1. The biometric modality is not:

□ fingerprint,

□ voice,

□ password or pin (personal identification number),

□ iris.

1. Determine the proper sequence of stages/phases in a general recognition process.   
   (1 – first, 2 – second, 3 – third, 4 – last).
2. \_\_\_ recognition,
3. \_\_\_ data acquisition,
4. \_\_\_ data registration,
5. \_\_\_ pre-processing.
6. Multi-level user identification uses:

□ user’s voice,

□ user’s face,

□ credentials as a username, password or PIN,

□ combination of user’s voice, face and credentials.

1. Match right definition to each kind of speaker identification system below :

|  |  |  |
| --- | --- | --- |
| Text independent systems |  | use extracted acoustic information e.g. formant frequencies, spectra, etc. |
|  |  |  |
| Text dependent systems |  | use precise phrases or passwords |
|  |  |  |
| Systems utilizing acoustic information |  | use extracted prosodic information e.g. speech dynamic, stress, pauses, etc. |
|  |  |  |
| Systems utilizing prosodic information |  | do not use precise phrases or passwords. |

1. Determine the right sequence of operations/ processes in the automatic speech recognition process.  
   (1 – first, 2 – second, 3 – last).
2. \_\_\_ dictionary search,
3. \_\_\_ digital speech signal,
4. \_\_\_ text sequence.
5. The most successful methods for extracting speech feature are:

□ Hidden Markov models,

□ Mel frequency cepstral coefficients,

□ Perceptual linear prediction,

□ Discrete cosine transform.

1. Match each automatic speech recognition (ASR) system with its proper input definition.

|  |  |  |
| --- | --- | --- |
| ASR recognizing isolated words require |  | the input to be natural speech without grammatical restrictions. |
|  |  |  |
| Dictation systems require |  | the input to be a single word from a dictionary. |
|  |  |  |
| ASR recognizing fluent speech require |  | the input to be fluent speech with some grammatical restrictions. |
|  |  |  |
| ASR recognizing natural speech require |  | the input to be a sequence of words with sufficient pauses separating adjacent words. |

1. Classify given spectral events/modifications to perceivable and non-perceivable by humans.

|  |  |
| --- | --- |
| **Perceivable** | **Non-perceivable** |
|  |  |
|  |  |
|  |  |

**A** – Number of formant frequencies,

**B** – Frequencies laying under the first formant frequency,

**C** – Location of formant frequencies,

**D** – Width of formant frequencies,

**E** – Overall tilt of the spectra,

**F** – Narrow band stop filtering.

1. The most significant speech recognition methods are:

□ Hidden Markov models,

□ Dynamic time warping,

□ Perceptual linear prediction,

□ Discrete cosine transform.

1. The main purpose of dynamic time warping in speech recognition domain is:

□ Feature extraction.

□ Comparison of two sequences of speech features that differ in time duration and to calculate their similarity.

□ Calculation of spectral coefficients.