COMMUNITY DAYS

MISC01 – Gesture Recognition con Kinect

Matteo Valoriani

mvaloriani AT gmail.com
Twitter: @MatteoValoriani

#CDays13 – 27 e 28 febbraio 2013

Grazie a

Sponsor

icubed

intel Software

Microsoft

Technical Conferences

Windows

WPC by DotNet

aspitalia.com

Devleap BridgetheGap

DomusDotNet

DotDotNet

DotNet Liguria

dotnet LOMBARDIA

DotNetSide

sharepointcommunity.it

UGIdotNET

Visual Basic tips&tricks
Dal 1996 la community su VB e VB.NET

WindowServer.it

COMMUNITY DAYS

Follow me on
Twitter or the
Kitten gets it:
@MatteoValoriani
COMMUNITY DAYS

Agenda

- Heuristic Based Gesture Detection
- Exemplar Matching Based Gesture Detection
- Common Problems
- Takeaways

Kinect’s magic

Immersive user experience

=

“Any sufficiently advanced technology is indistinguishable from magic” (Arthur C. Clarke)

Skeleton Data

- Maximum two players tracked at once
- Six player proposals per Kinect
- 20 joints in standard mode
- 10 joints in seated mode

Heuristic Based Gesture Detection

Heuristics

- Experience-based techniques for problem solving, learning, and discovery

- Cost effective
- Helps reconstruct missing information
- Helps compute outcome of a gesture

Select the Right Triggers

- Use skeleton view to analyze whole skeleton behavior
- Use joint view to isolate and analyze specific joints and axis behavior
- Use data sheet view: to get the real numbers

- Not all joints are needed
- Player location in the play area can cause some joints to become occluded

Define Key Stages of a Gesture

- Determine
  - When the gesture begins
  - When the gesture ends
- Determine other key stages
  - Changes in motion direction
  - Pauses
  - …

Be careful!!

- Some players have more energy (or enthusiasm) than others
- Some players will “optimize” their gestures
- Most players will not perform the gesture precisely as intended

Heuristic Based Gesture Detection: HandOnHead

DEMO

Heuristic Based Gesture Detection: FAAST

DEMO

Pros & Cons

PROs

- Easy to understand
- Easy to implement (for simple gestures)
- Easy to debug

CONs

- Challenging to choose best values for parameters
- Doesn’t scale well for variants of same gesture
- Gets challenging for complex gestures
- Challenging to compensate for latency

Recommendation

Use for simple gestures

- Hand wave, Head movement

Exemplar Matching Based Gesture Detection

Gesture Definition

- Define gesture as pre-recorded animations
  - Motion capture animations
  - Record different people doing same gesture
  - Each person doing same gesture multiple times

Exemplar

- Definition: ideal example to compare against
- Pre-recorded animations are exemplars

Exemplar Matching

- Need to compare skeleton frames
  - Define error metric for skeleton
  - Angular difference for each joint in local space
  - Peak Signal to Noise Ratio for whole skeleton

$$MSE = \frac{1}{N}\sum_{1}^{N} \text{Distance}_i^2$$

$$PSNR = 10 * \log_{10}(MAX^2 / MSE)$$

Exemplar Matching

- Search for best matching frames
  - Best matching frame has strongest signal
  - Different classifiers can be used
    - K-Nearest
    - Dynamic Time Warping (DTW)
    - Hidden Markov Models (HMM)

Exemplar Matching

DTW Based Gesture Detection: Swipe

DEMO

Pros & Cons

PROs

- Very complex gestures can be detected
- DTW allows for different speeds
- Can scale for variants of same gesture
- Easy to visualize exemplar matching

CONs

- Requires lots of resources to be robust
- Optimize by reducing exemplar matches

Recommendation

Use for complex context-sensitive dynamic gestures

- Dancing, fitness exercises

User Posture

User posture may affect design of a gesture

Posture Abstraction

Kinect SkeletonData depend to:

- Kinect's location
- User location

Distance Model

Use distance between center of body and joints

Distances vector:

d1: 33
d2: 30
d3: 49
d4: 53
…

Displacement Model

use displacements between center of body and joints (as distance but using difference of vector).

Displacement vector:

v1: 0, 33, 0
v2: 15, 25, 0
v3: 35, 27, 0
v4: 43, 32, 0
…

Hierarchical Model

Skeletal body model as a tree where joints are nodes and the spine joint is the root. A feature represents the displacement between joint and its parent position.

Hierarchical vector:

h1: 0, 33, 0
h2: 15, -7, 0
h3: 20, 9, 0
h4: 18, 9, 0
…

Normalization

One dissimilarity source between the captured data from different individuals is related to their height.

The acquired skeletal data can be scaled properly, simply by dividing all limb lengths by a value that is proportional to a given user’s height.

Relative Normalization

Use the distance between spine and head joints to normalize all information

Unit Normalization

Scale all limb segments connecting two joints to unit length before computing the aforementioned features. This way, the vectors lose their length and keep their directions only.

The challenges

- Environment
- Input variability

WWW.PHDCOMICS.COM

Takeaways

A system, not just a detector

Invest equally in other components

A good design gesture can resolve a lot of problems

Collect real user data

Q&A

Tutto il nateriale di questa sessione su

http://www.communitydays.it/

#CDays13

@MatteoValoriani

Grazie a

Sponsor

icubed

intel Software

Microsoft

Technical Conferences

Windows

WPC by DotNet

aspitalia.com

Devleap BridgetheGap

DomusDotNet

DotDotNet

DotNet Liguria

dotnet LOMBARDIA

DotNetSide

sharepointcommunity.it

UGIdotNET

Visual Basic tips&tricks
Dal 1996 la community su VB e VB.NET

WindowServer.it

COMMUNITY DAYS

So Long and Thanks for all the Fish

COMMUNITY DAYS

Resources and tools

- http://channel9.msdn.com/Search?term=kinect&type=All (Others projects)
- http://kinecthacks.net/ (Others projects)
- http://www.modmykinect.com (Others projects)
- http://kinectforwindows.org/resources/ (Microsoft SDK)
- http://www.kinecteducation.com/blog/2011/11/13/9-excellent-programming-resources-for-kinect/ (resources)
- http://kinectdtw.codeplex.com/ (gesture recognition library)
- http://kinectrecognizer.codeplex.com/ (gesture recognition library)
- http://projects.ict.usc.edu/mxr/faast/ (gesture recognition library)
- http://leenissen.dk/fann/wp/ (gesture recognition library)