Lecture - 3

Active I/P mode = I/P is provided intentionally. Command to the computer directly. Ex : Speech, Poiting to something, gesture, writing something. Computer will capture the I/P and process accordingly.

Passive I/P mode = I/P is naturally provided. Not give a command directly to the computer. Computer will capture the I/P and process accordingly. Ex : Gaze, Brain wave patterns, Lip movements, Facial Expressions.

Virtual Environment = Input may be anything (Active or Passive) and O/P can be provided virtually. Real life experience.

NOTE : Recognition process will be done for both the active and passive input.

Modality Relations :

CARE properties

CARE = Complementarity, Assignment, Redundancy, Equivalence

Complementarity = multiple modalities are used together at a time to reach O/P. Ex : Speech + Gesture will work together. (Put the red triangle in that place) . All the inputs combined together to provide O/P. (ALL MODALITIES WILL WORK TOGETHER)

Assignment = Only one modality will be selected for use as information. Multiple modalities will be there at I/P side but only one modality will be selected at a time to give O/P. (ONLY ONE MODALITY WILL WORK AT A TIME)

Redundancy = only a part of information will be used. Ex : (IN MIT pointing hand + voice) -> Put the red triangle in the blue circle. If only one triangle is present then pointing hand is redundant and only voice is enough.

Equivalence = Any available modality can be used in any manner. We can give modalities in sequential manner or parallel manner. (SYSTEM CAN WORK WITH ANY COMBINATION OF INPUTS)

Types of multimodal systems :

Fusion based multimodal systems (Blended multimodal systems/interfaces) : Input should have atleast one active I/P and atleast one passive I/P to get the O/P. Ex : LipNet = based on CNN , Lip movement is passive and speech is active. LipNet is used in autonomous vehicles.

Temporally cascaded multimodal system : Process two or more user modalities sequenced in temporal order (modalities working in order one after the another). Input can be anything, that is active or passive. Ex : first gaze + then gesture.