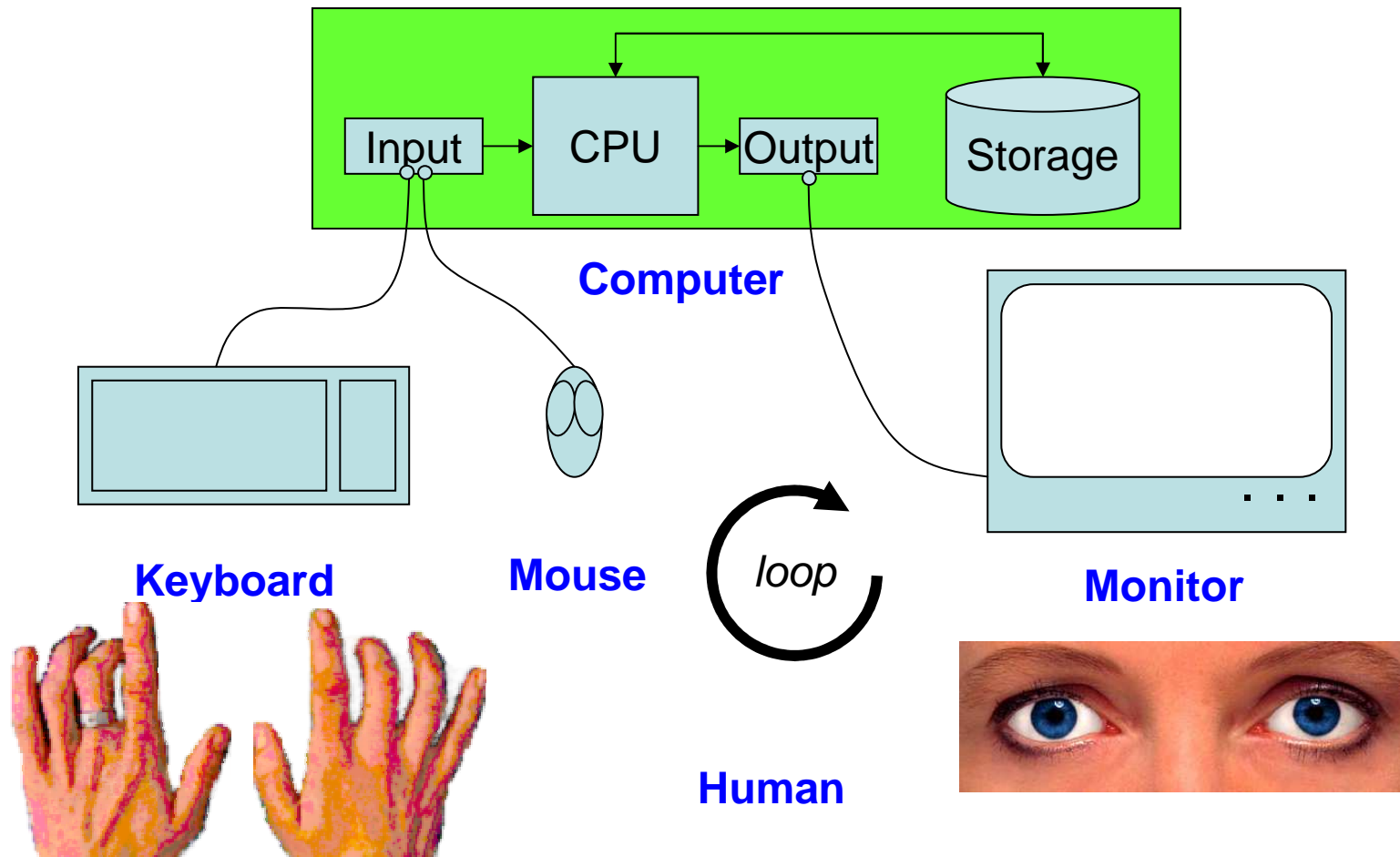




*Распознавание образов
для офисных
приложений*



3 Computing Loop Components between Humans & Computer





Dimensions of Computing Loop Components

Portable Mini-Computer (smaller than Netbooks)

Input	USB-Link
CPU	Single Board Computing
Output	... with embedded graphics
Storage	flash drive (no disc any more)



small
small
small
small

Human Output Devices (Feedback by Graphics)



Monitor	LCD (14")
Monitor-Glasses	Virtual Screen
Mini-Beamer	MPro 110 (Disagu, 14")



large
small
small

Human Input Devices (Control by Gesture)

Keyboard	Size of hands & fingers
Mouse	Size of hands & fingers

large
large

ViControl (new) Entrance to a Tool free human control

small

... everything is already small, but not the input devices



3 Current “Toolfree” Human Control Devices

Touch Screen (limited information: x, y, vx, vy)

Monitor (needs large screen)

Mobile (needs pencil)

large, no Tool

small, but Tool

Speech Detection

Microphone (limited numbers of environments)

small, but Noise

Gesture Detection (wide range of information)

CyberGlove (Virtual Technologies, limb angles)

Flying-Mouse (Logitech, limited information: x, y, z, vx, vy, vz, ax, ay, az)

small, but Tool

small, but Tool

Video Measurement (no tool necessary, wide range of information)

small, no Tool



... finger size is not benchmark for input device anymore



Two Technologies for Toolfree Measurement of Gestures

Radar, Ultrasonic, Lidar

Fingers can not be distinguished easily

not feasible

Active Video (needs additional light source)

PMD (Photonic Mixer Device) (pel with 100 μ m)

low resolution

Fast-Shutter (Microsoft, IR \rightarrow low texture, light power \rightarrow portable?)

feasible

Passive Video (standard environment illumination)

Mono

no distance

Stereo (synchronized imagers)

feasible

... but now, how to extract gestures?



Measurement of Gestures

Gesture: Temporal State of Human Limbs

shake the head

beckon with hands

cross the fingers

shake the legs

move the body

... and at last keystroke of keyboard & movement of mouse

Measurement of Depth and Motion

Depth only (Fast-Shutter-Video or Stereo-Video)

not enough

Motion only (Mono-Video Dense Optical Flow Measurement)

not enough

Combined Measurement of Depth & Motion

6D-Stereo (Depth & Motion Measurement with ViTechnology)

already proved

3 **Vi** Our Video Experiences

Dr.-Ing. Jens Schick

1986-1992

Dr.-Thesis

“Simultaneous Estimation of Shape and Motion from monocular Vision”

1992-1994

EU Commission Prometheus-Project for Daimler

“First Autonomous Driving in public Traffic based on image processing”

1994-1999

Project Manager at Daimler Process Development

“Video based ...Painting Control, ...Gap Measurement, ...Industrial Robot Control”

1999-2000

Project Manager at Daimler Research

“Emergency Brake of Trucks”

2001-2003

Constitution and directing the Bosch Automotive Video Team

“Mercedes Benz Night Vision”

2003-2008

Directing platform project Bosch Video

“Algorithm Development, Architecture Definition, System Integration”

2009

Vivid-Video-Vision

“Video Solutions with Embedded Systems”



ViTechnology

A general solution (based on Gestalt Primitives, Prof. Wertheimer 1905)

Massive data reduction without loss of information (more than 1:100)

ViAlgorithms

ViSpot

ViShape

ViString

ViMatch

ViFlow

ViDepth

Analysis of spots and dark holes (light, holes, ...)

Extraction of defined shapes (signs, plates,...)

Connection of Edges (lanes, borders, contours, ...)

Matching of texture (recognition, classification,...)

Gathering of common flow (detection by motion)

Measurement of Distances (dense 3D clouds, separation, ...)

Human Limb Perception

ViCluster

Statistical Combination of ViAlgorithms (object candidates)

... which types of Gestures do we start first with?

Application Point of View

coming from existing input systems (common habits)
minor changes of products (market acceptance)
minor technological risks (lost investments)

Technological Point of View

basic tasks first

- *detection from motion (model free detection)*
- *measurement from spatial clouds (safe distance measurement)*

keep it simple

- *no texture analysis (second step of recognition)*
- *no type classification (deep model assumptions)*

... for the consumer market (overview)

Notebooks

Substitution of Mouse Pad Hardware (later Extension of Mouse Functionality)

Substitution of Keyboard Hardware (later Extension of Keyboard Functionality)

Spatial Mouse & Keyboard (changing location and orientation of virtual devices)

Gesture Understanding (perception of finger, hands, head, body gestures)

Mobile phones

Introduction of Mouse Functionality (later Extension of Mouse Functionality)

Introduction of Virtual Keyboard (later Extension of its Functionality)

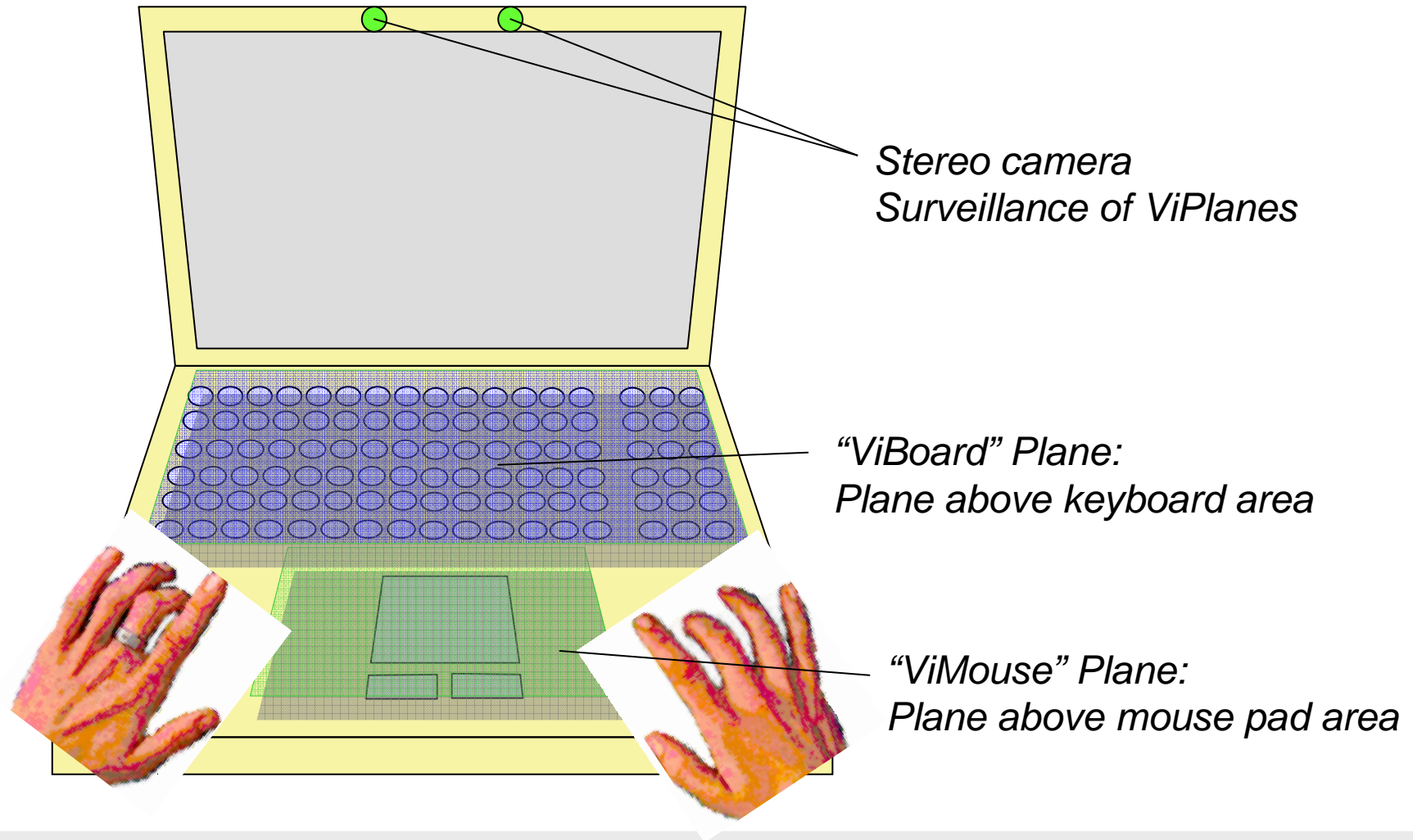
Consoles

Spatial Mouse & Keyboard (see notebooks)

Gesture Understanding (see notebooks)

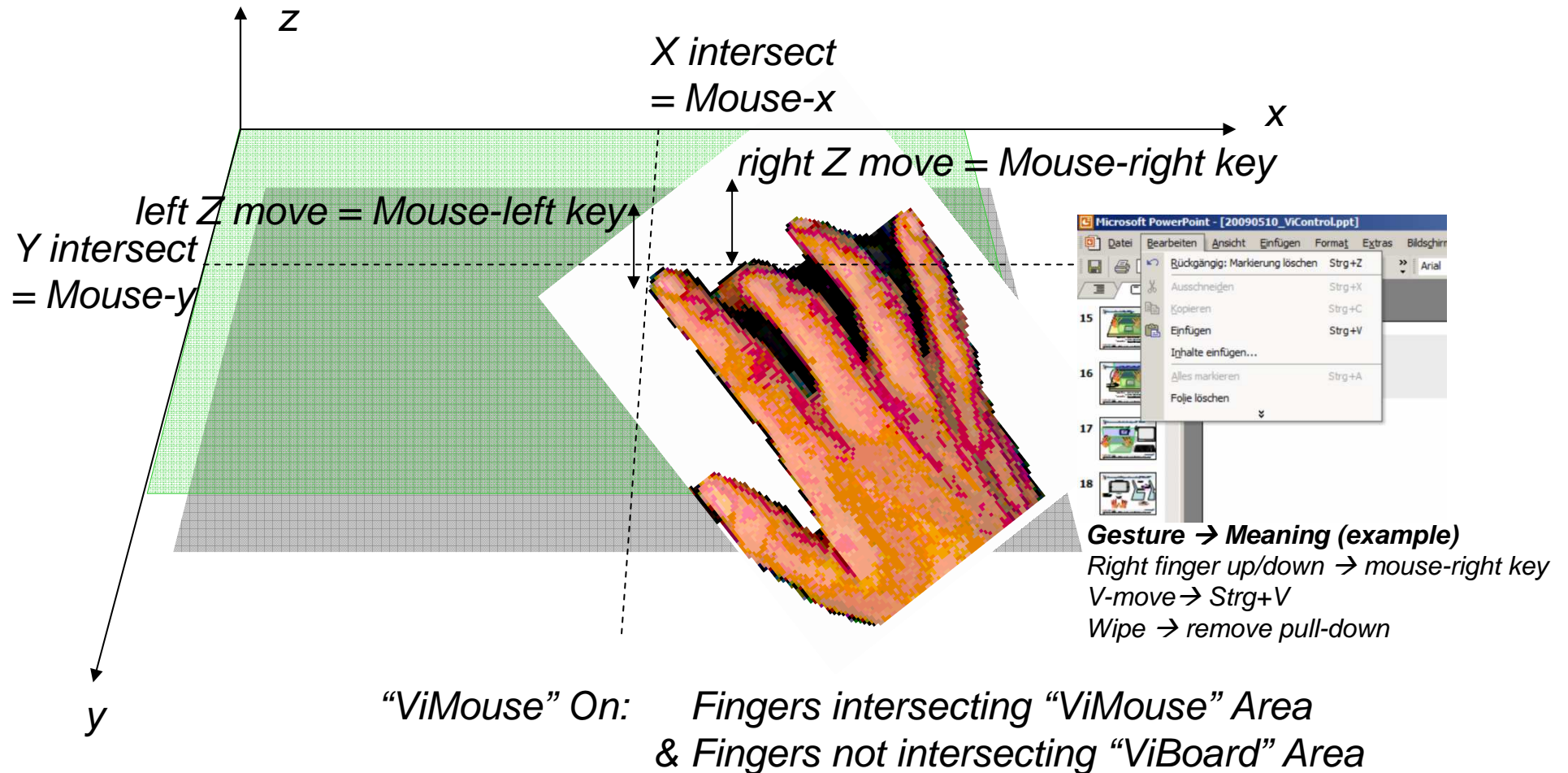


Definitions of Virtual Planes (ViPlanes)





Phase 1: First Version of the "ViMouse"





Phase 2: Improvements for the “ViMouse”

Other finger movements

*starting from up and down
with different speeds
and different directions*

Multiple finger detection

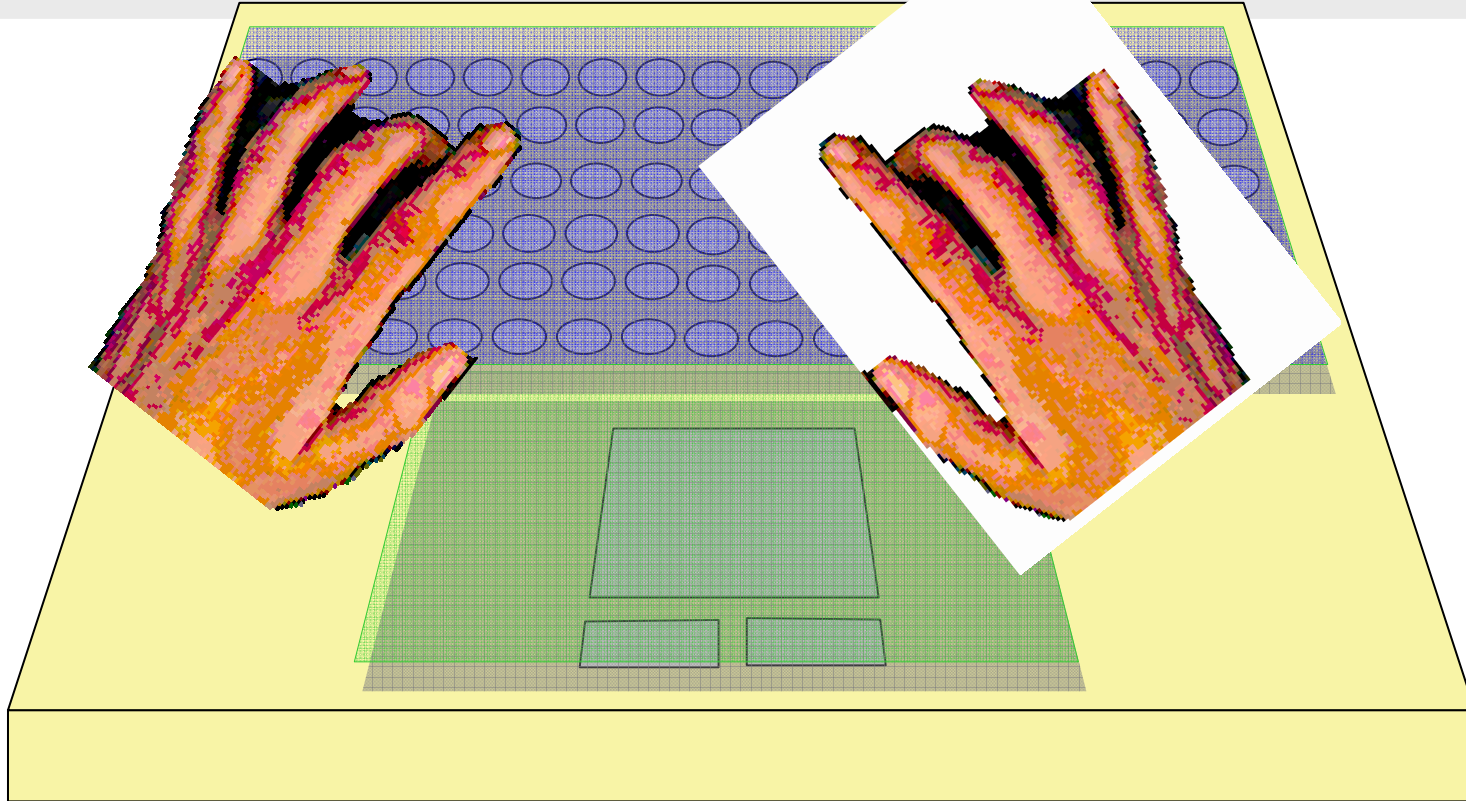
*separation of fingers (through motion)
relative motion detection (kind of coding)
... have your own ideas*

Double hand detection

*separation of hands (coming from the left and from the right)
starting with one finger each (quicker input)
relative positioning of fingers (kind of coding)
... were are close to the virtual keyboard*



Phase 3: Just an Extension of the “ViMouse” → “ViBoard”



“ViBoard” On: *Fingers intersecting “ViBoard” Area*

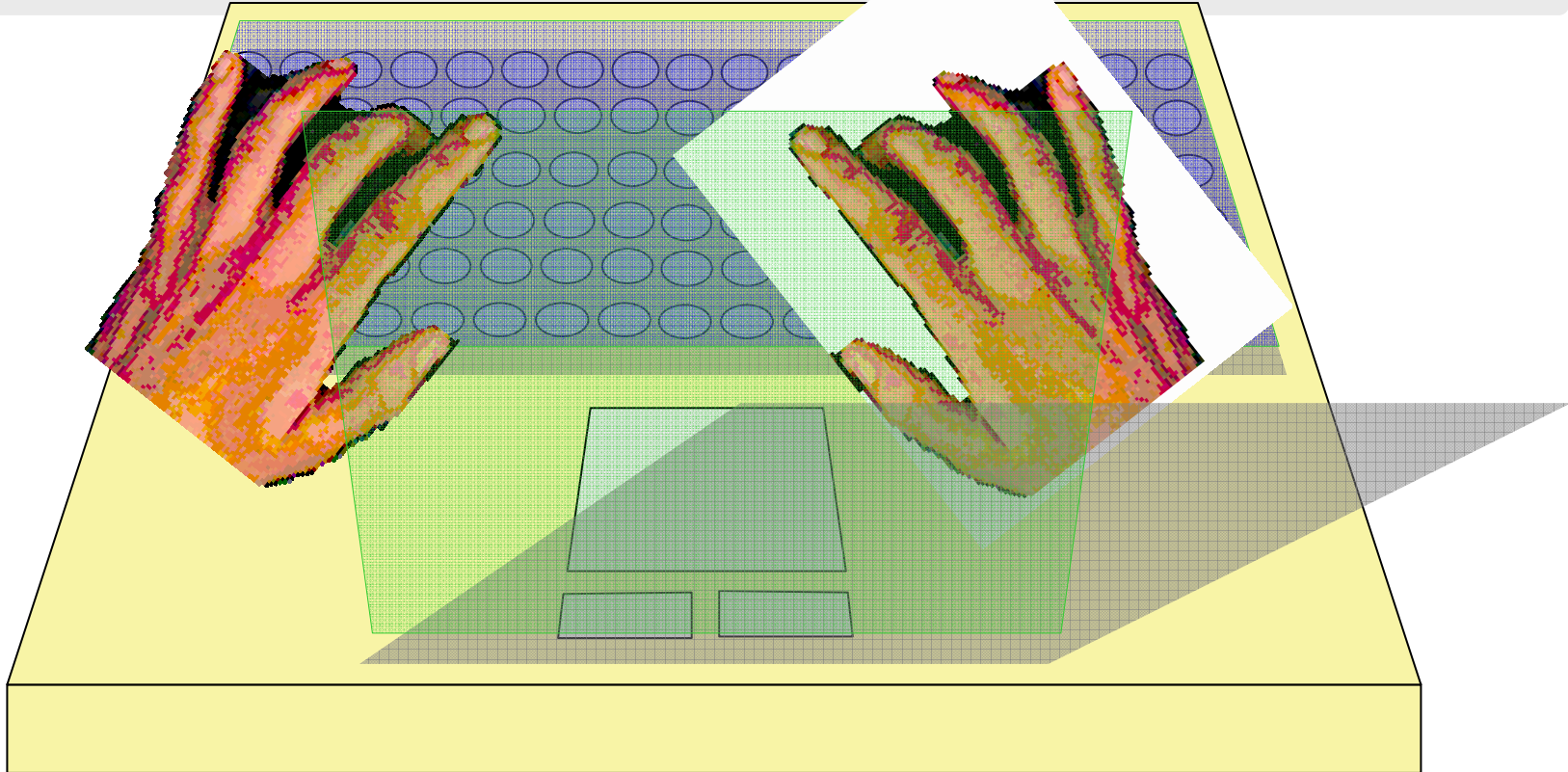
(gestures trained before with “ViMouse”)

“ViMouse” On: *Simultaneous usage possible*

(shifted “ViMouse” area, fingertip detection, motion understanding, ..)



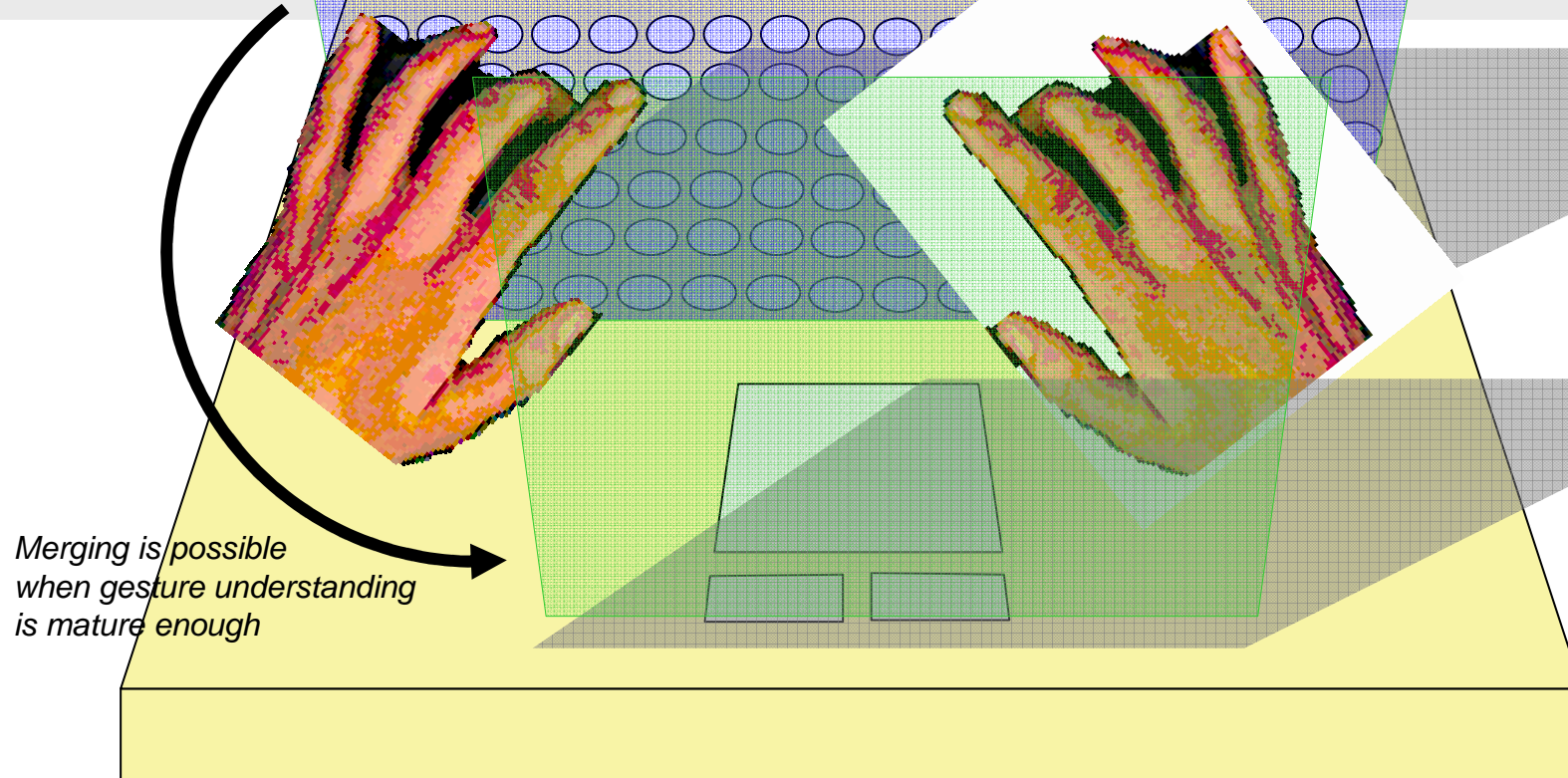
Phase 4: Lifting the “ViMouse” Area



*“ViBoard” On: Fingers intersecting “ViBoard” Area
& usually not intersecting lifted “ViMouse” area*
“ViMouse” On: Simultaneous usage possible
(hands have to be lifted to intersect “ViMouse” area)



Phase 5: Merge of "ViMouse" and "ViBoard"

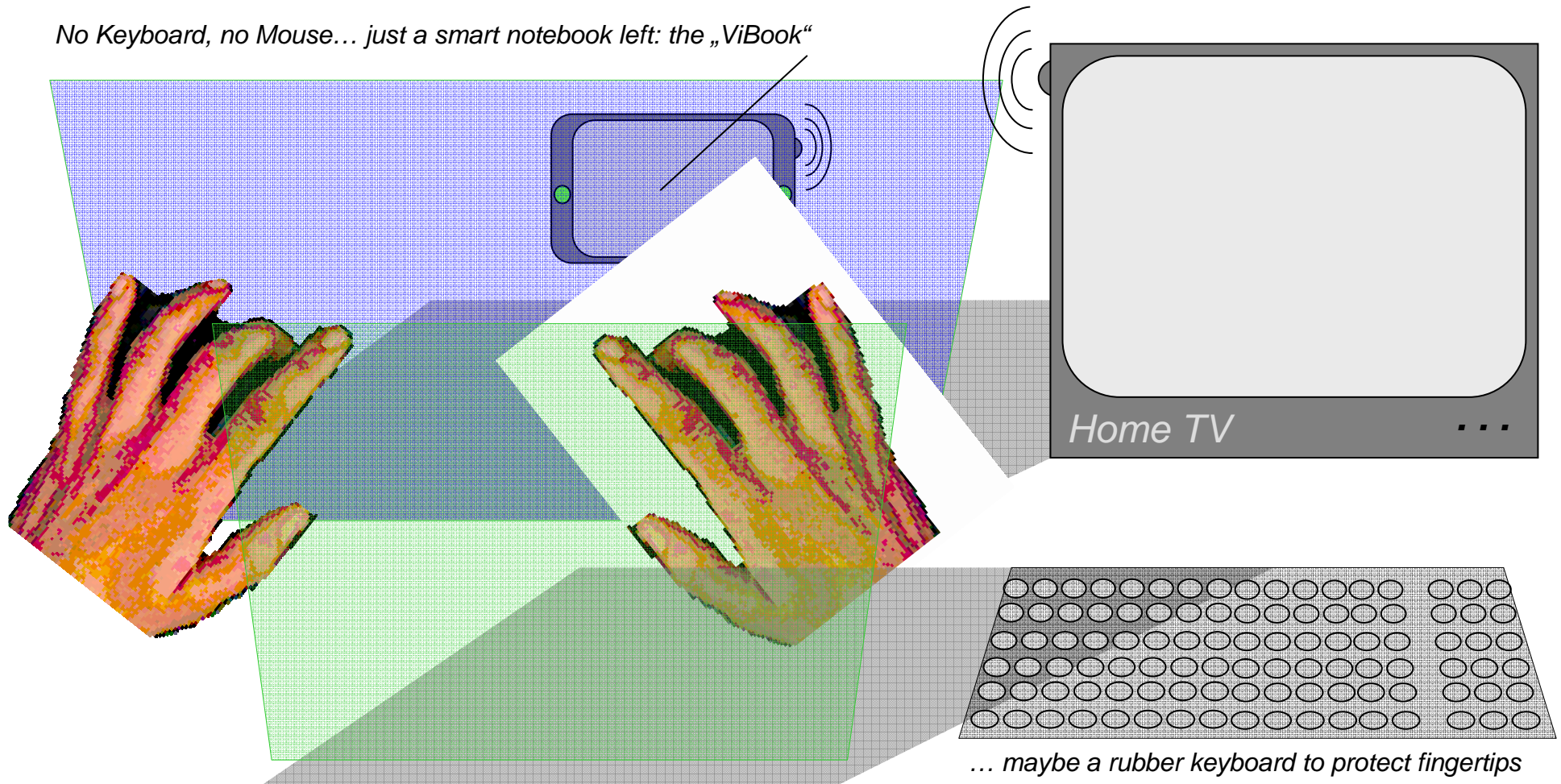


*"ViBoard" On: Fingers intersecting "ViBoard" Area
& usually not intersecting lifted "ViMouse" area
"ViMouse" On: Simultaneous usage possible
(gesture understanding leads to mouse access)*



Phase 6: Loss of Keyboard and Mouse → “ViBook”

No Keyboard, no Mouse... just a smart notebook left: the „ViBook“





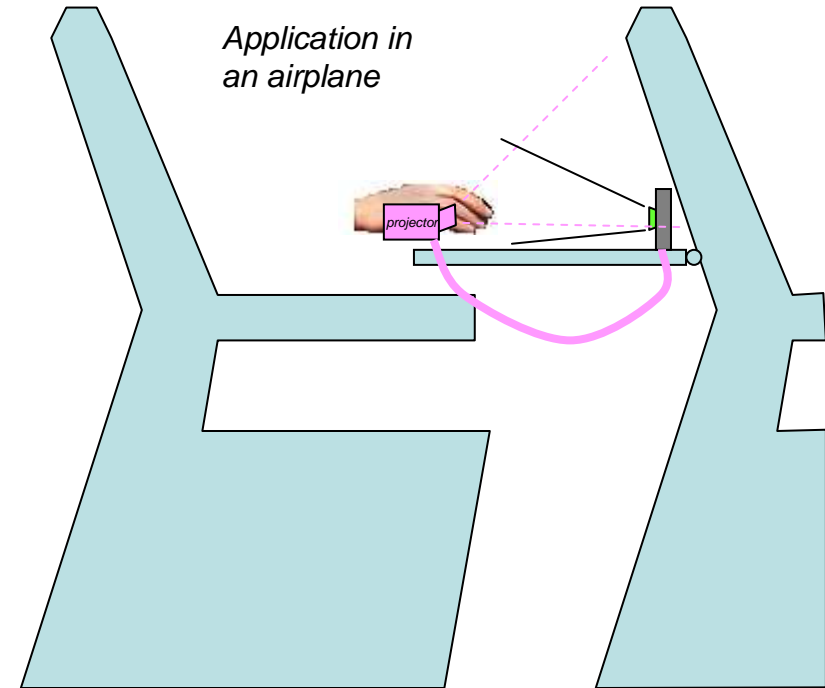
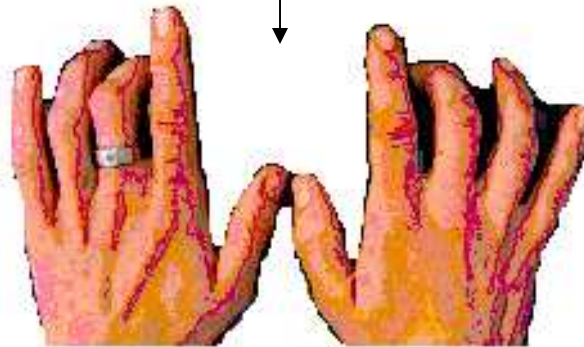
Phase 7: Maybe with Microphone and GSM → “ViPhone”

Original sized Mobile Phone



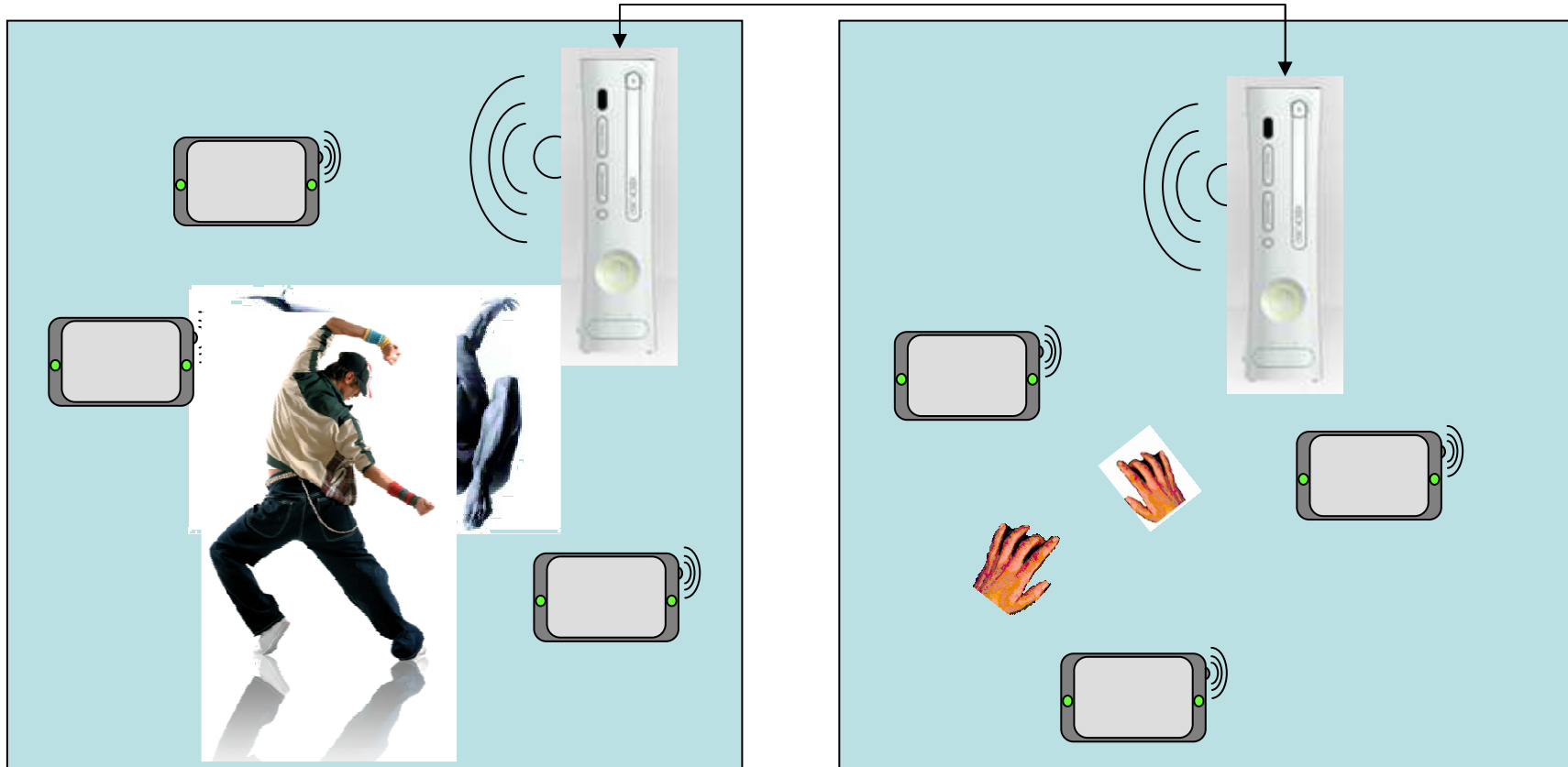
With direction sensitive microphone for videoconferencing (own development)

ca. 30 cm





Phase 8: Body Gesture Perception for Gaming → “ViControl”



Multi-ViControl applications for

- *Multi-User with Mono-ViControl each*
- *Single-User with multi-ViControl for better spatial resolution*
- *Multi-User with Multi-ViControl*



First Integration of ViTechnology with “DeepView”

DeepView

Low Power Embedded Video System

Embedded Pre-Processing on Hardware

Post-Processing on Microcontroller

Features

Image Control by Statistical Analysis (for critical illumination conditions, stereo auto calibration)

ViTechnology Algorithms on FPGA (proved at Bosch and Daimler)

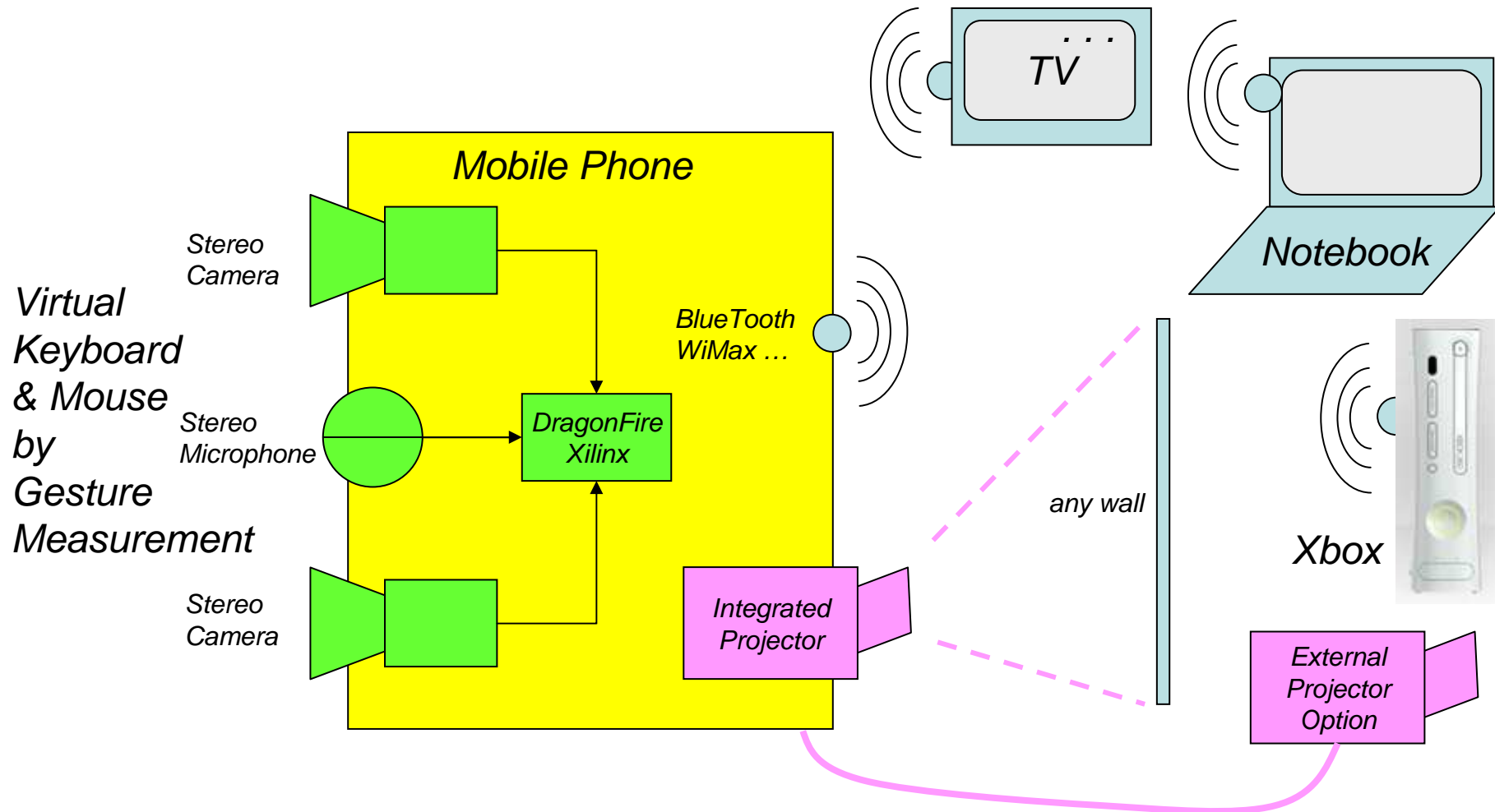
Scene Understanding by Cluster Interpretation on μC (Gaming, Office, Surveillance, ...)

Computational Power available (Low Power, affordable FPGA available since 2009)

Video Chip Development has been started (paused since 01/2009)



First Level Architecture of "DeepView"





Application Fields for “DeepView”

Business - Conferencing

- ... Speaker-Tracking*
- ... Zooming*
- ... Focusing*
- ... 3D-Data Exchange (lots of opportunities: Growing roundtables, augmented reality, ...)*

Business - MS-Office Applications

- ... 6D-Mouse to substitute the mouse and keyboard*
- ... embedded MS-Office software*
- ... projecting of video data (no Monitor needed)*
- ... wireless video contact (e.g.)*
- ... remote control of Notebooks*

Gaming

- ... Remote Control of GameBox (X-Box)*
- ... substitutes Wii*
- ... Pear to Pear Gaming by Mobile Network*

Services

- ... Private Surveillance*
- ... Car Safety*
- ... Socializing (electronic Baby sitter, pay attention to seniors ..)*
- ... Medical supervision*
- ... selling usage of games (even distributor of games...)*



Components of “DeepView”

A pair of Stereo-Cameras (4 MPel, Color, 5-10cm baseline, synchronized, auto-calibration)



Stereo-Microphone (Areus-Technology)



Xilinx DragonFire (FPGA, ARM11, 2 Watt, small sized,...)



ViTechnology (newest Realtime Vision Software)



Wireless Video Transfer Technologies (WiMax,...)



Compact Video Projector (small sized, optional, ...)



Embedded OS (Symbian, Linux, WinCE, ...)



Office-Software



... all components available