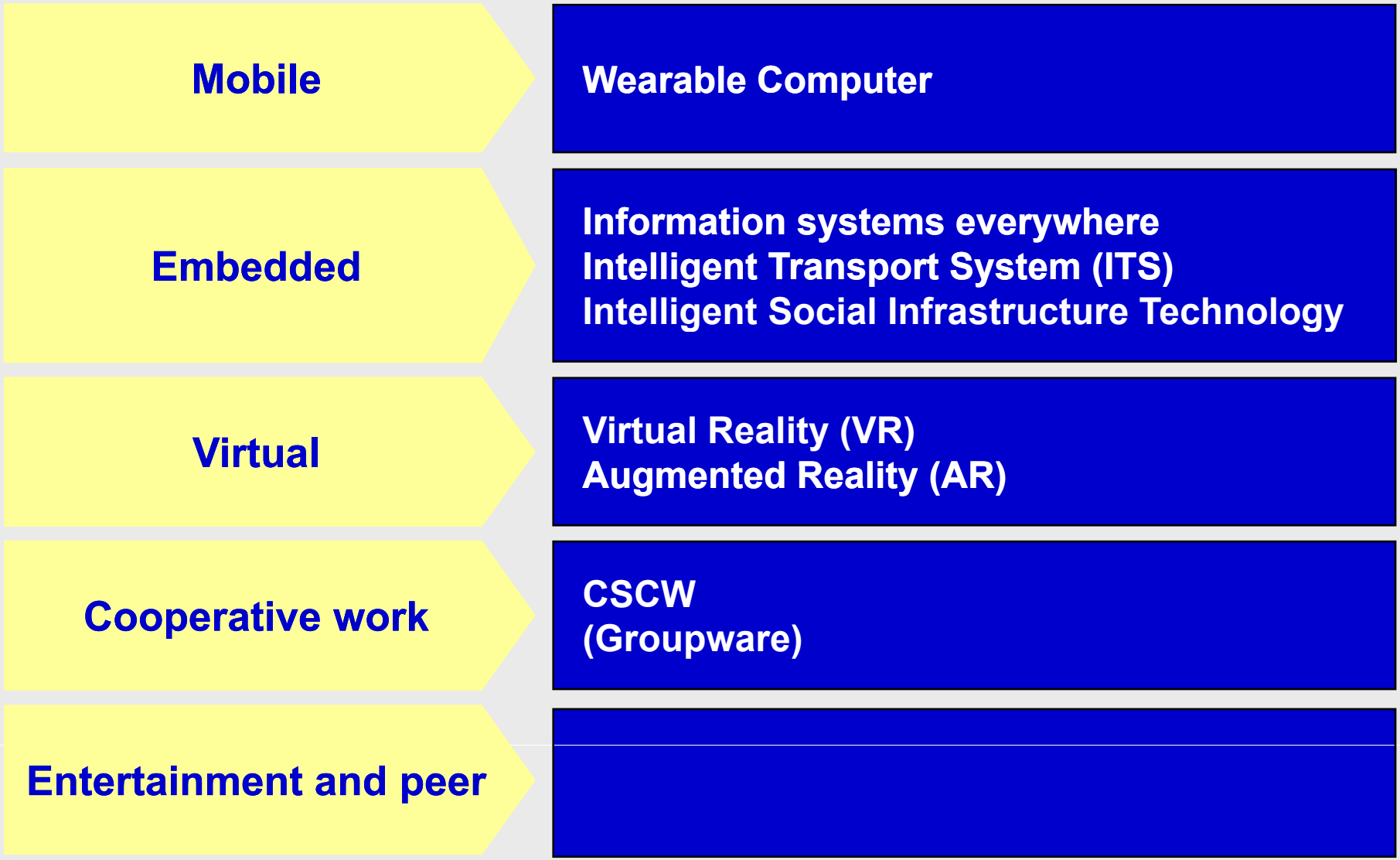
Characteristics of Multimedia Equipment



Recognition・Understanding・Synthesis techniques

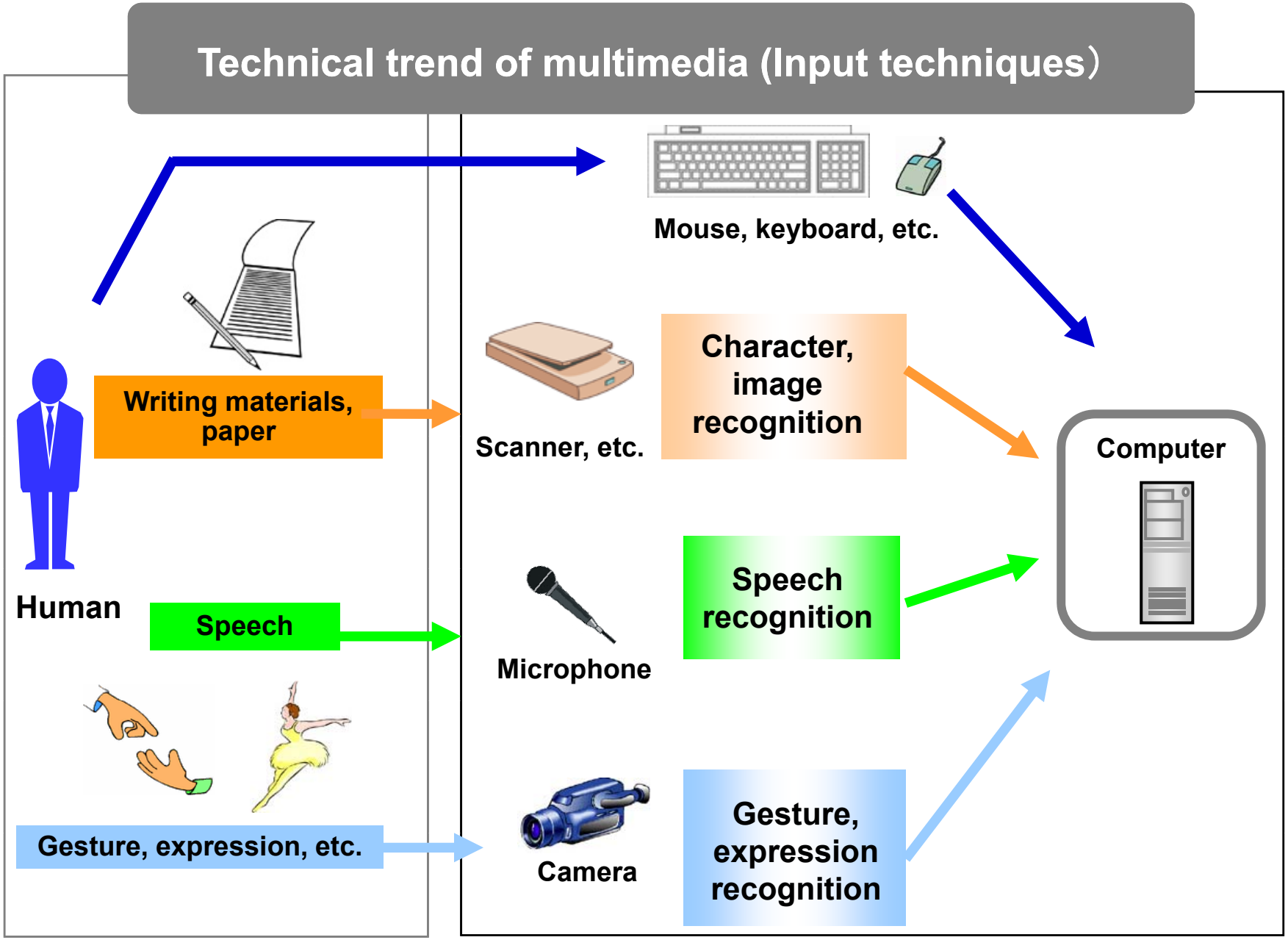
(1) Basic techniques

(a) **Recognition/understanding technique** which converts meaning, concept, and intention conveyed by media, such as speech, character, figure, and image, used by human-beings, to linguistic code that can be processed by computers.
Human image reader, which extracts/measures human gesture and eye movement and understands the meaning, is also included.

(b) **Media generation/synthesis technique** which converts code information, such as meaning, concept, and intention, that computers want to produce to human-beings, to various media such as character, figure, image, and speech.

(c) **Media fusion technique** which integrate multiple media to be recognized and understood.

Technical trend of multimedia (Input techniques)



Recognition・Understanding・Synthesis techniques

(2) Character recognition techniques

(a) Reading object：Numbers, alphabets,
Kana (katakana, hiragana), Chinese characters

(b) Degree of character variation：
Printing
Restricted handwriting（writing along guidelines）
Normal handwriting（following specimens, like JIS）
Freehand writing（with/without a frame）

Recognition method which can understand the meanings of characters and words (knowledge processing) is needed

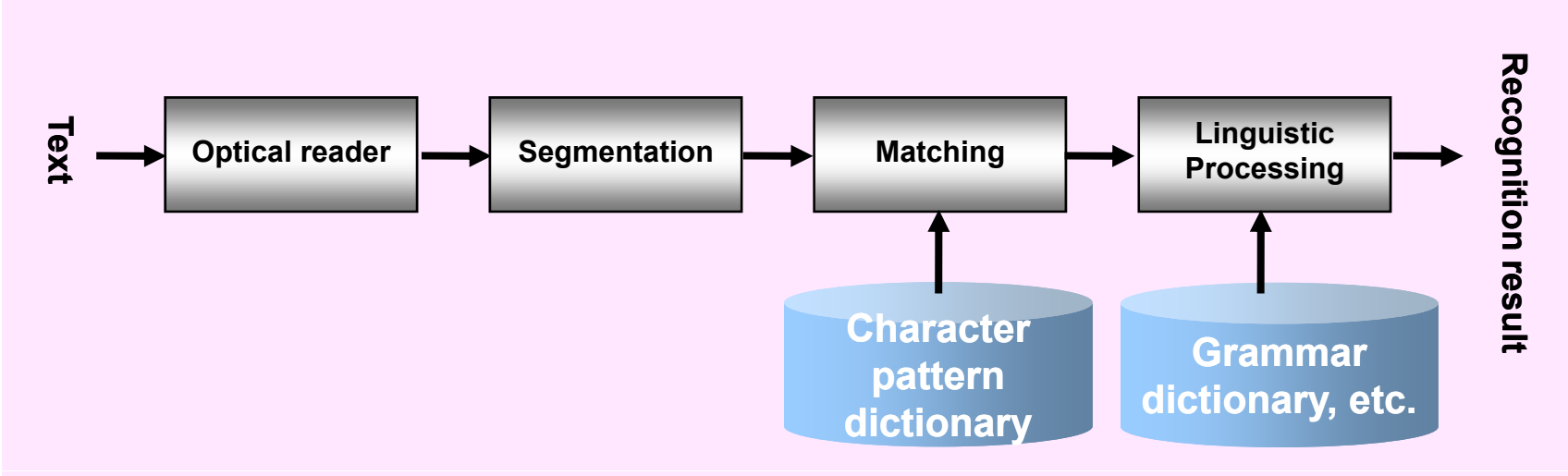
(3) Character generation techniques

Technical trend of Multi-media
(Character recognition techniques)

(a) Character recognition techniques

- Off-line recognition : optically reading the characters on the paper (postal code recognition, account list input, facsimile OCR, etc.)
- On-line recognition: detecting and recognizing handwritten strokes, including the orders by using an electronic pen (pen computer, etc)

(b) Flow of character recognition process



Recognition・Understanding・Synthesis techniques

(4) Speech recognition techniques

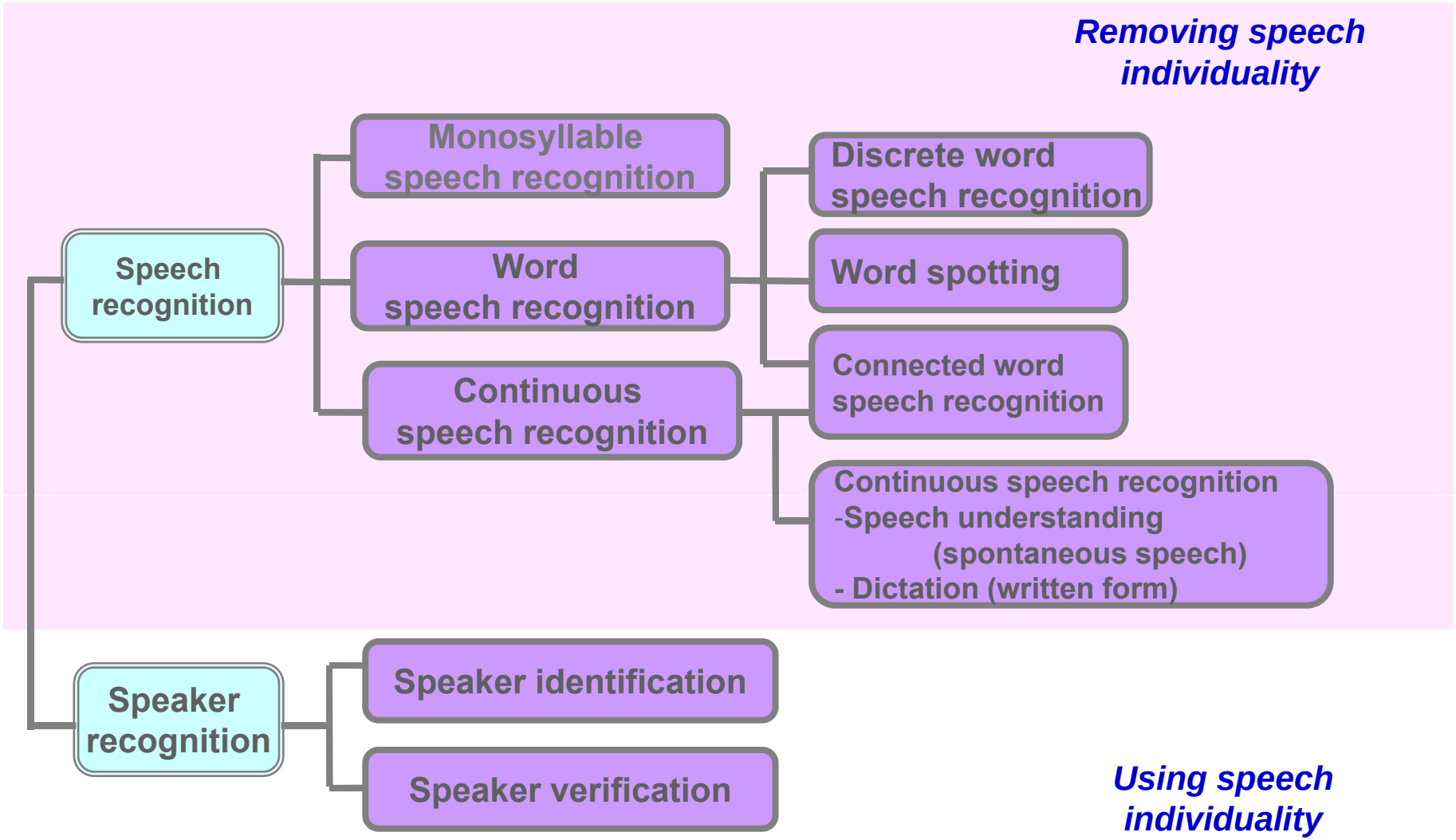
Object for recognition	Size of vocabulary for recognition	Speakers
(From isolated words to continuous sentences)	(From 10 words to several tens of thousand words)	(Speaker dependent, Speaker independent)

Speech is like cursive writing
（Time warping and co-articulation problems)

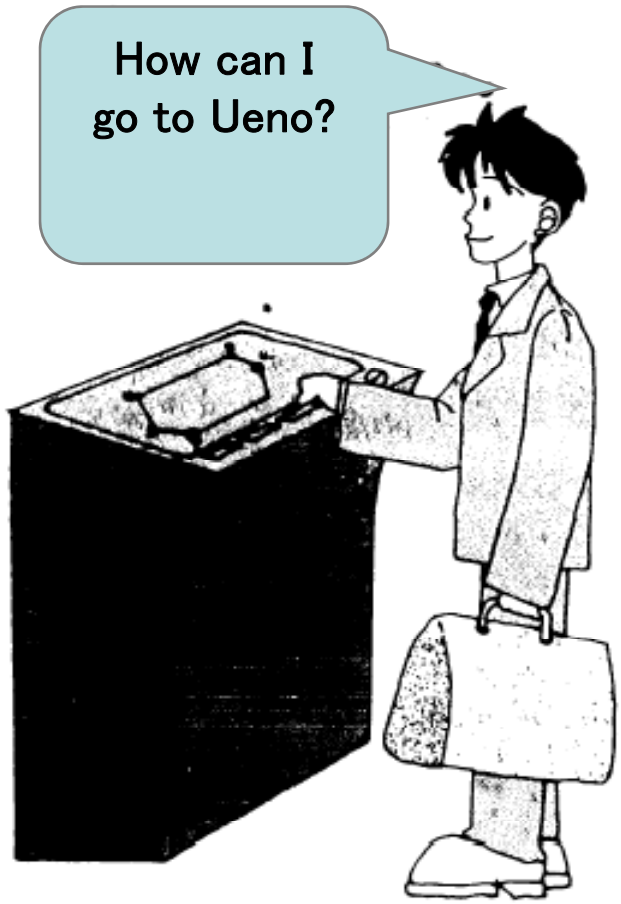
It is necessary to properly combine linguistic knowledge, such as grammar, meaning, context and task-dependent knowledge, with the phonetic information.

Speaker-independent speech recognition：
- **Express voice variations by statistical models**
- **Automatically adapt to the voice variation**

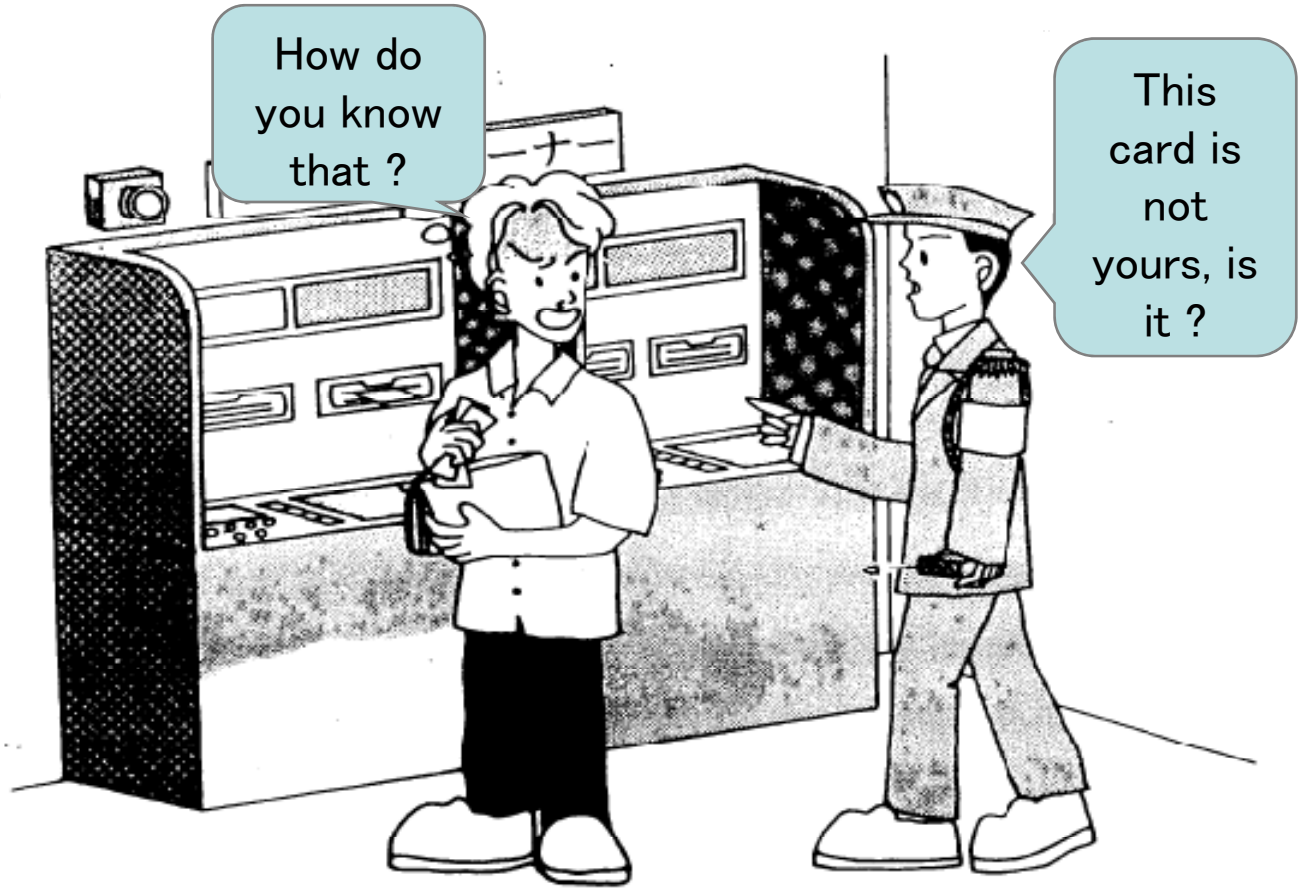
Technical trend of multimedia
(Classification of speech recognition techniques)



Information service at the street corner



Security by voice



Recognition・Understanding・Synthesis techniques

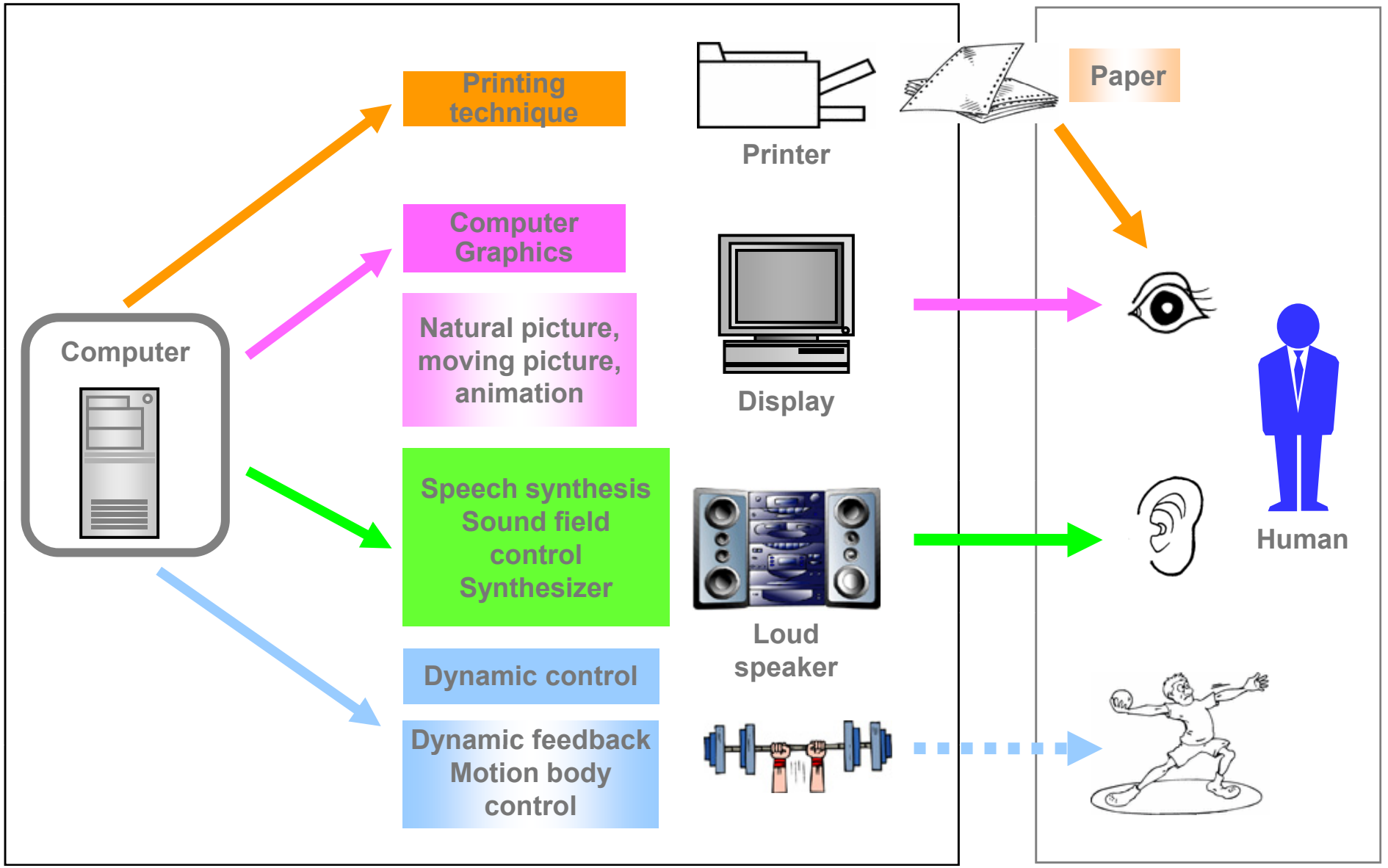
(5) Speech synthesis techniques

Storing waveforms	Storing parameters	Text-to-speech conversion
In case that the variation of vocabulary and sentences to be output is limited	＊ Store as a parameter set and synthesize	Arbitrary sentences can be converted into speech

In order to achieve high quality text-to-speech conversion, it is necessary to imitate human ability (language analysis and speech production mechanism) in reading written text.

In the future, to realize automated translation telephony, it is indispensable to be able to recognize and synthesize arbitrary speech in multiple languages.

Technical trend of multi-media (Output techniques)



Recognition・Understanding・Synthesis techniques

(6) Graphic and image recognition

Object

From binary images
to gray-scale images,
natural images,
3D scenes (view)

Processing stages

Feature extraction

（Thinning, description of carves and region shapes, binary image, edge detection, texture analysis, region segmentation, etc.）

↓

Structure analysis

↓

Content understanding

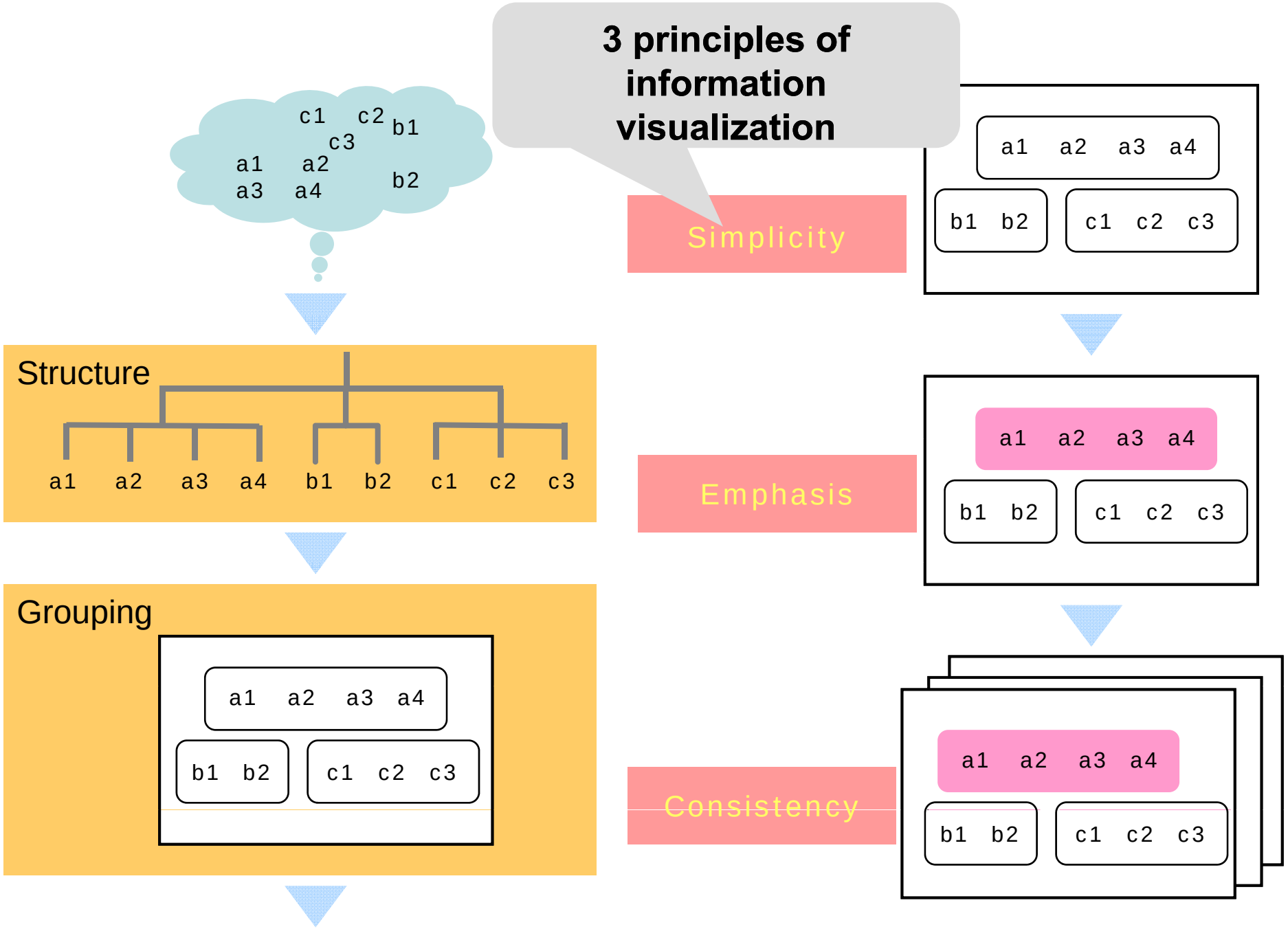
Future research: Describing real-world images based on reasoning for analyzed images and knowledge sources related to the object world

Recognition･Understanding･Synthesis techniques

(7) Graphic and image generation

It becomes difficult when the level of code information is in a concept level, instead of graphic commands (attribute information describing dots, straight lines, curves, colors, etc.).

Model-based CA (computer animation) techniques, hardware, software, artificial intelligence (knowledge processing), etc.


3 principles of information visualization
c1 c2 b1 c3 a1 a2 a3 a4 b2
Structure
a1 a2 a3 a4 b1 b2 c1 c2 c3
Grouping
a1 a2 a3 a4
b1 b2
c1 c2 c3
Simplicity
Emphasis
Consistency

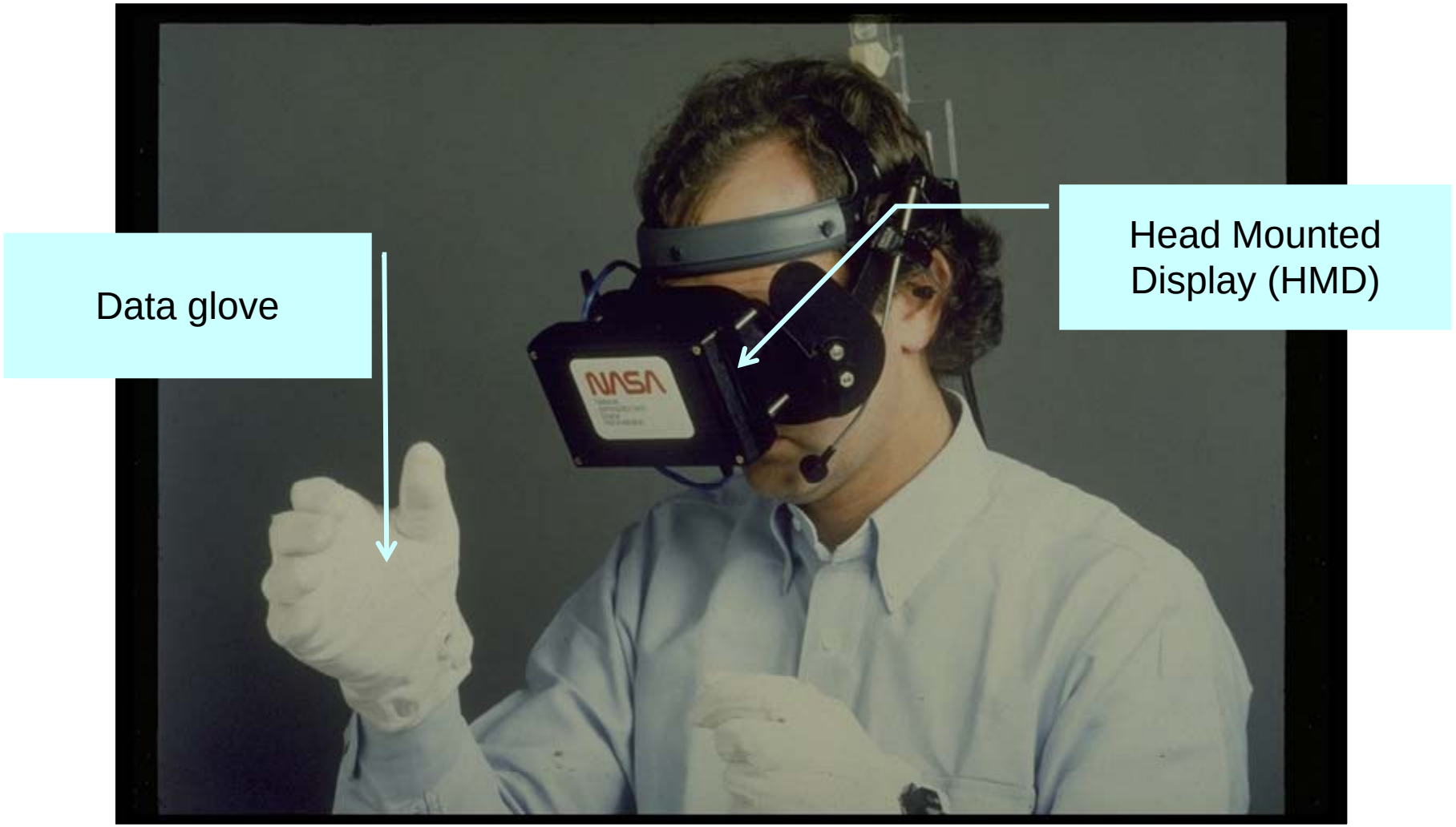
Principal factors concerning three dimensional space recognition

Factor		Property	Valid distance (m)	Artificial realization
Monocular	**Adjustment**	Adjustment of the focus of the lens in the eyeball	within 2～3m	Difficult
	Perspective transformation	Size variation of the image on the retina according to the distance	within 500m	Possible
	Motion parallax	Interaction of the movement of the images generated when the observer or the object moves	within 200m	Partially possible
	Viewing area	Restriction of the observable area	more than 50m	Difficult
Binocular	**Congestion**	Cross-eyed when observed closely	within 30m	Partially possible
	Parallax	Difference between the locations of the images on the retina of the eyes	within 100m	Possible

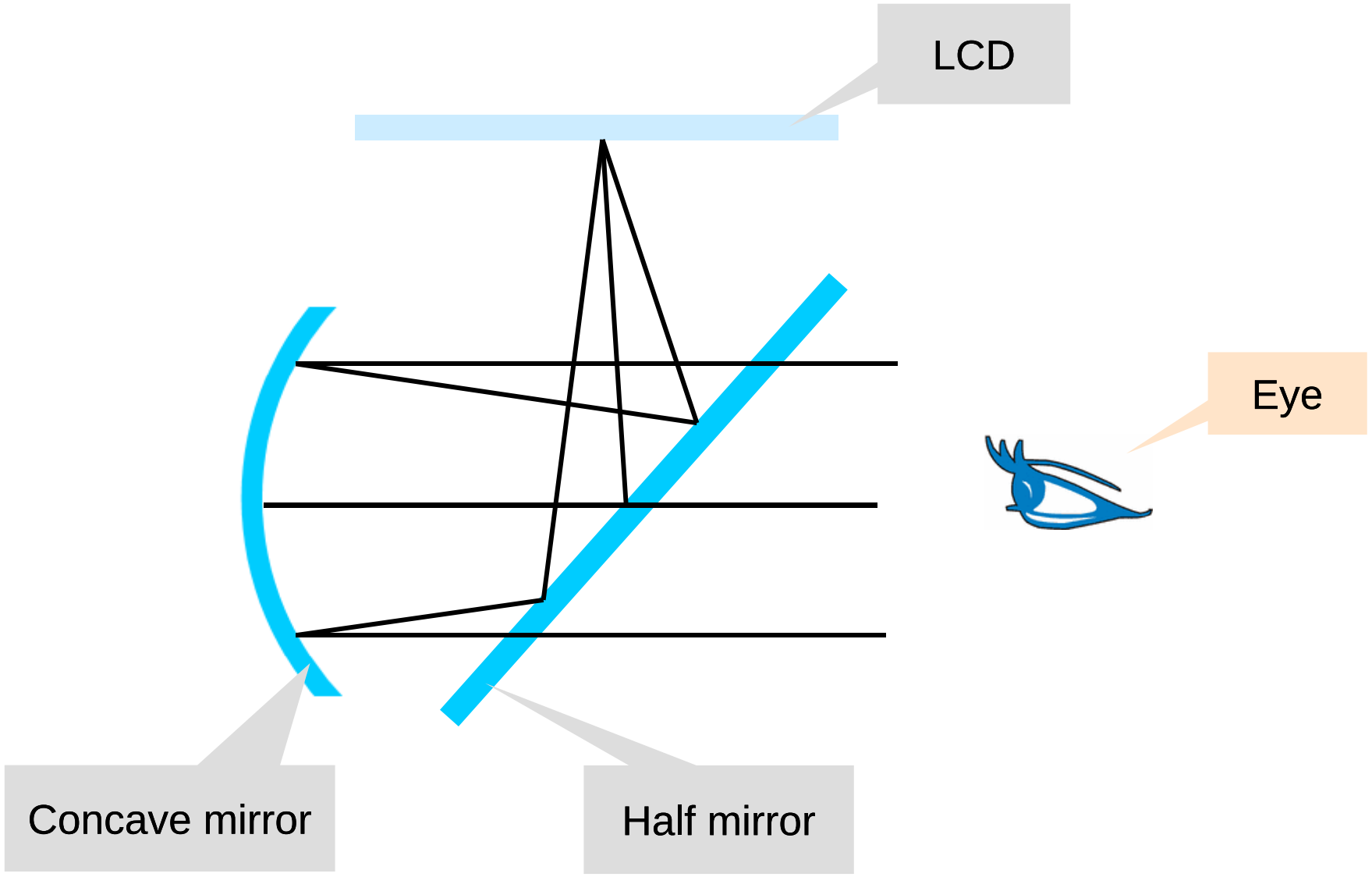
Characteristics of various three dimensional displays

Method	Glasses	Color	Video	Big screen	Viewed by many people	Width of visual area	Moving eye points	Applicable domains
HMD	**Needed**	○	○	○	×	○	○	VR
Anaglyph	**Needed**	×	○	○	○	○	×	
Polarizing glasses	**Needed**	○	○	○	○	○	×	Movie event
Time division shutters	**Needed**	○	○	○	△～○	○	×	Movie event
Parallax barrier	**Needless**	○	○	△	×	×	×	
Lenticular	**Needless**	○	○	○	×～△	○	○	Personal use (for terminal)
Back light division	**Needless**	○	○	△	△	○	○	Medical care
Depth samples	**Needless**	○	△	×	△	○	○	
Integral photography	**Needless**	○	×	△	△	○	○	Medical care
Holography	**Needless**	△～○	×～△	×～△	△	○	○	Art

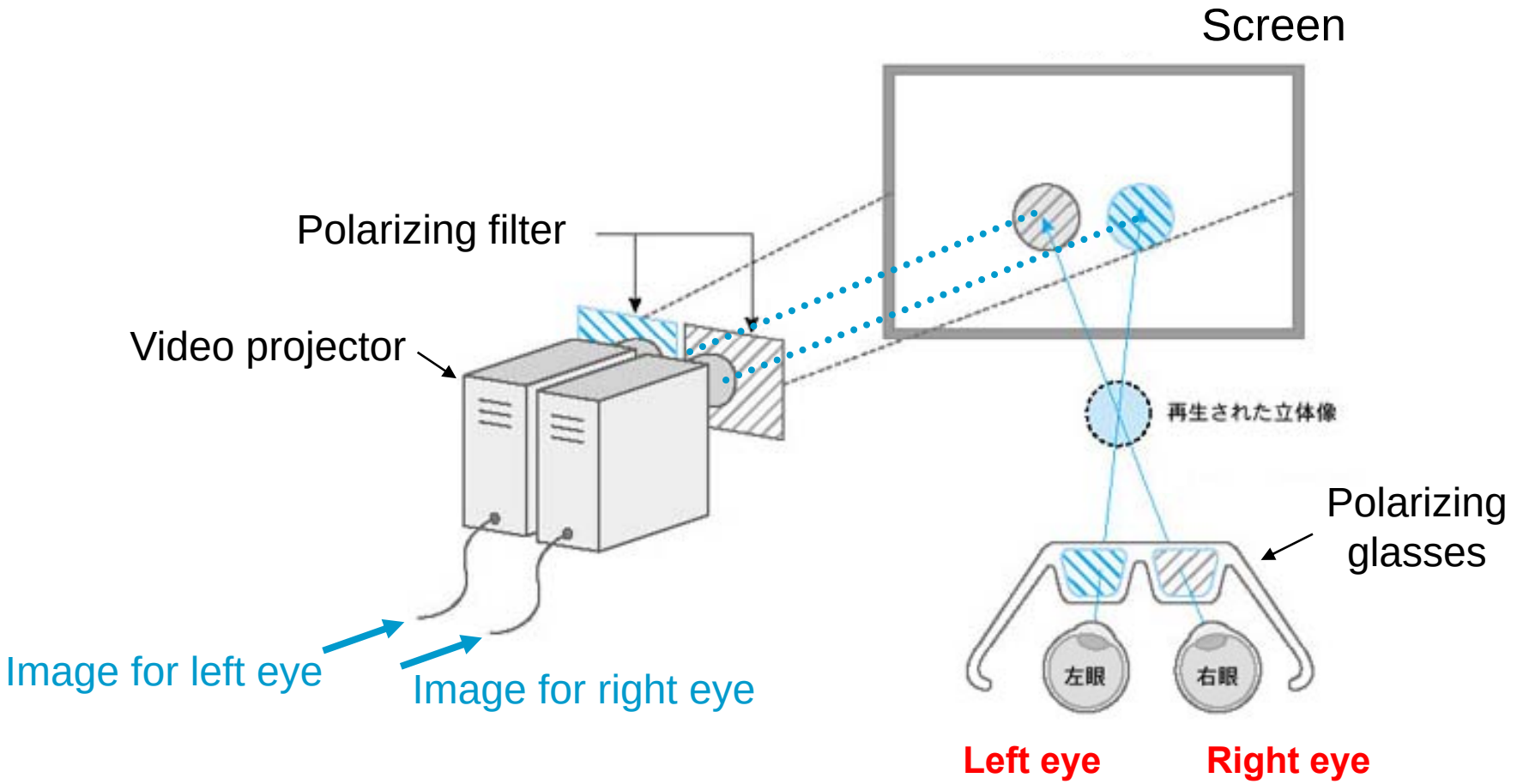
A system which can freely operate three dimensional image world (VPL Co. in 1989)



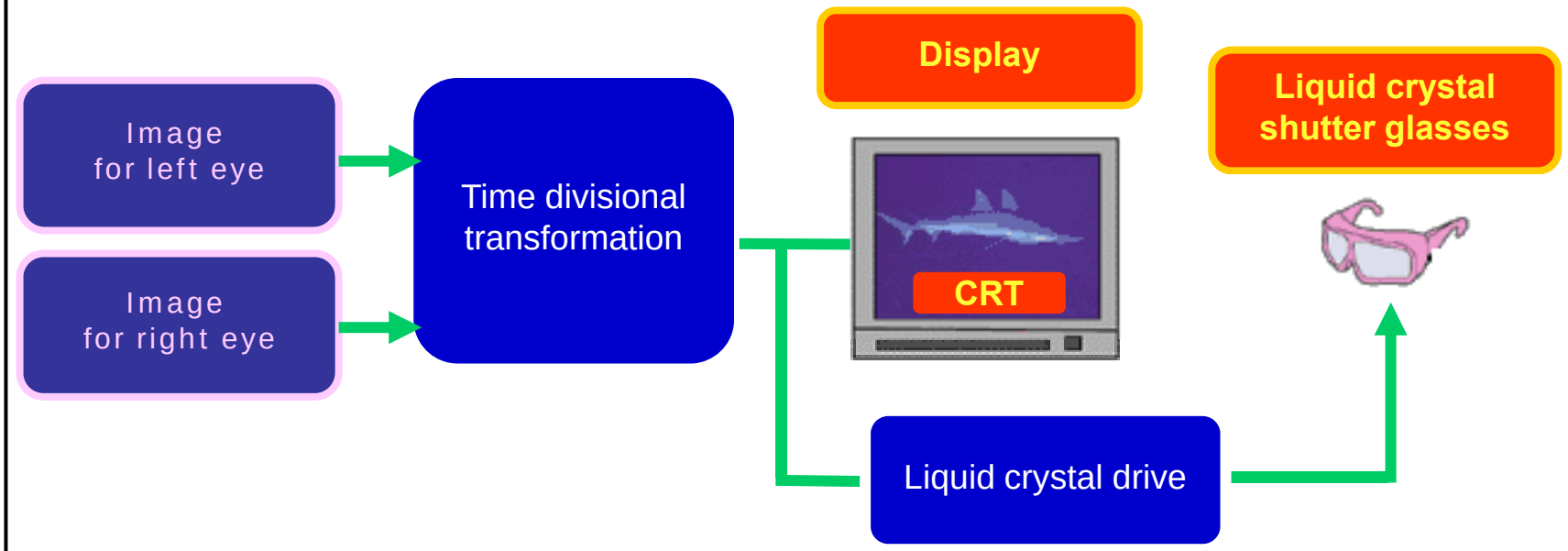
Principle of HMD



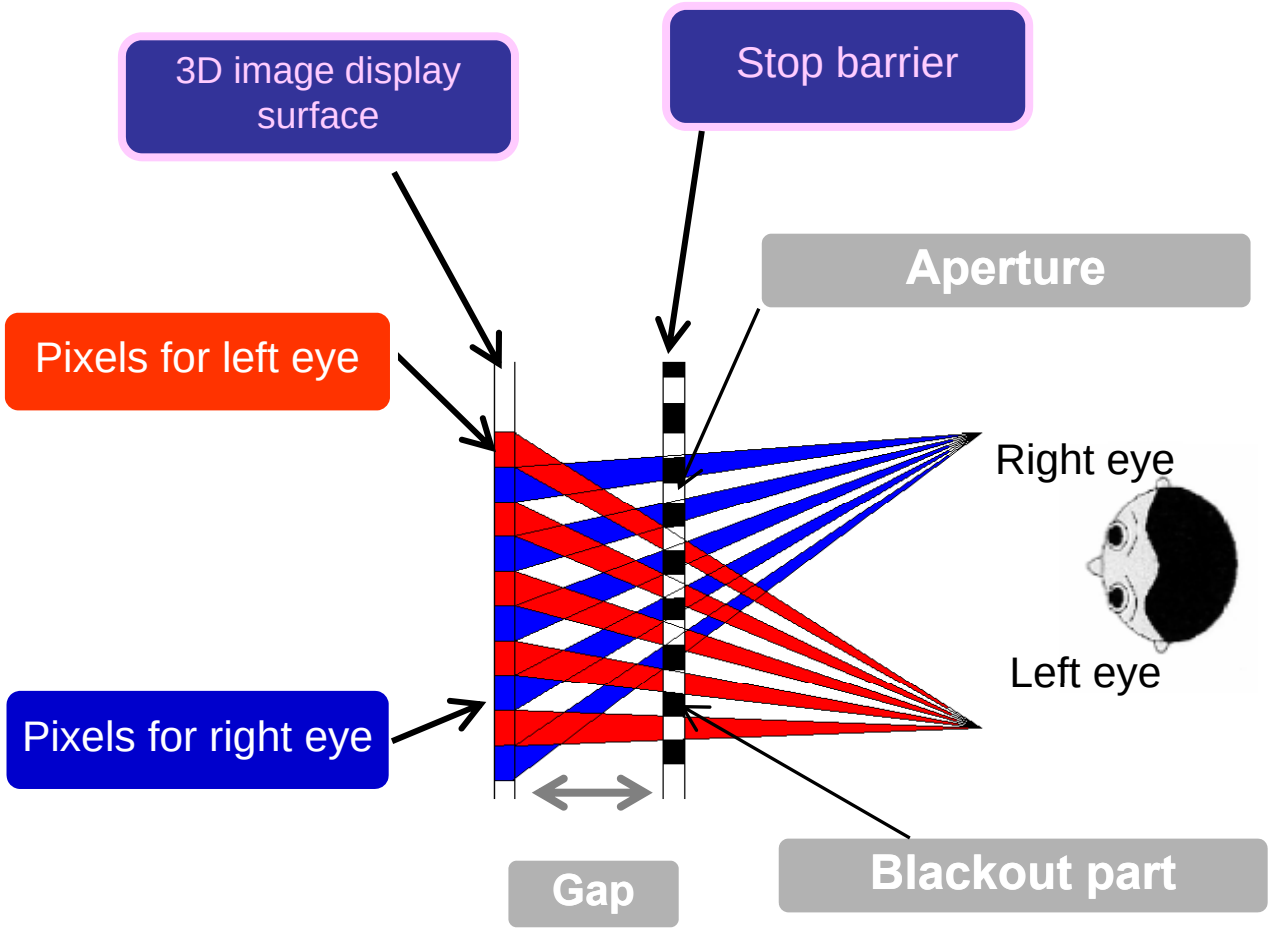
Principle of polarizing glass method



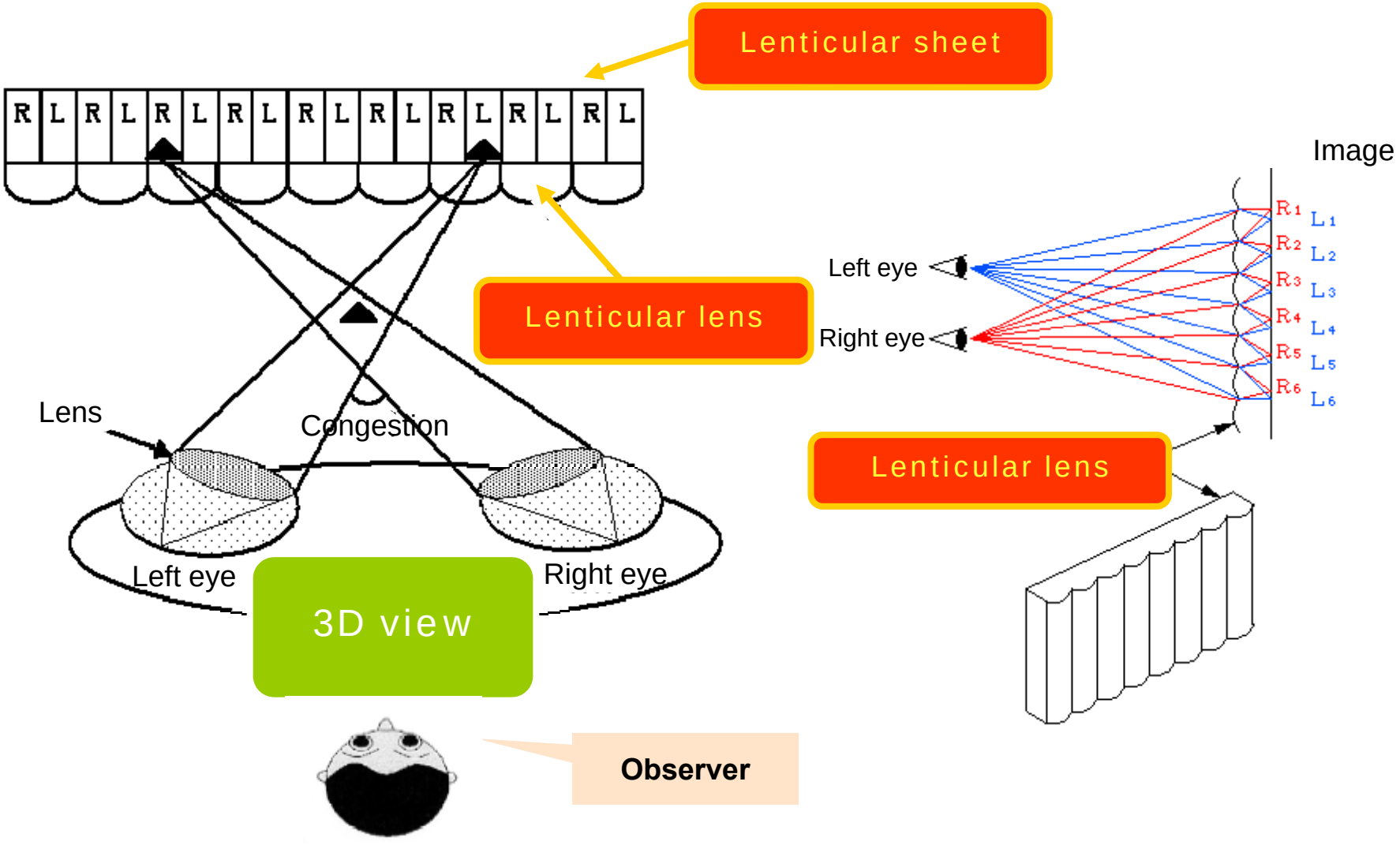
Principle of time divisional shutter method



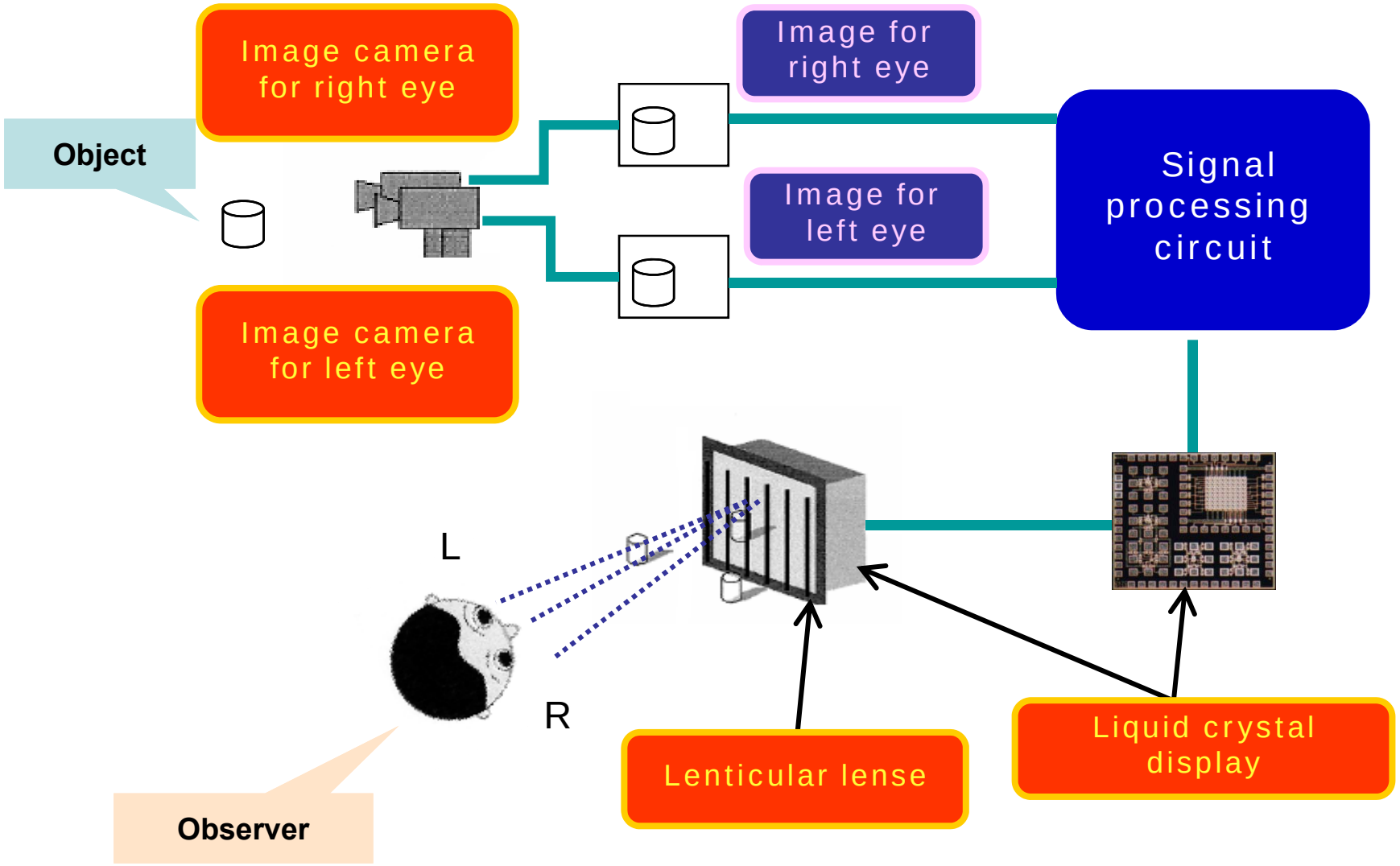
Principle of parallax stereogram



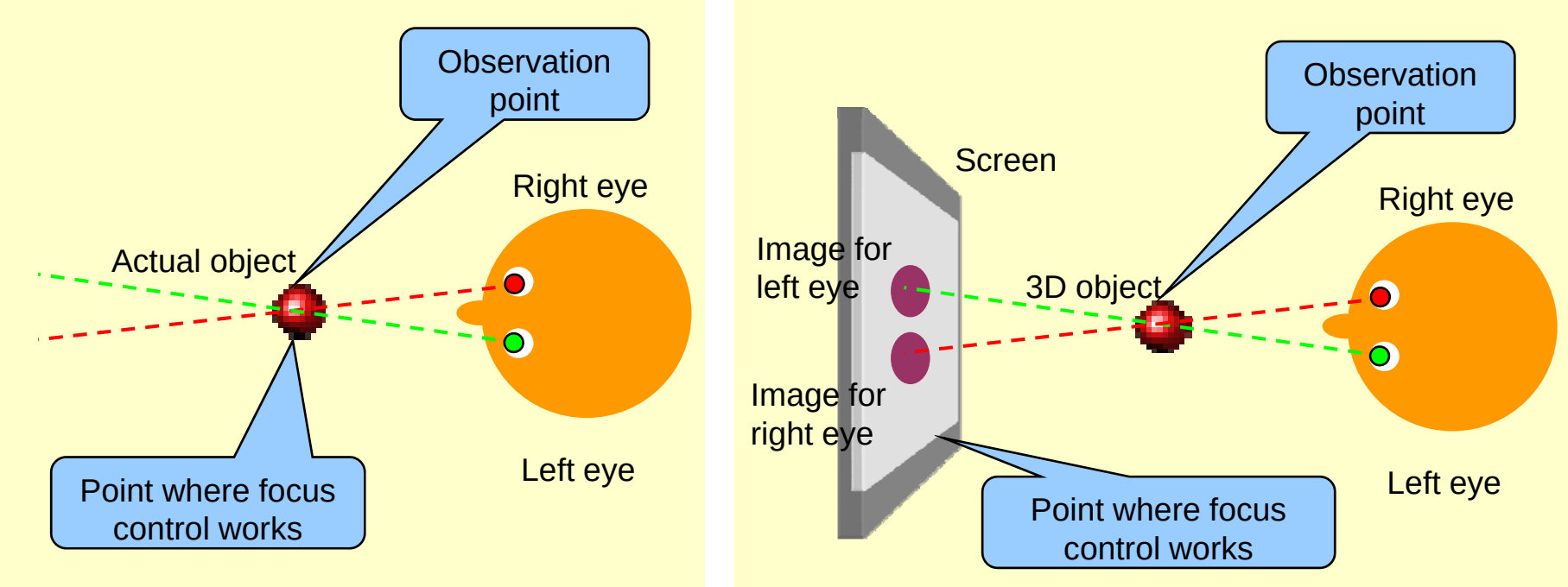
Principle of stereo vision using lenticular sheet



Principle of 3D presentation by lenticular method



Difference between observation of actual object and 3D image



(a) Actual object observation

(b) 3D image observation

Technical trend of multimedia
(Media conversion･editing techniques)

Media conversion techniques

Image (bit data) → character (language)

Character recognition

Character → speech

Text-to-speech conversion (Speech synthesis)

Speech (acoustical signal) → language

Speech recognition

Video (video signal) → position (pointer)

Gesture recognition

Editing techniques

Audio Editing:

Speech and audio signal editing, etc.

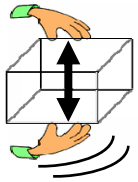
Image editing:

Automatic image indexing, etc.

Roles of two level ontology

Building a house

With a roof like temples in Thailand



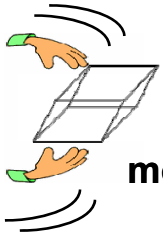
Height

Translation to shape level

Conceptual level ontology

Sharp-pointed sphere

Auxiliary expression by hand movement



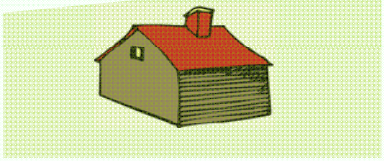
Shape modification

Translation to graphic level

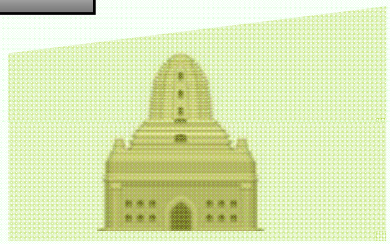
Formal level ontology

Parameter set

3D graphics presentation



Virtual Space



Multi-media Applications

Media capture	**Process to convert information in analog domain to digital expressions in a computer（Image scanner, MIDI code generation tool, character recognition, speech recognition, etc.）**
Media preparation	**Media transformation （Adobe PhotoShop, audio mixer, post-processing of video, syntax analysis, etc.）**
Media integration	**Integration of different media data （Audio video editing, presentation generation tool, WYSIWYG, etc.）**
Media interaction	**Interaction with users （Interactive television, home shopping, intellectual media agent, etc.）**
Media communication	**Communication between multiple users （Video phone, video meeting, CSCW, multimedia mailing, etc.）**

Ubiquitous computing

Mark Weiser, Xerox PARC
“yard, foot, inch”

The PalmPilot

used mainly to keep track of personal calendar and contact information

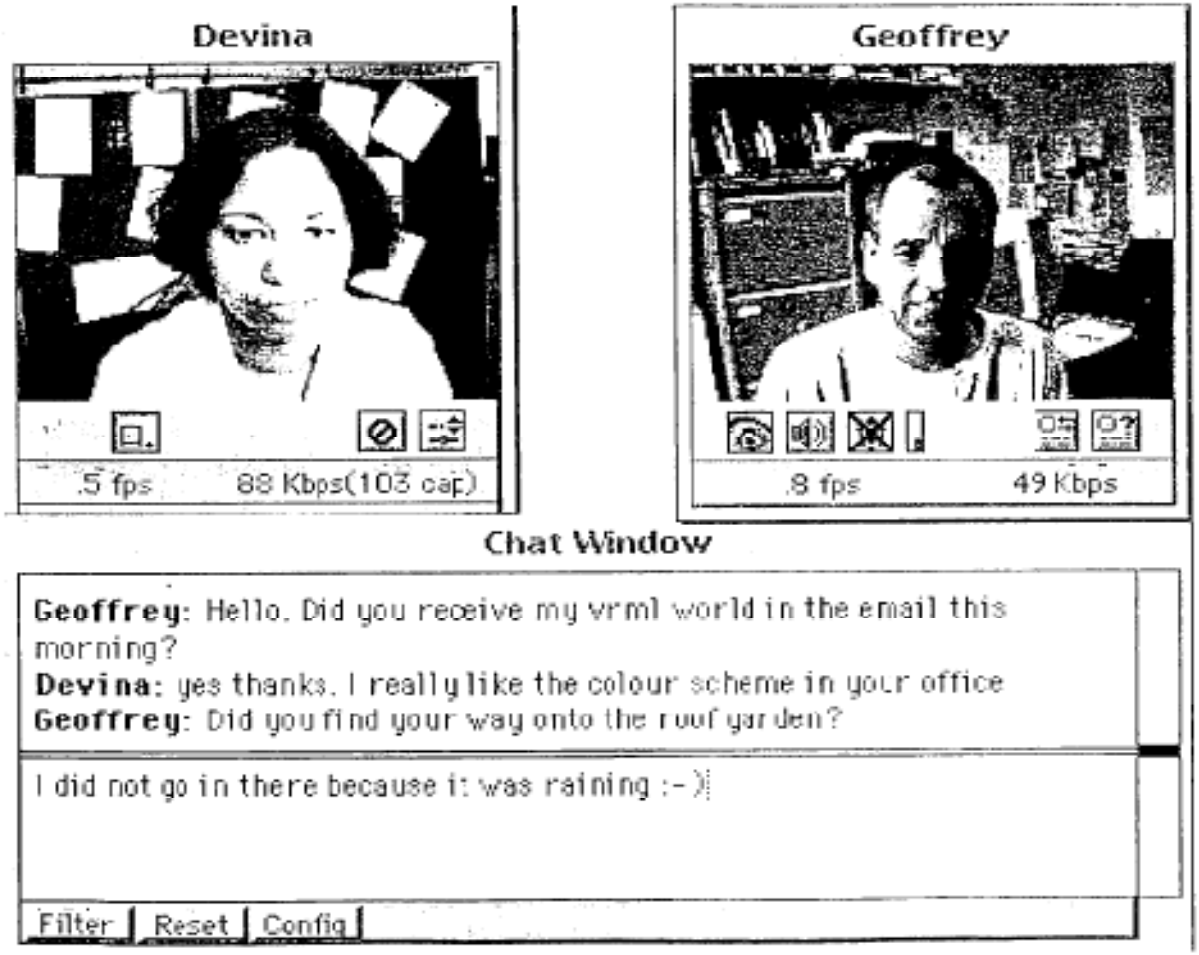
An example of a tablet computer – a Fujitsu Stylistic 1200

Teleconferencing (remote meeting) system

Camera

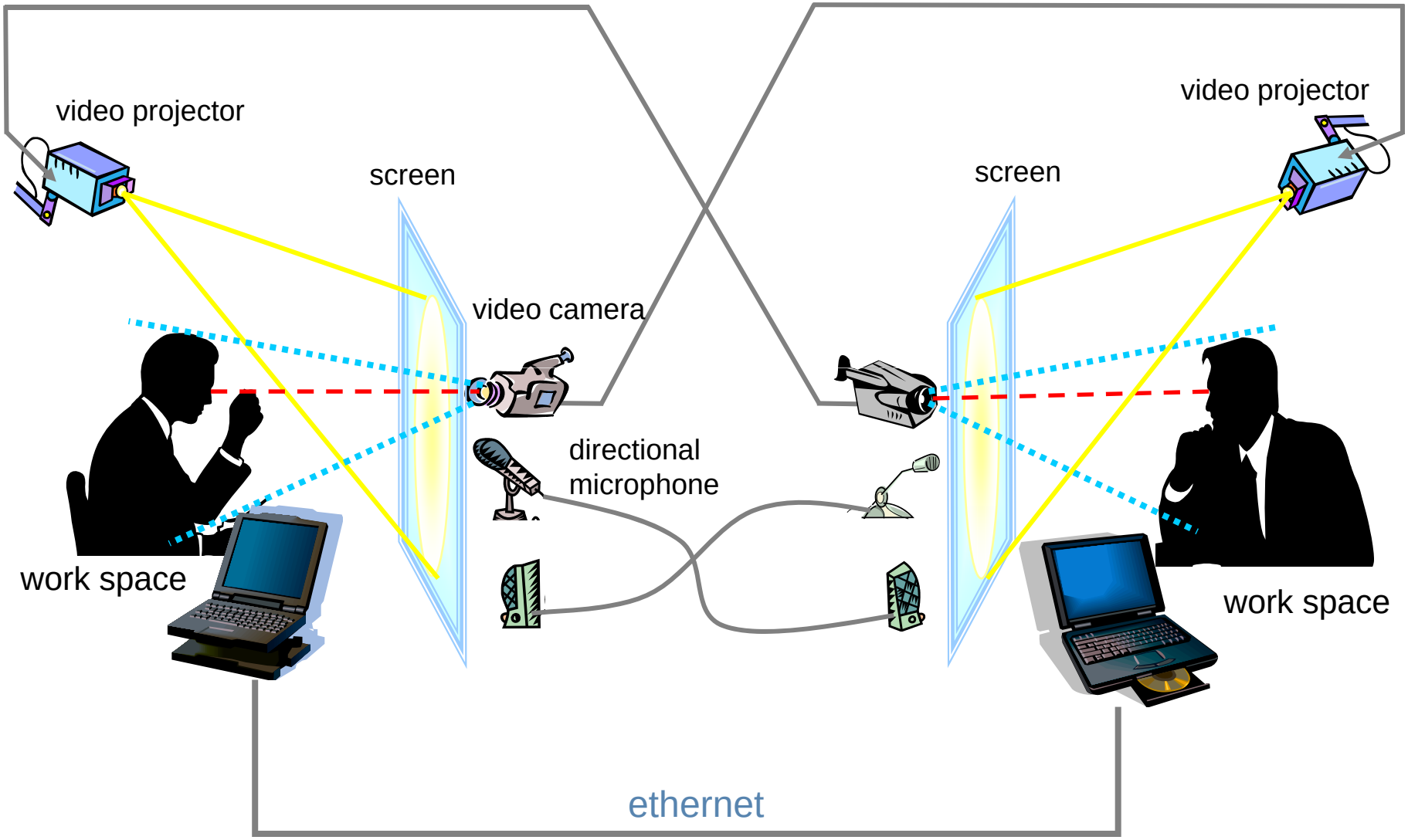
Camera

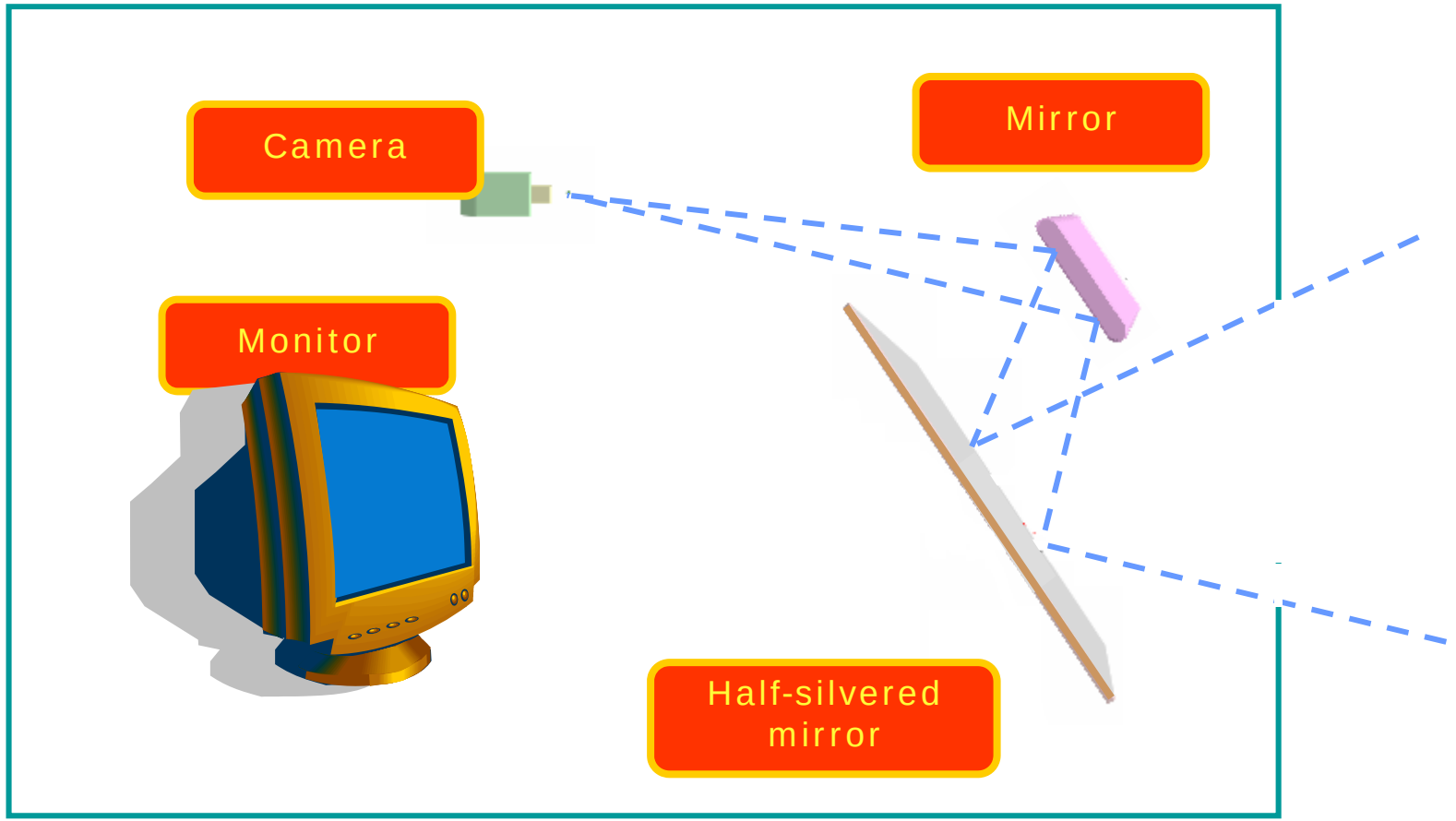
Documents



CuSeeMe – video conferencing on the Internet

Principle of gaze agreement





Video tunnel

Computer supported cooperative work
‘Groupware’

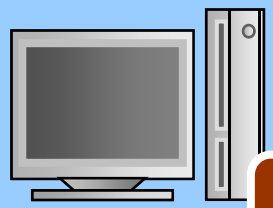
CSCW (computer-supported cooperative work):

Supporting human cooperative work
by computer

||

Dynamic interaction between humans
mediated by computer and network

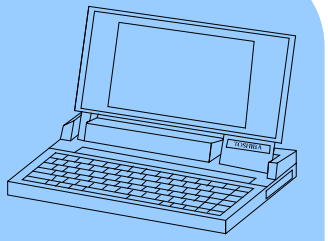
Progress from basic systems to groupware



Organization

Main frame system

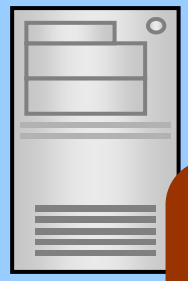
1965-



Individual

PC application

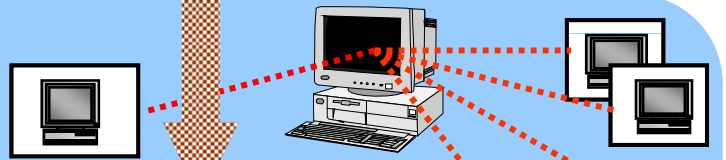
1975-



Project

Mini-computer distributed processing

1970-



Group

Network Group-ware

1985-

Personal computer environment

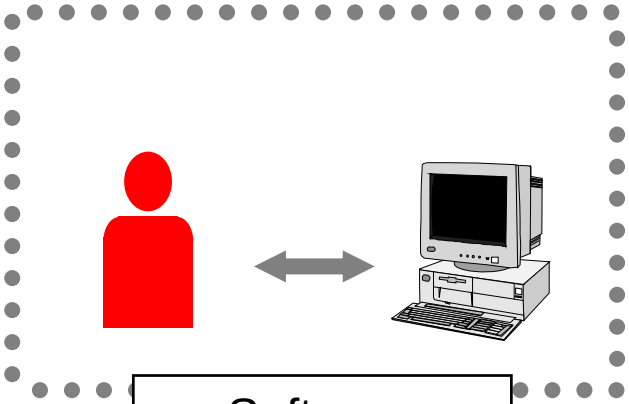
(Past)

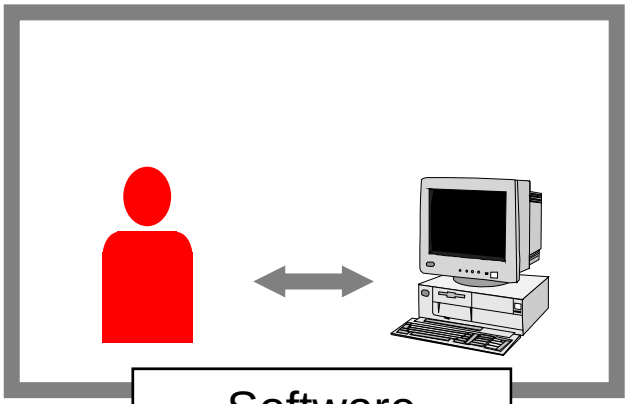
Support of individual work

Human computer interface

User

Computer





Software

Hardware

Software

Hardware

Network

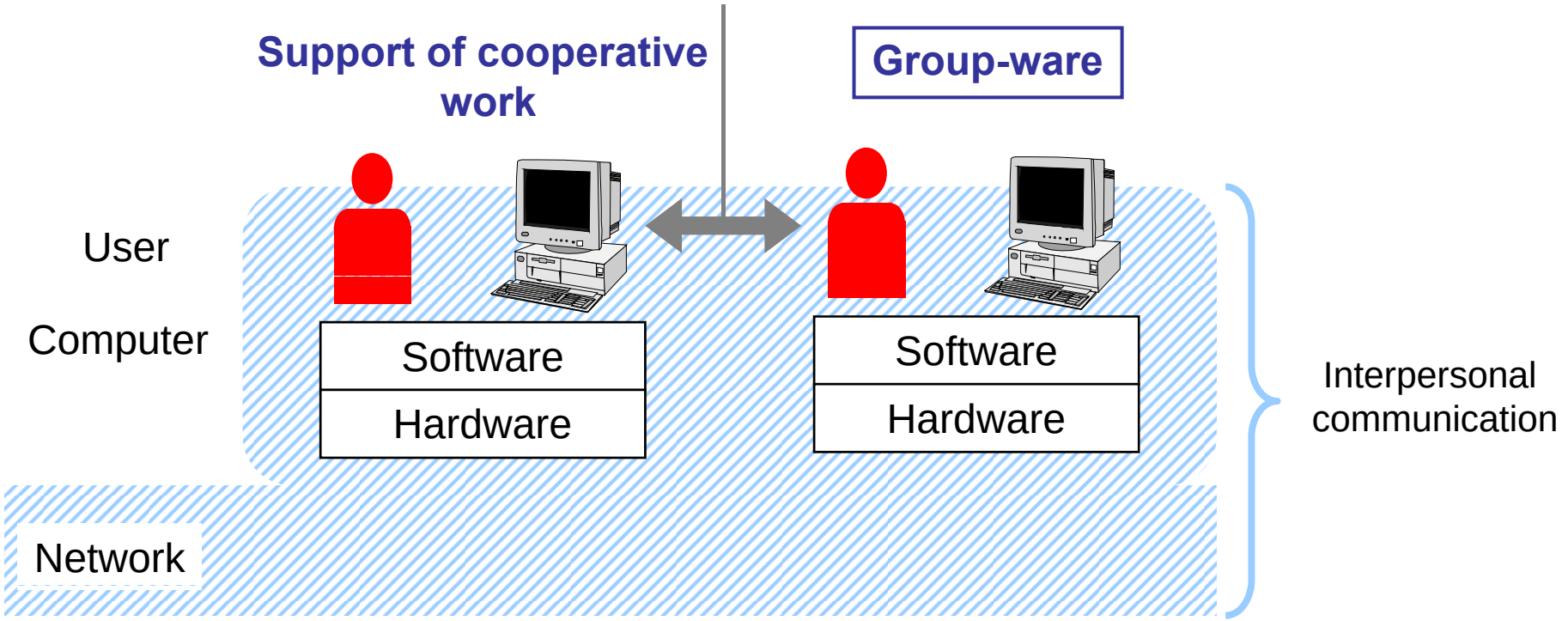
Tele communication

CSCW environment

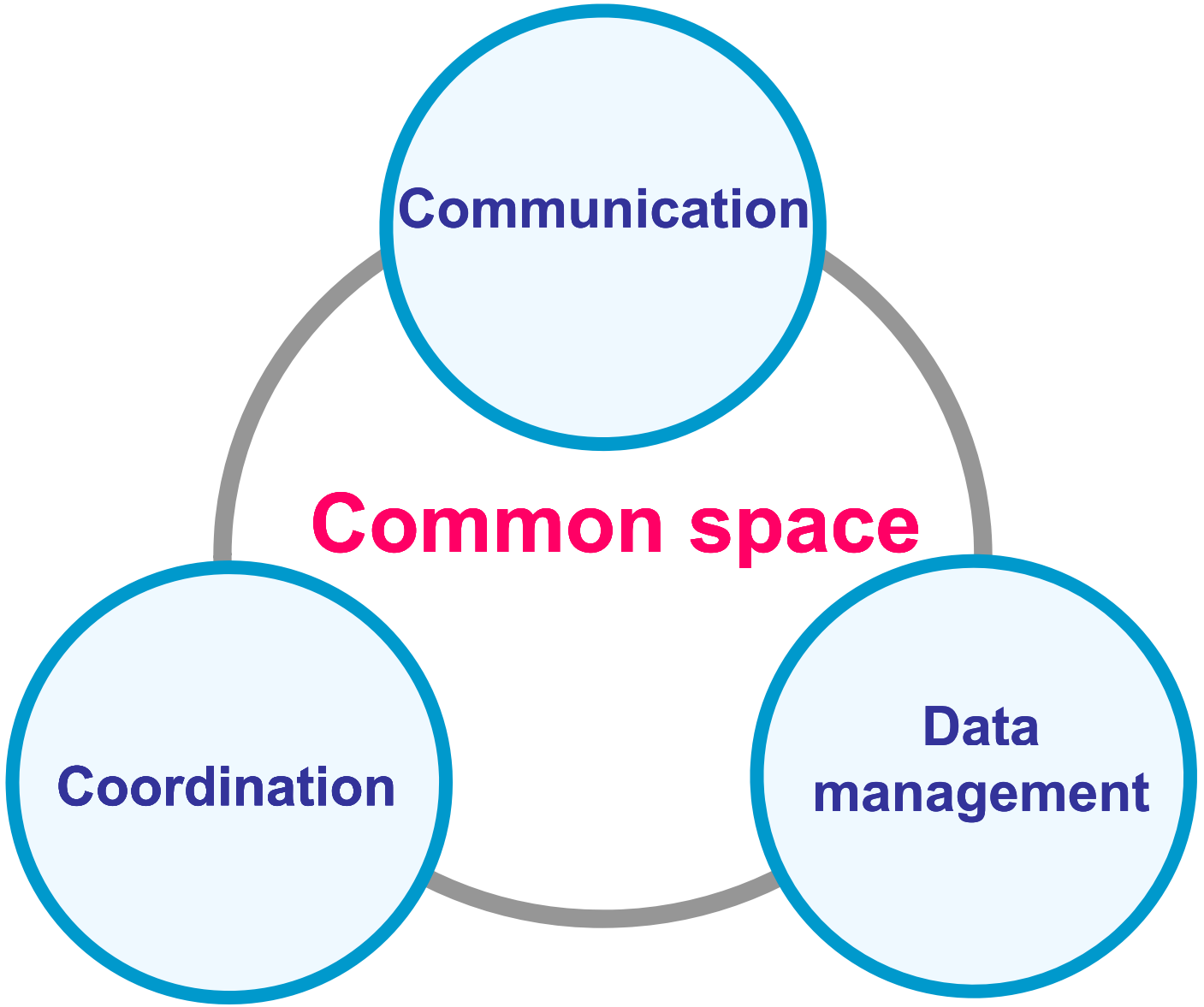
（Future）

Human Computer Interface

（Multi-user interface）



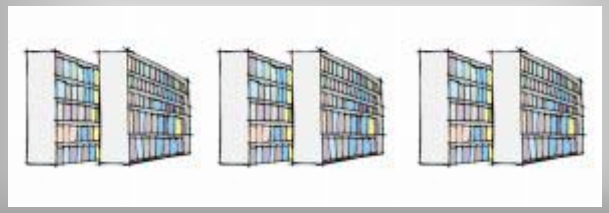
Basic functions of groupware



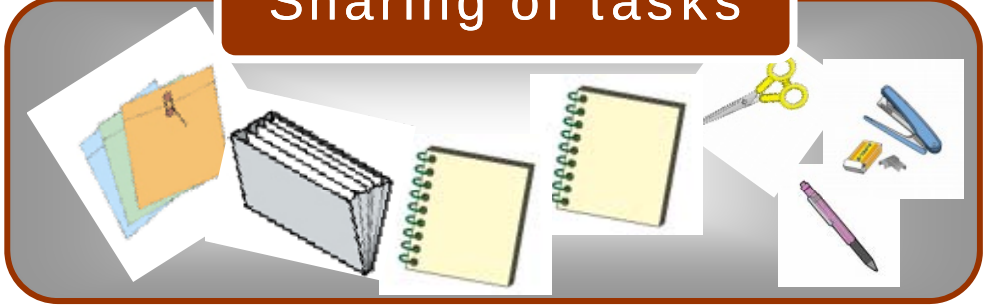
Three sharings necessary for cooperative work

Sharing of awareness

Sharing of information

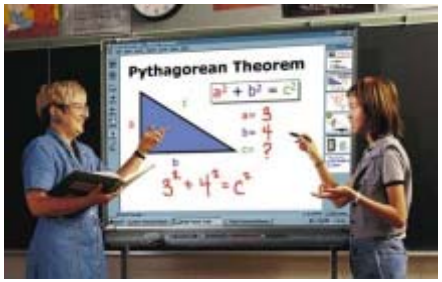


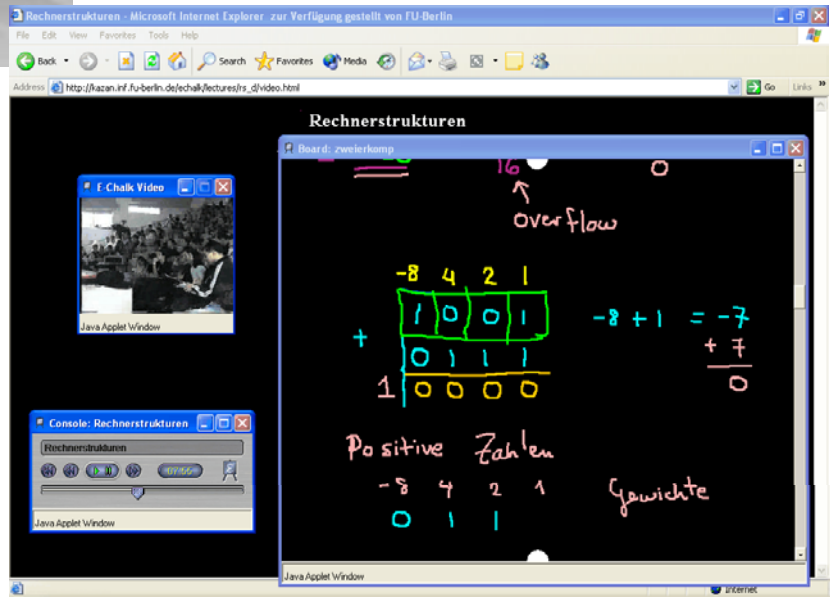
Sharing of tasks



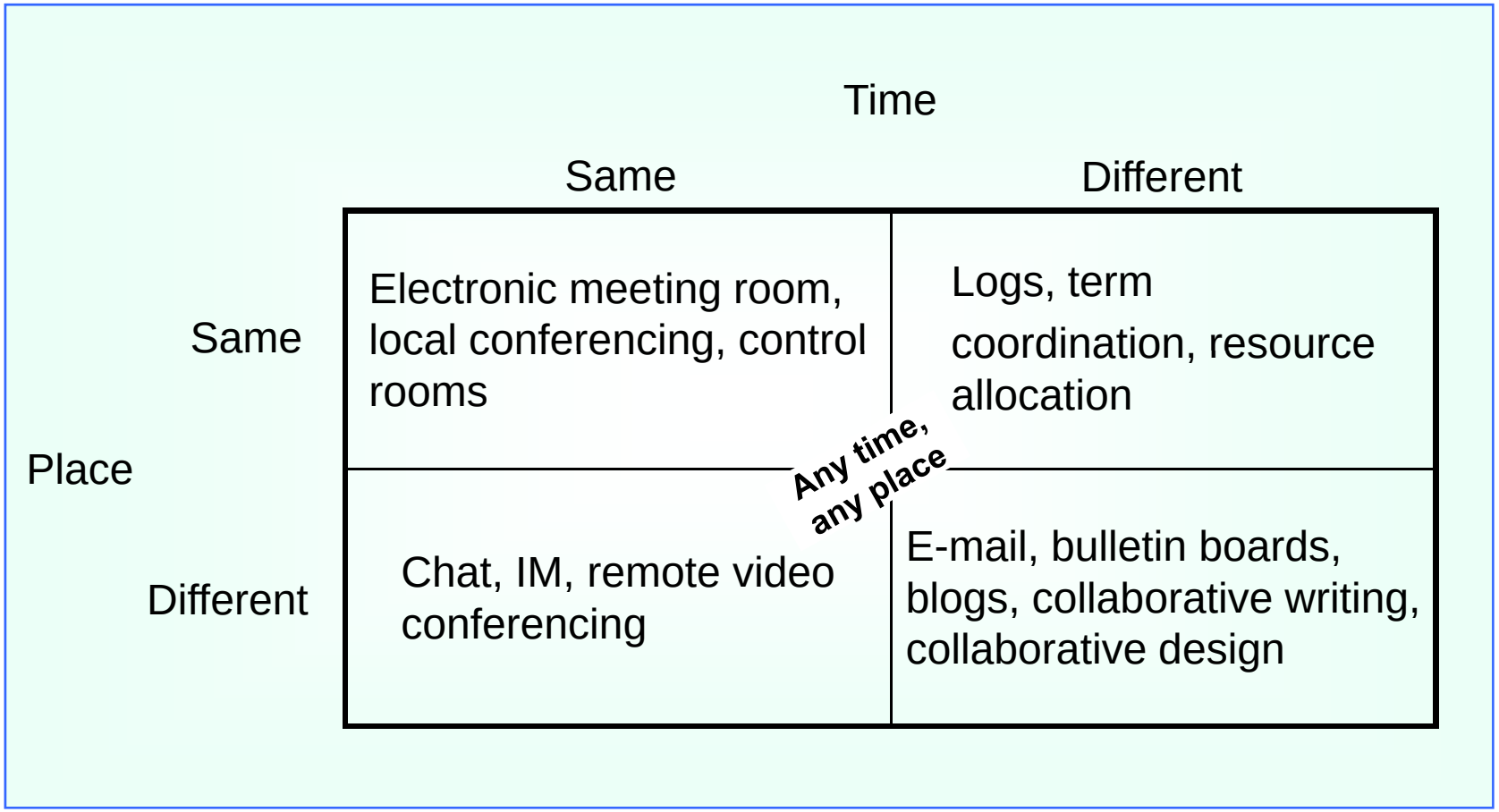
Clear Board

Electronic whiteboard

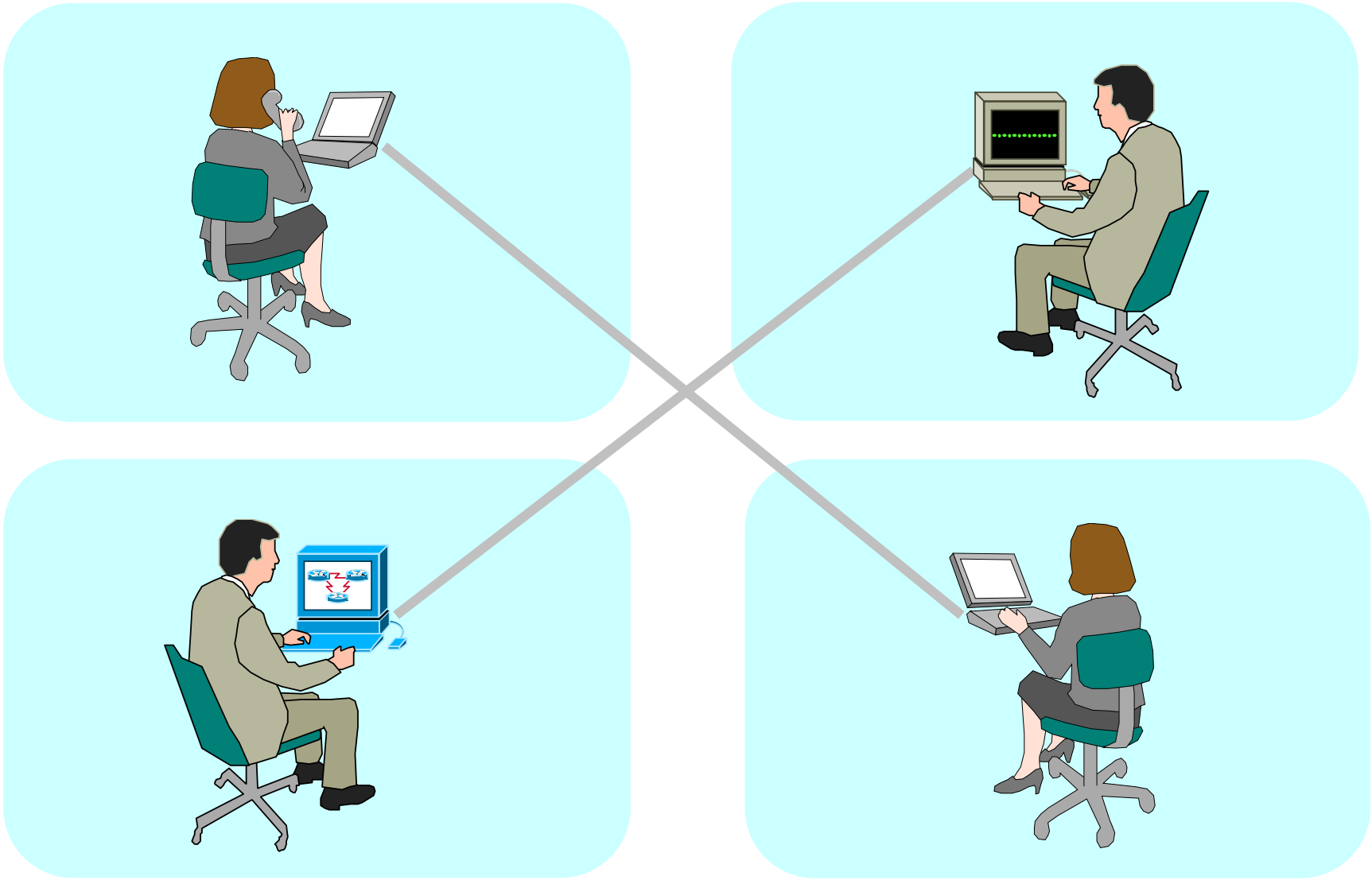




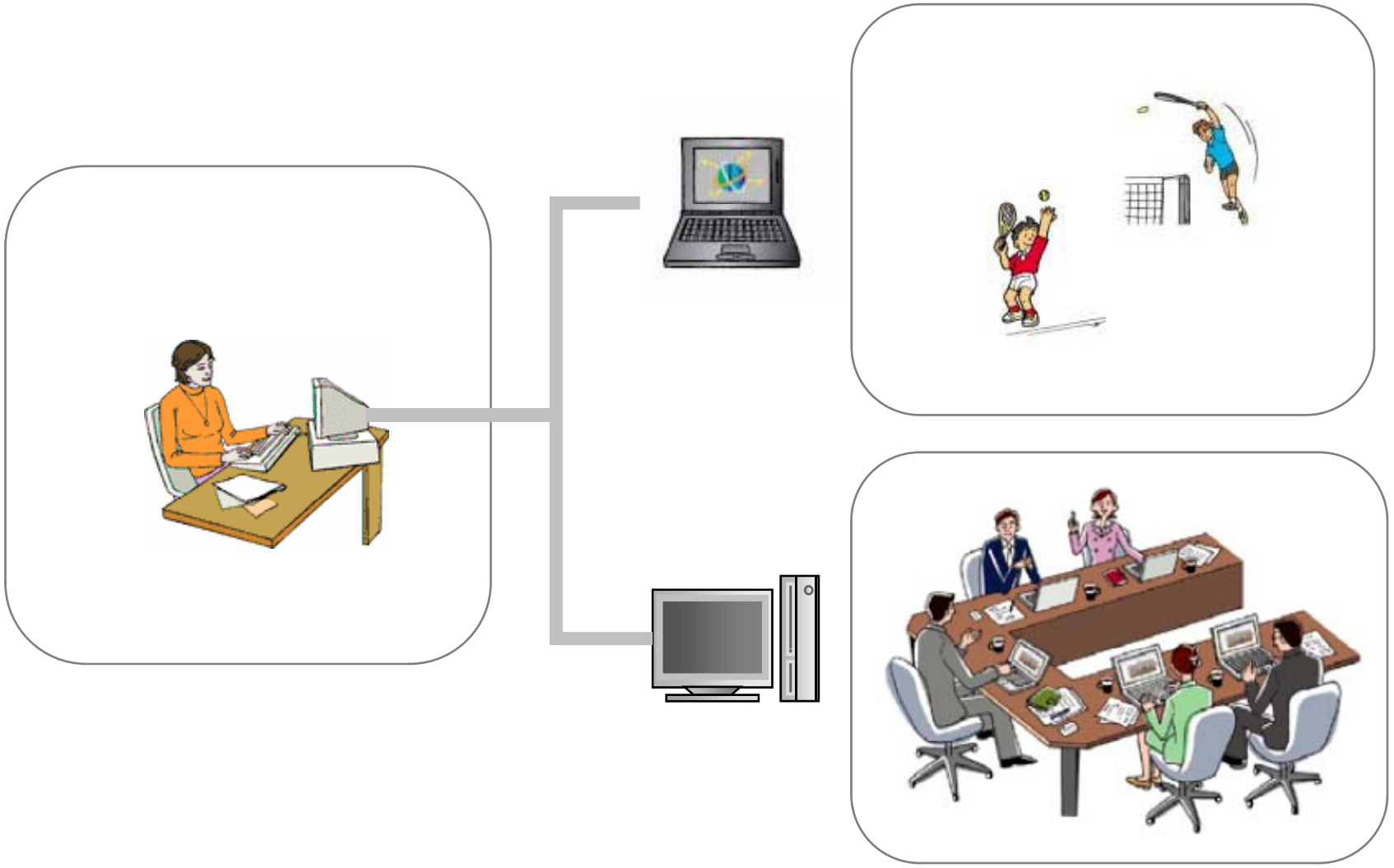
Classification of CSCW



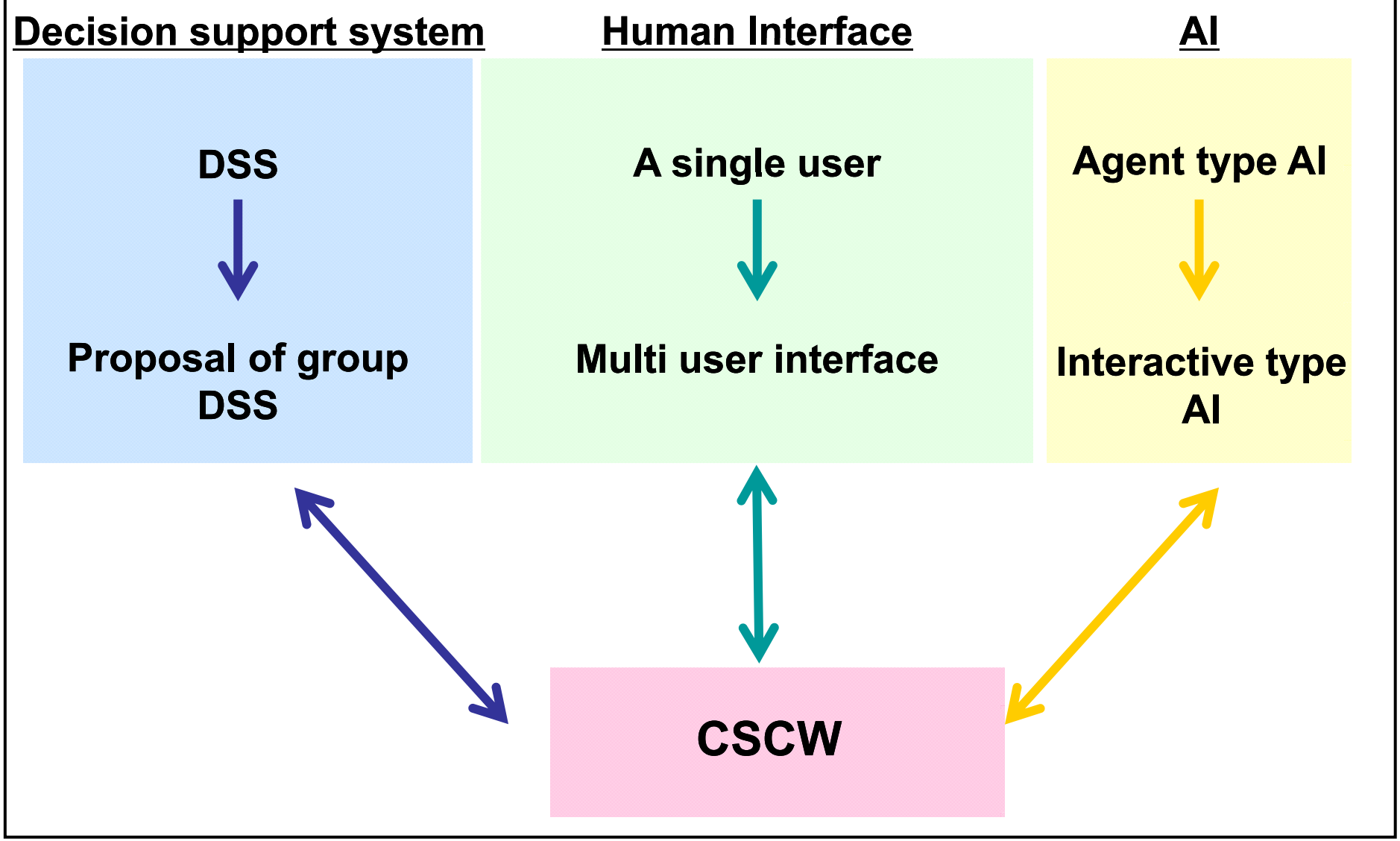
Synchronous and distributed working style



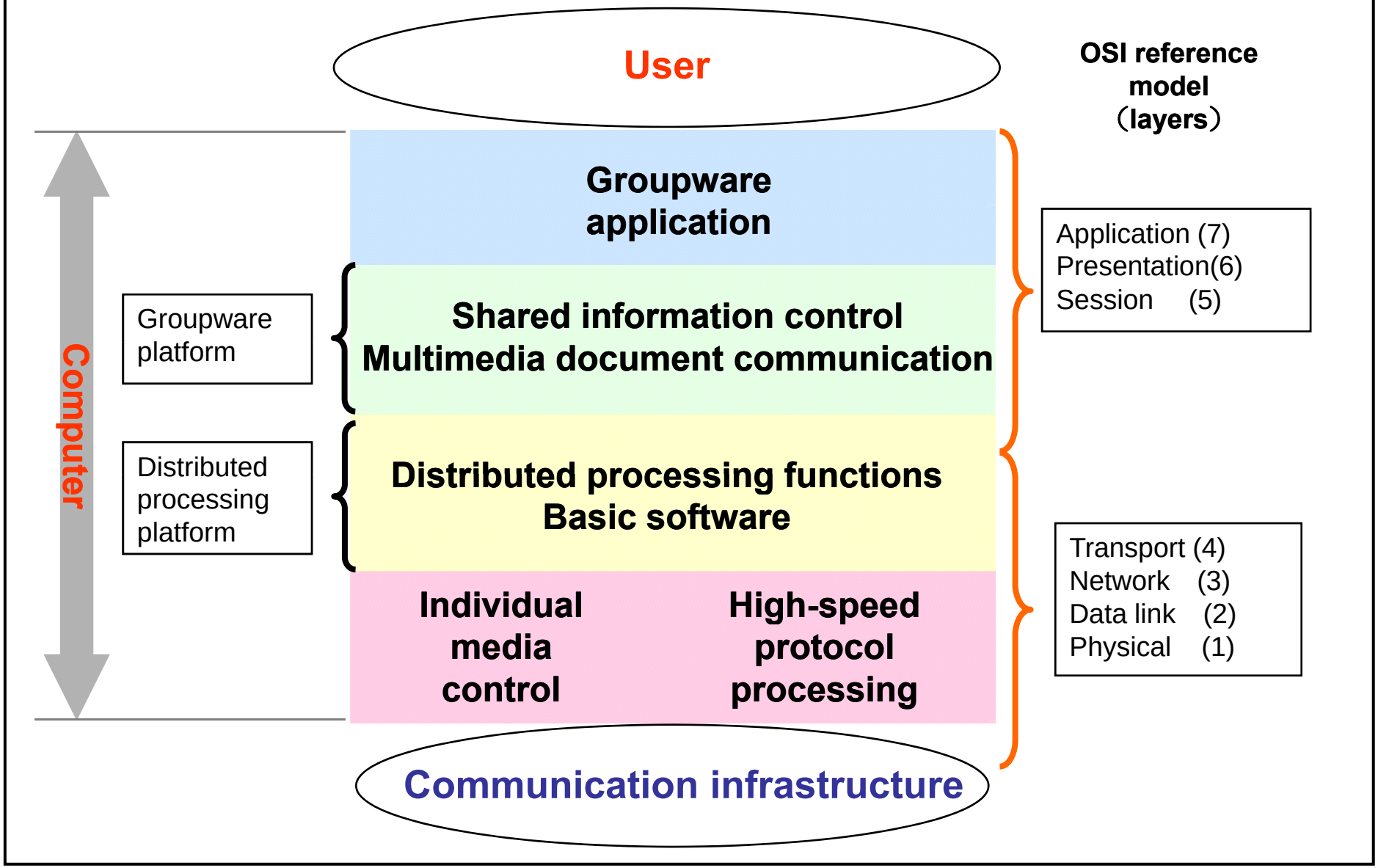
Asynchronous and distributed working style



Background of CSCW research



General architecture of groupware system

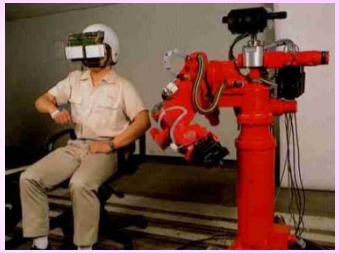


Artificial reality and tele-existence

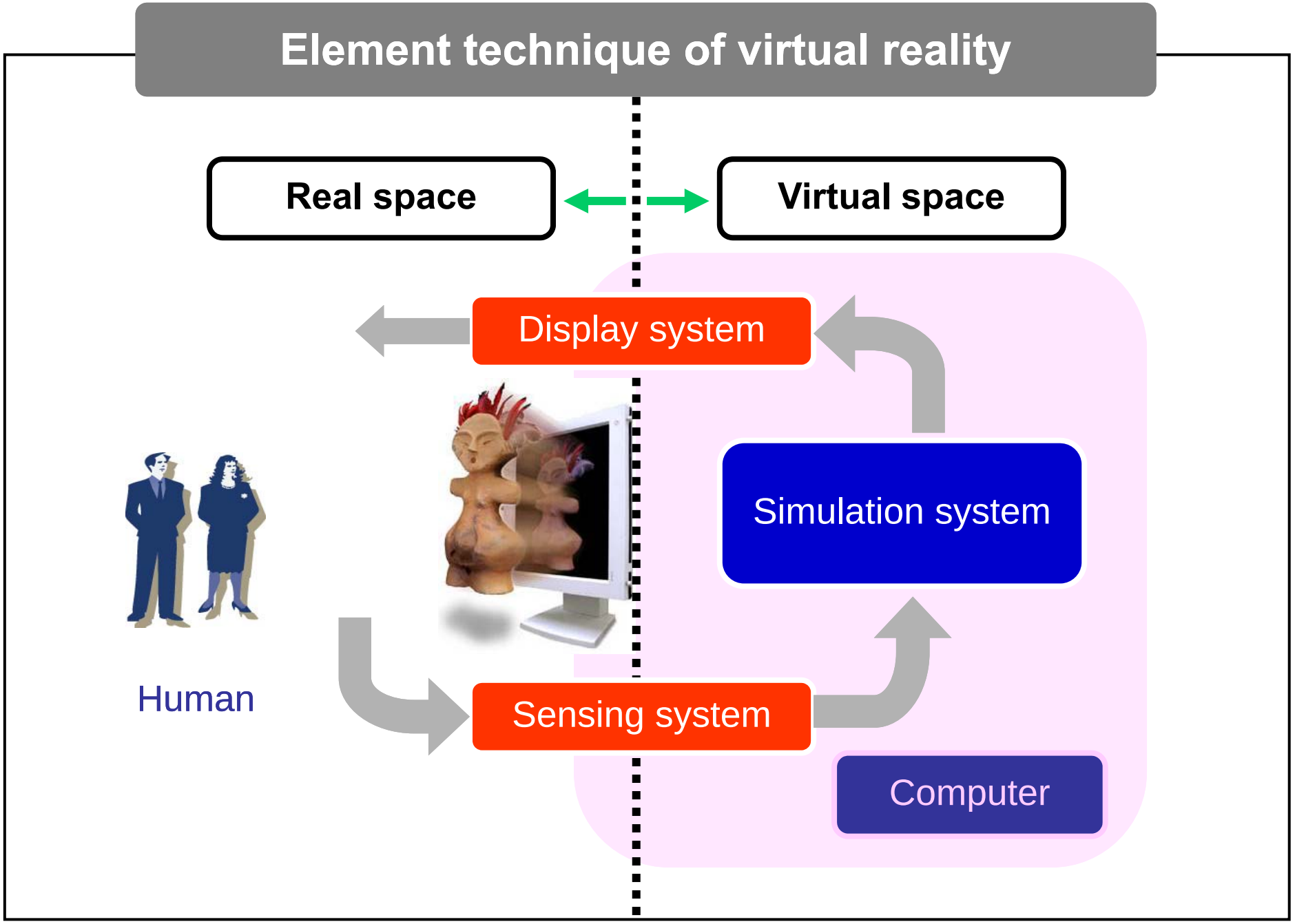
Artificial reality

Virtual reality

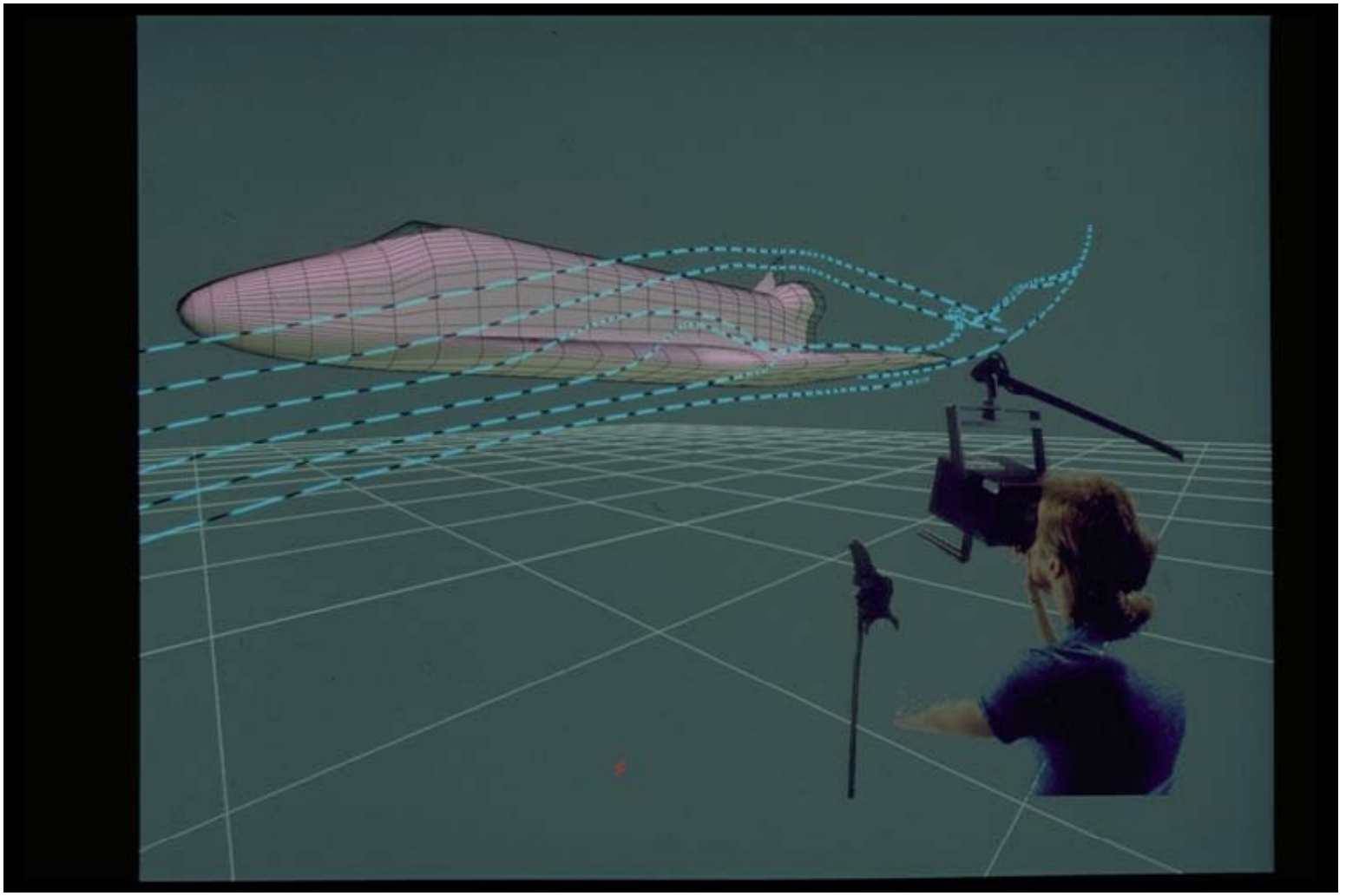
Tele-existence



Element technique of virtual reality



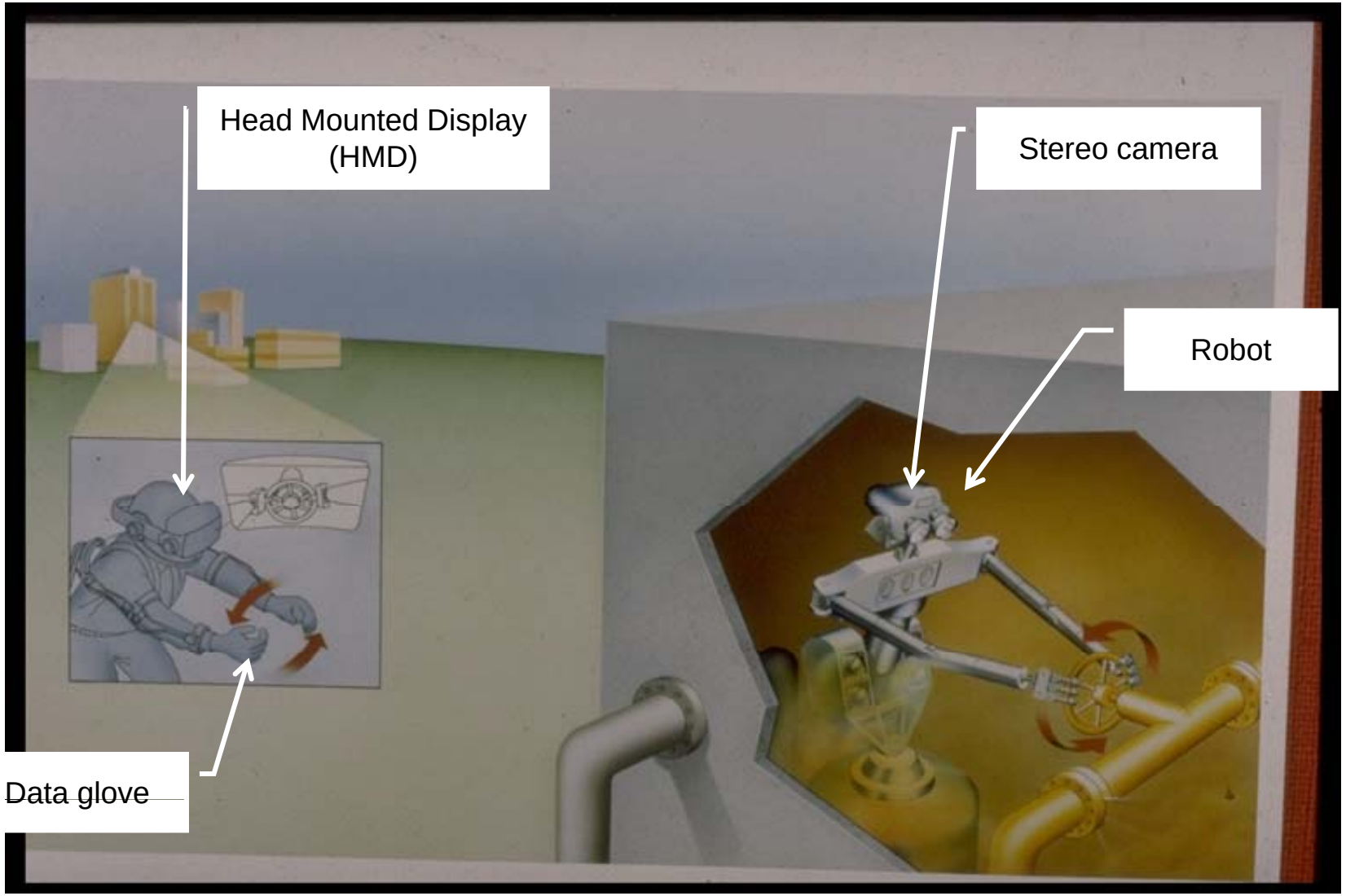
Virtual wind tunnel



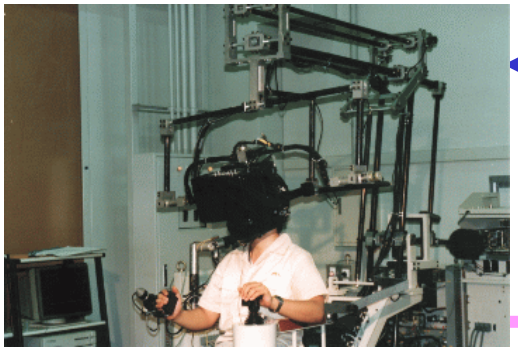
MR type presentation ‘MR car, it is here’

Virtual car is represented with full size. The system is developed jointly by MR System Research and ART (Germany) Com.
Content data is provided by Daimler-Chrysler Co. in Germany.

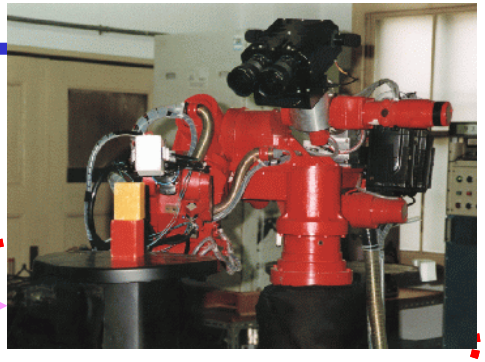
Virtual work environment at NASA Ames Research Center



Tele-existence
(Mechanical Engineering Laboratory, Agency of Industrial Science and Technology)



Presentation of perceptual information
(Visual, audio, tactile)



Transfer of motion information
(Head, eye, hand, body, etc.)

Robot motion control
(Sensor, manipulator, movement)

Virtual space meeting system

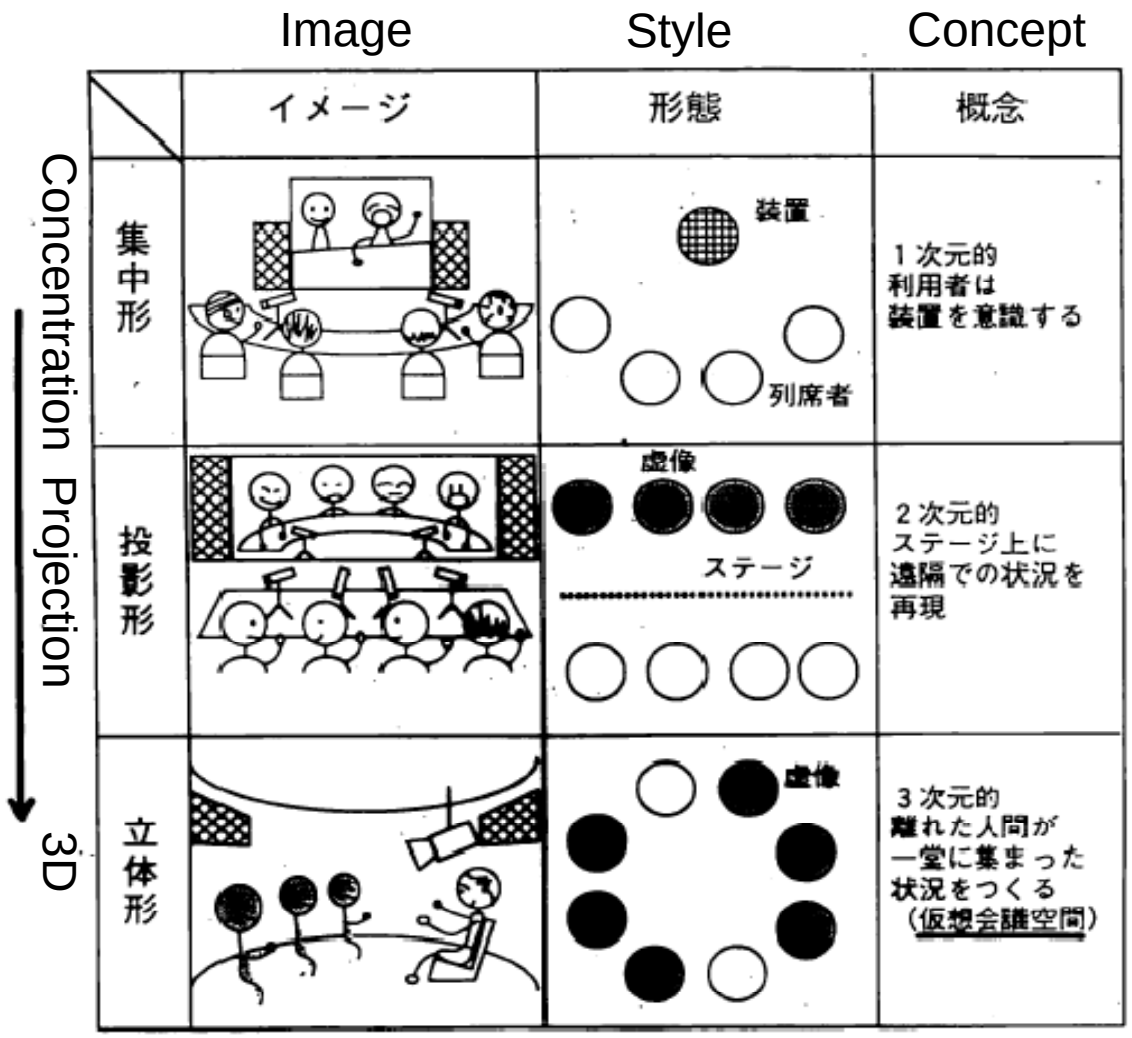
Image communication using large, high definition, 3D display

Sound field communication including room acoustic characteristics

Multiple-party virtual space meeting system (realistic sensation communication meeting)

1. Feeling as if meeting is held at the same place
2. 3D vision
3. Eye contact between participants
4. Image of every desired direction is easily presented
5. Simulation of 3D images is realized at the same space
6. Natural meeting irrespective of the number of multiple parties

Future research: **Extraction of human image from arbitrary scene, motion recognition from human image, 3D CG, operation in virtual space, 3D display, sound field generation in the virtual space (virtual sound source generation, sound source representation, compensation of head movement, etc.)**



One-dimensional

User is aware of equipment

Two-dimensional

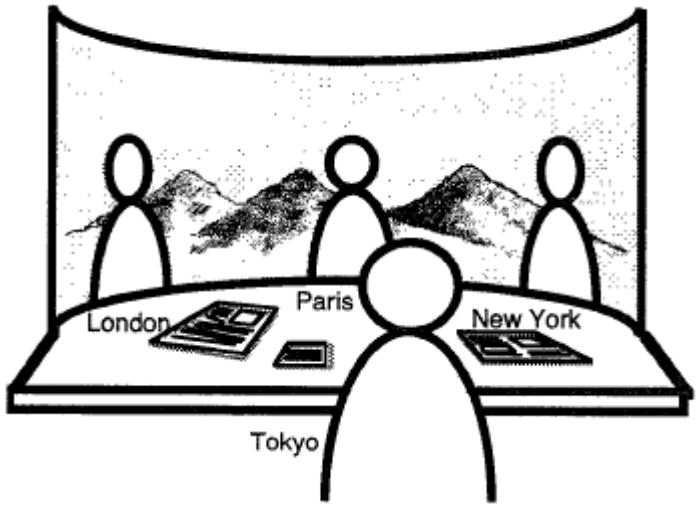
Reproducing the situation of remote place on the stage

Three-dimensional

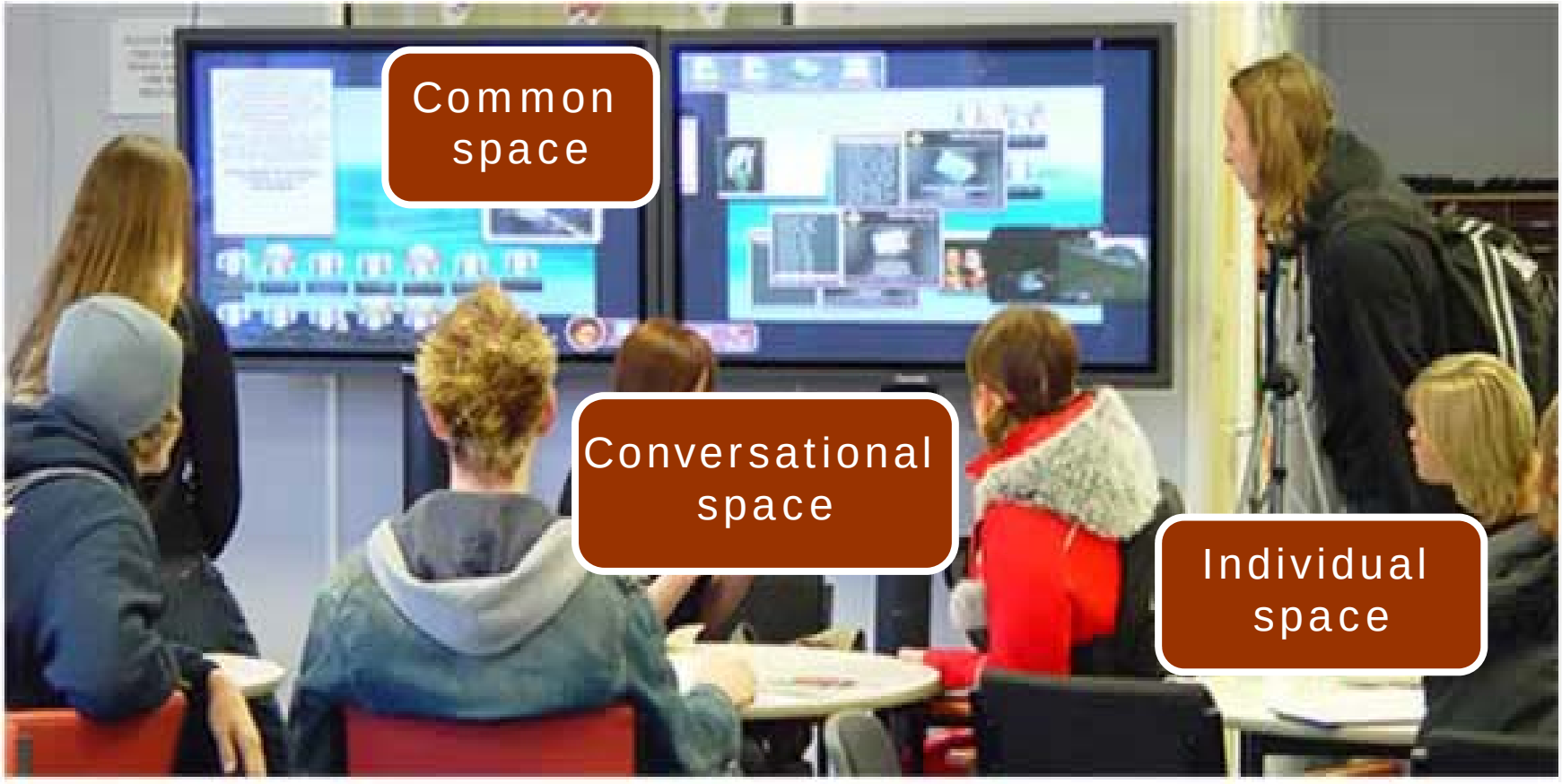
Make the situation where people in the remote places get together in one place (virtual meeting place)

Image of the evolution of tele-conference

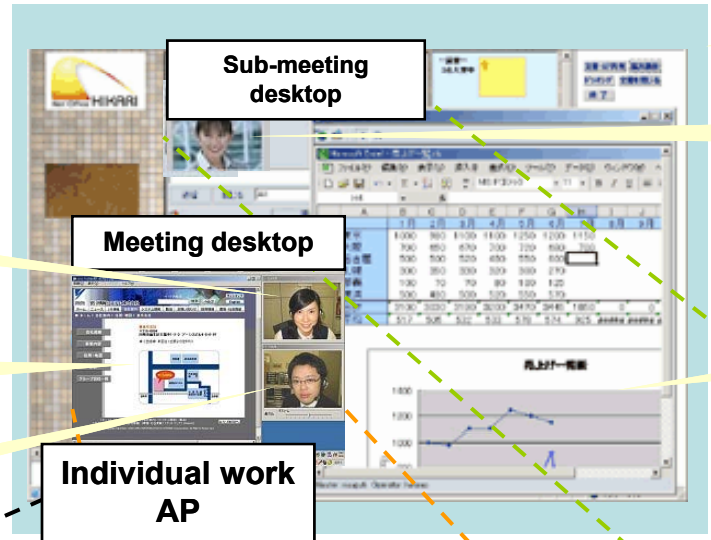
3D video conference



Working space and conversation space



Virtual meeting space

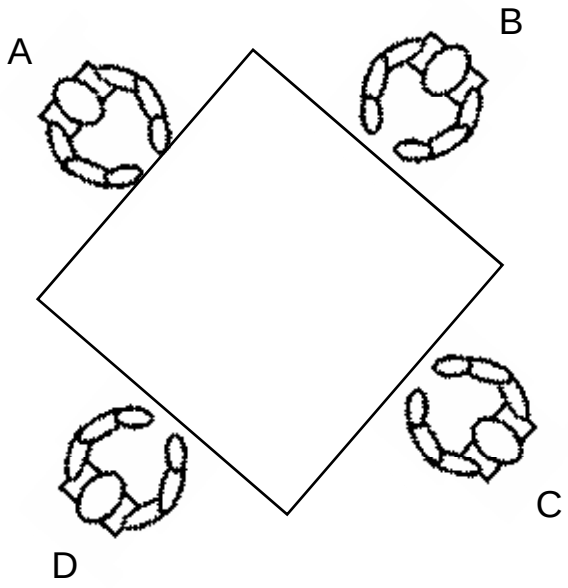


Personal space
(Video phone, making documents)

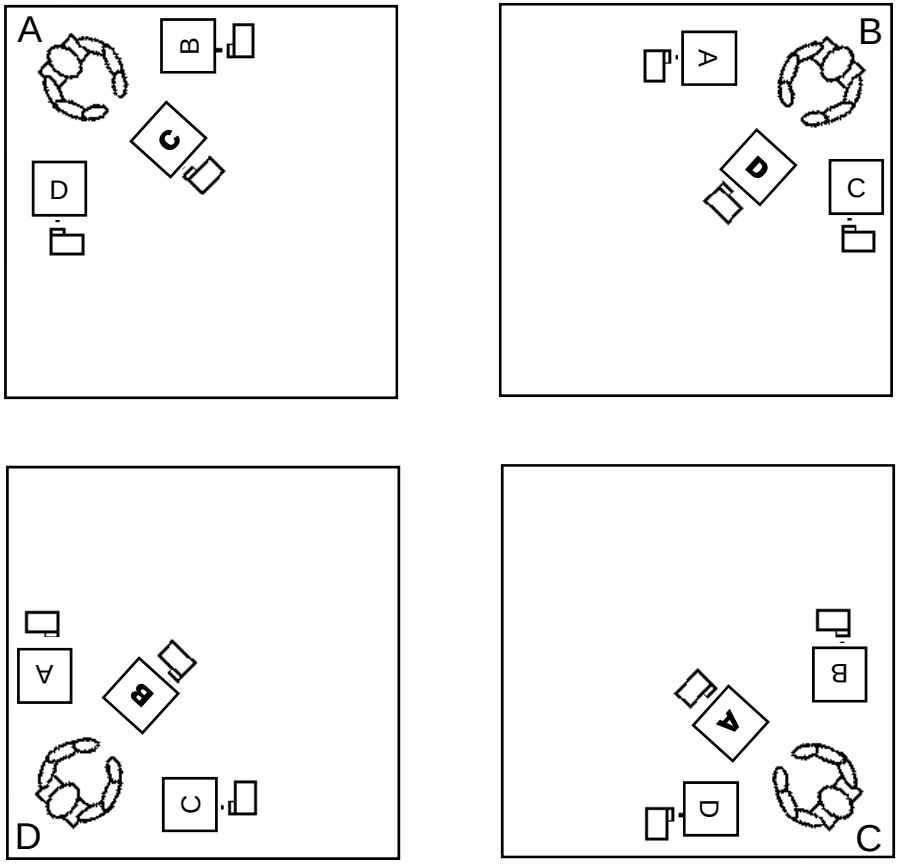
Common space
(Document explanation, presentation, collaborative editing, etc.)

Closed area space
(Private talk during meeting, etc.)

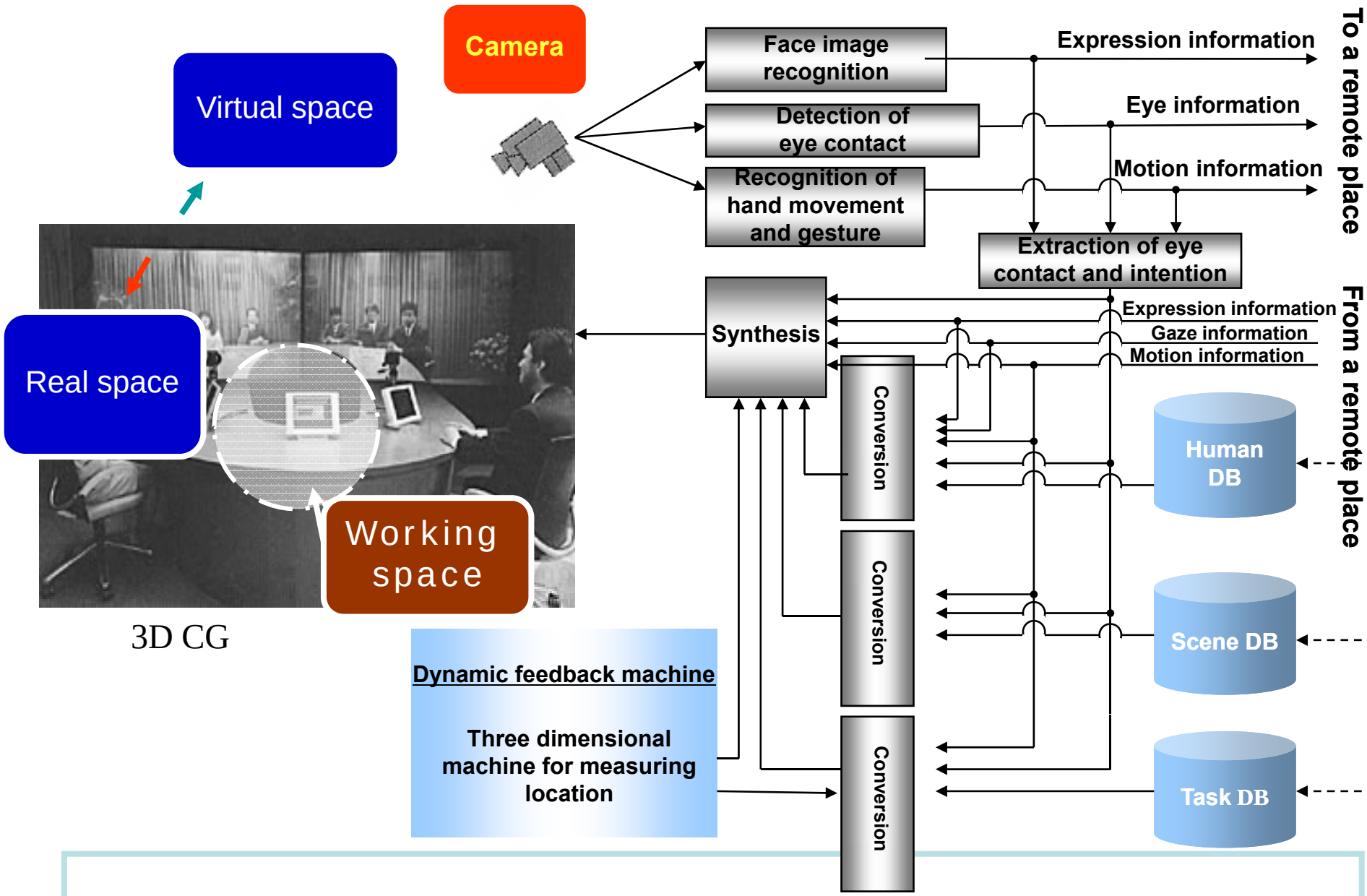
Multiple party virtual conference system (BNR; Bell Northern Research)



Real conference



Structure of the virtual conference system



Realistic sensation tele-conference

3D live theater

Network・virtual reality

Distribution type network VR

Virtual world on the network

Server client model
（World generated by a computer is distributed to several sights）

Sharing type network VR

Interactively shared virtual world

Real-time distributed system, in which individual users have their own avatars in the virtual world and they communicate each other.

Mobile type network VR

Making use of wearable computers and mobile phones

Users moving around the real world with computers having access to the virtual space.