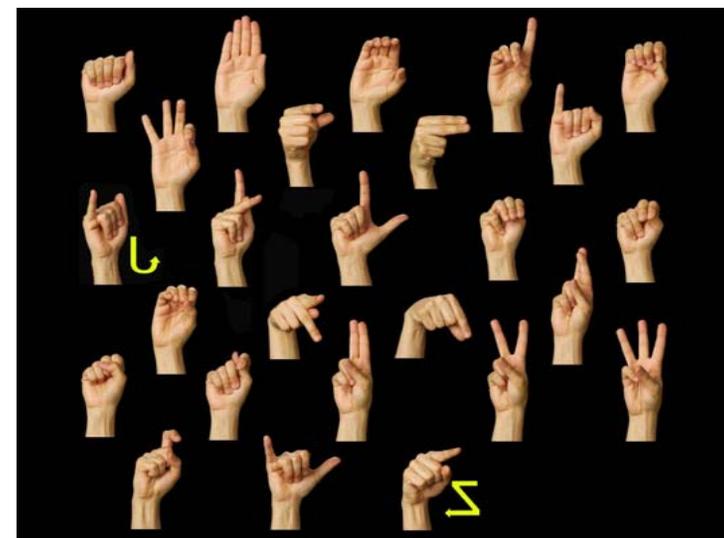
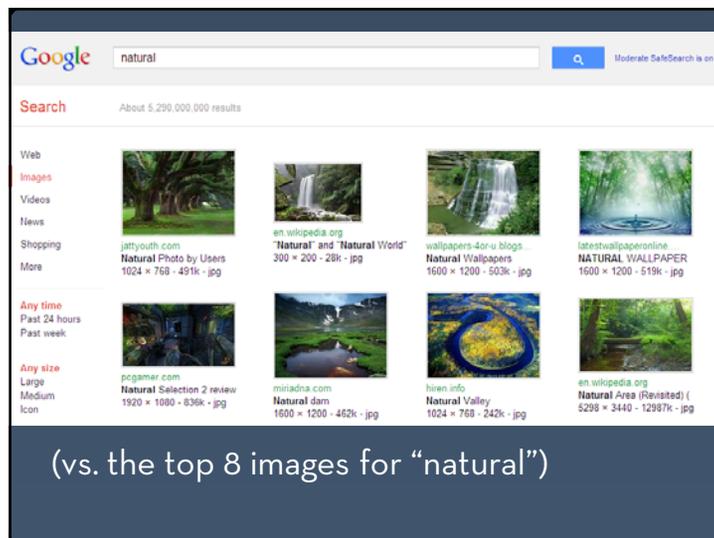
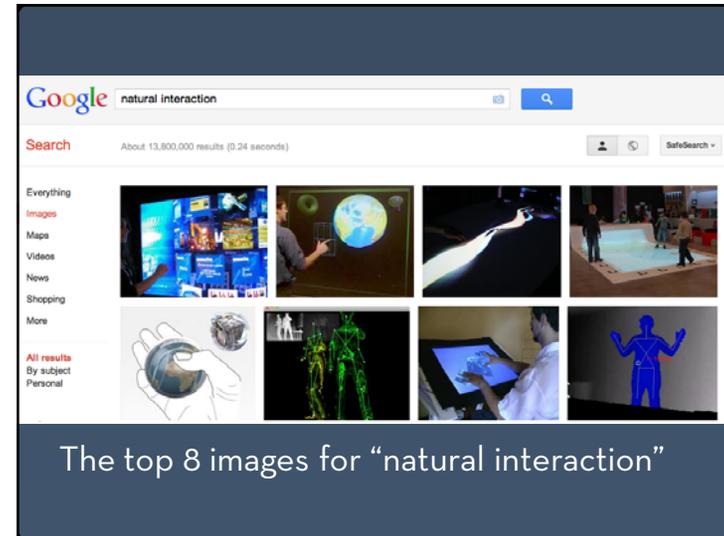
Deixis & Proxemics

Gesture Input Technology

Gesture Design

“Natural”
User Interfaces

What makes an input method “natural”?



This is an **ill-posed question!**

A reasonable working definition?

A user interface is “*natural*” if:

The experience of using a system matches expectations, such that it is always clear to the user how to proceed, and that few steps (with a minimum of physical and cognitive effort) are required to complete common tasks.

Hinckley & Wigdor

Wait... isn't this just *usability* by another name?

It is a **common mistake** to attribute the naturalness of a product to the **underlying input technology**. A touch screen, or any other input method for that matter, is not inherently natural.

Hinckley & Wigdor

“Fluent” experiences depend on the **context and expectations** of the user, often relying on prior **learning** and **skill acquisition**.

What do we “do”
with gestures and
body posture?

Deixis

You, here, now!

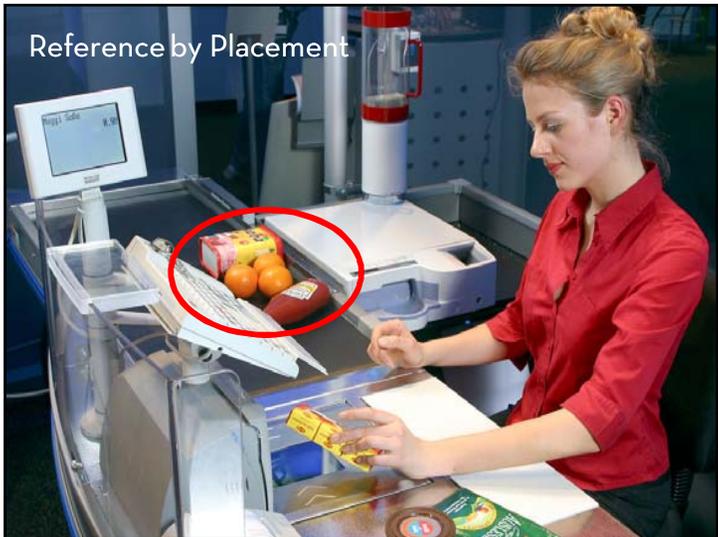


Deixis: referencing the world

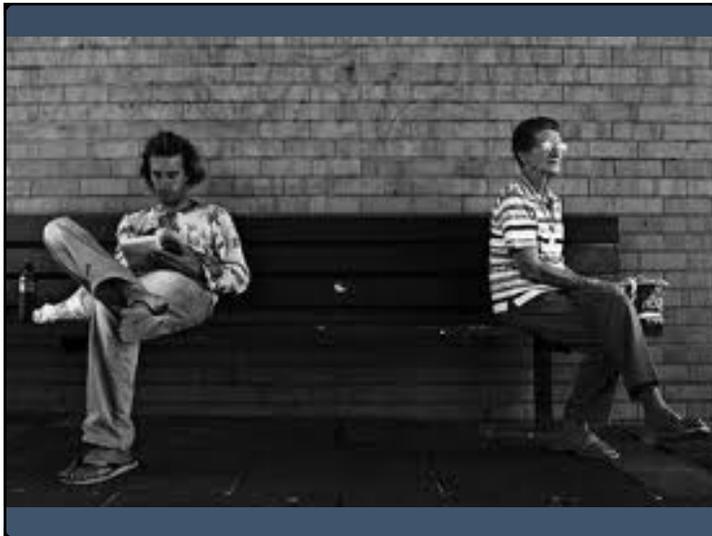
We continuously reference elements in the world in ambiguous ways, yet for the most part we seem to convey our intentions quite well.

Deixis: Reference by means of an expression whose interpretation is relative to the (usually) extralinguistic context.

Common methods of physical reference:
pointing & placing [Clark 2003]



Proxemics





Proxemics

Proxemics is the study of measurable distances between people as they interact. [Hall 1966]

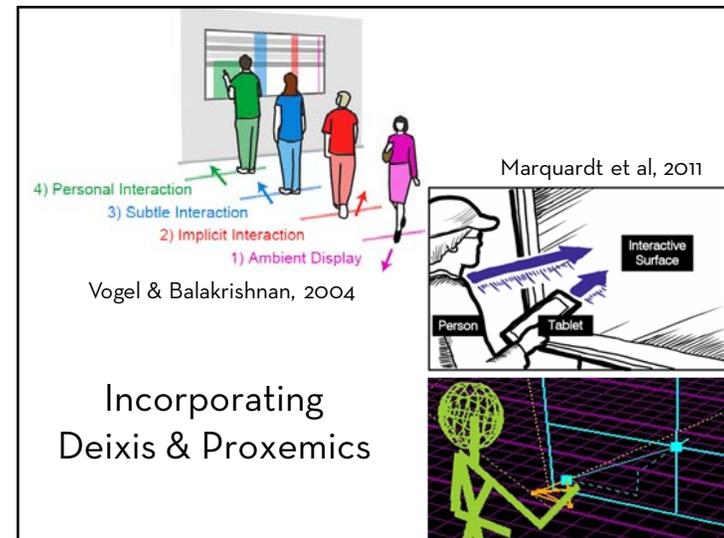
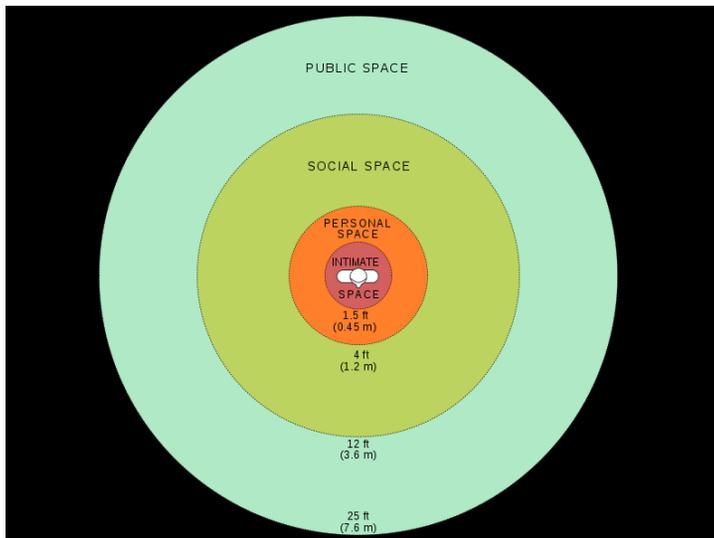
Taxonomy of Distance:

Intimate: embracing, touching or whispering

Personal: interaction among friends / family

Social: interactions among acquaintances

Public: distance used for public speaking



Gesture Input Technologies

Kinect Sensor



Depth Cameras

Structured IR light

- ✓ **cheap, fast, accurate**
- ✗ **missing pixels, shadows**

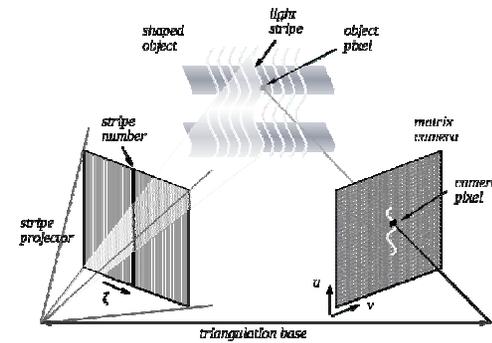


missing pixels (not IR reflective)



How Kinect Works

Structured Light 3D Scanner



RGB vs. Depth for Pose Estimation

RGB

- Only works when well lit
- Background clutter
- Scale unknown
- Clothing, skin colour

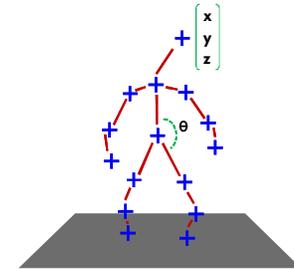
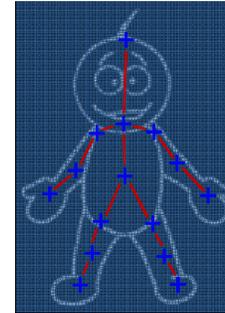
Depth

- Works in low light
- Person 'pops' out from bg
- Scale known
- Uniform texture
- Shadows, missing pixels



Much easier
with depth!

Human Pose Estimation



Kinect tracks 20 body joints in real time.

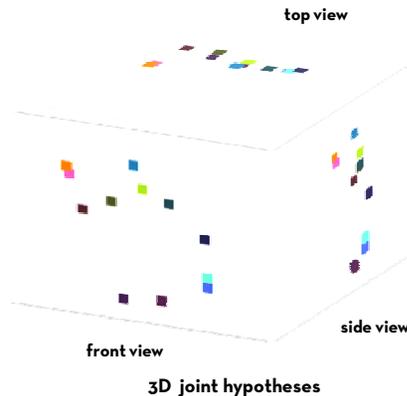
Skeletal Tracking



Input depth image



Inferred body parts &
overlaid joint hypotheses



3D joint hypotheses

Kinect SDK

Input

Image Data Streams: RGB, Depth images
Skeletal Tracking
Audio (Microsoft Speech Platform)

Constraints

Latency: data analysis introduces lag
86cm to 4m range
Not outdoors (too much IR noise)
Not too close to other Kinects (IR interference)
Track 1-2 people only; full bodies must be in view (?)



Designing Gestural Interfaces

Designing Gestural UIs

A designer must consider:

(a) the physical **sensor**

Input Device Properties

Property Sensed: position, force, angle, joints

States Sensed: contact, hover, ...

Precision: accuracy of selection

Latency: delay in property/state sensing

Acquisition Time: get pen, move hand to mouse

False Input: accidental touches

Of “clutches” and “live mics”

| Device | Property | State Tracked |
|----------------|----------------|---------------------|
| Mouse | 2D Position | Hover, Button-Press |
| Stylus | 2D Position | Hover, Contact |
| Touch | 2D Position | Contact |
| Gesture | 2D/3D Position | ?? |

In-air gestures may involve a **live mic**, increasing chances of *false positives* and *false negatives*.

Clutch: differentiate actions intended to drive the computing system from those that are not.

Managing a “live mic”

Reserved Actions

Design gestures that will not be triggered unless specifically desired by the user.

Reserved Clutches

Use a special gesture to indicate that the system should now monitor for input commands.

Multi-Modal Input

Use another modality such as buttons or voice input to engage tracking by the system.

Designing Gestural UIs

A designer must consider:

- (a) the physical **sensor**
- (b) the **feedback** presented to the user
- (c) **ergonomic** and industrial design
- (d) the **interplay** between all interaction techniques and among all devices in the surrounding context
- (e) the **learning curve**

Gesture Design Exercise

How to design gestures?

Observation: generate potential gestures by observing (and participating in) situated activity.

Participatory design: have representative users generate potential gestures for you.

One methodology [Wobbrock et al 2009]:

1. Show participant start and end states of UI
2. Participant performs gesture for that effect
3. Analyze collected gestures from population

Must still consider **interplay** with task/context!



Design Exercise

Context: *virtual post-its* - the primary interface elements are movable, resizable squares.

Your task: design a consistent touch gesture vocabulary for a set of operations.

You may assume that:

- (a) Users can use both of their hands.
- (b) The system identifies the hands/fingers being used.
- (c) You may introduce additional widgets or graphical elements as part of your vocabulary.

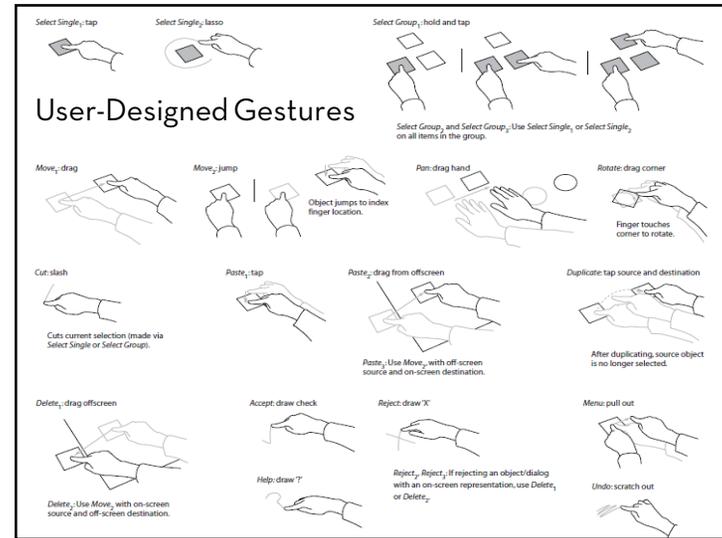
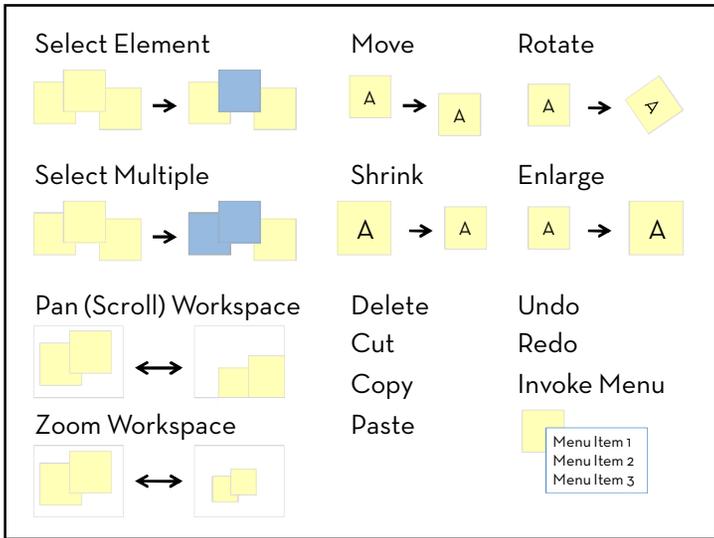
Design Exercise

Overview:

- 5 min Individually develop your own gestures
- 15 min Share with table, revise as a group
- 15 min Share with class

Consider:

- Learnability
- Mechanics of repeated use
- Consistency / compatibility across operations



Final Thoughts

Leverage the unique opportunities provided by a particular input technology. Don't "shoehorn" new modalities where old techniques excel.

Consider "perceptual" vs. "symbolic" input.

Prevent accidental (vs. intentional) input via unambiguous design and/or clutching.

Respect existing conventions of spatial reference and social use of space.