Multimodal Systems
2025 – 2026

- **This project is part of the final exam.**
- **You deliver your results at least one week before the date of the exam.**
- **The project is presented and discussed with the teacher on the day of the exam.**

A. Development of Python scripts

Results include:

- **A short report (2-3 pages) on the work you made, explaining and justifying possible design choices**
- **The output of the project, i.e., the Python code and possibly related material.**
- **You will need to show the work to the teacher (and possibly answer questions about it) on the day of the final exam.**

Project #01

You have been asked to develop a simple interactive system that enables the user to listen to two audio files by touching two virtual buttons appearing on the left and right. The user is placed at the center of the image captured by the video camera (i.e., the x-coordinate of their center of gravity falls within an interval $[x_A, x_B]$, defining a central stripe in the image). When the user enters the scene, no audio is played back. When the user reaches the button on the right, the first audio file is played back. If the user then reaches the button on the left side, the second audio file is reproduced. The user can continue to activate one audio file or the other by touching the corresponding button.

1. Identify the sensor devices you can use to develop the system.
2. Design the system and identify possible critical aspects in its implementation.
3. Build an initial prototype of this system in Python.

Project #02

You have been asked to develop a system to analyze movement performance. The system receives the 3D trajectories of a collection of body landmarks and computes the average kinetic energy of the movement of selected subsets of them for each detected gesture. Unitizing can be performed by applying a threshold to the kinetic energy. The following subsets of landmarks are considered: whole body, upper body, lower body, right arm, left arm, right leg, and left leg. Then, the system shows how the kinetic energy of the whole body is distributed among the six subsets. This can be represented, for example, as a bar chart with the six subsets on the x-axis and, for each of them, the percentage of computed kinetic energy (with respect to the total kinetic energy of the body) on the y-axis. Moreover, the system displays whether one of the subsets is characterized by a very large amount of kinetic energy compared to the others and displays an appropriate message, e.g., “The right leg is moving a lot” to indicate that the user is primarily moving the right leg.

1. Identify the sensor devices you can use to develop the system.

2. Design the system and identify possible critical aspects in its implementation.
3. Build an initial prototype of this system in Python.

Project #03

You have been asked to develop an application for rehabilitation purposes that measures, in real-time, the variation in the position of the edges of a user's bounding rectangle over time. Such a variation can be computed as:

$$\Delta B_L(t) = \frac{|x_L(t-1) - x_L(t)|}{\Delta t} \text{ for the left edge and}$$
$$\Delta B_R(t) = \frac{|x_R(t-1) - x_R(t)|}{\Delta t} \text{ for the right edge}$$

where x_R and x_L are respectively the *x* coordinates of the right and left edges of the bounding rectangle, computed at times $t - 1$ and t, and Δt is the sampling period (e.g., 40 ms). You have to implement the computation of such a variation as a Python script. To show the result of the computation, make two graphs that display ΔB_L and ΔB_R over time.

1. Identify the sensor devices you can use to develop the system.
2. Design the system and identify possible critical aspects in its implementation.
3. Build an initial prototype of this system in Python.

Project #04

You have been asked to develop an alarm system for detecting possible dangers for a fragile patient. Dangers consist of (i) going too close to a stair and/or (ii) falling. The system plays back a warning message when the user comes too close to the stairs and triggers an alarm message for assistance when the user falls (you can simulate the alarm message by just displaying the message string). Consider the user moving in a space captured by an input video camera, with stairs placed on the left side of the image. If the x-coordinate of the center of gravity falls below a given threshold, the user is considered to be too close to the stairs. If the y-coordinate of the center of gravity of the user crosses a threshold corresponding to the floor and remains there for more than 10 seconds, a fall is detected, and the alarm message is triggered.

1. Identify the sensor devices you can use to develop the system.
2. Design the system and identify possible critical aspects in its implementation.
3. Build an initial prototype of this system in Python.

Project #05

You have been asked to develop an interactive system that uses full-body movement to control switching between videos in an archive. For the sake of simplicity, assume the archive contains six videos. When the user enters the scene, a default video is displayed. Every time the user stops moving, one video in the archive is displayed: when the user stops for the first time, the system shows the first video; when the user stops for the second time, the system shows the second video; when the user stops again, the system switches to the next video, and so on. After the last video is displayed, the next pause in the movement makes the system switch back to the first video.

1. Identify the sensor devices you can use to develop the system.
2. Design the system and identify possible critical aspects in its implementation.
3. Build an initial prototype of this system in Python.

Project #06

An automatic audio composition system utilizes four audio fragments, each associated with a distinct range of the user's movement energy and contraction. That is, the first fragment is associated with high values of both energy and contraction, the second one is associated with high values of energy and low values of contraction, the third one is associated with low values of energy and high values of contraction, and the fourth one is associated with low values of both energy and contraction. When the end of a motion unit is detected, the average values of energy and contraction are computed, and the audio fragment located in the corresponding quadrant of such a feature space is played back.

1. Identify the sensor devices you can use to develop the system.
2. Design the system and identify possible critical aspects in its implementation.
3. Build an initial prototype of this system in Python.

Project #07

You have been asked to develop an interactive system for an art installation. The focus is on a visitor to the installation. Every time the visitor stops moving, a picture of the visitor is taken (thus showing their posture). The picture is then blended into the background image. You may implement the blending of the visitor's captured posture by cutting their silhouette and pasting it into the current video frame. At the next stop, a new picture replaces the old one. For the sake of simplicity, suppose that the installation is experienced by just one visitor at a time.

1. Identify the sensor devices you can use to develop the system.
2. Design the system and identify possible critical aspects in its implementation.
3. Build an initial prototype of this system in Python.

Project #08

You have been asked to develop a system to analyze movement performance in terms of Laban's Effort. The system tracks the 3D trajectories of a collection of body landmarks and computes a simplified representation of Laban's Effort for one representative landmark selected in the collection. The following features are computed: the Directness Index for the Space component of Laban's Effort, the normalized average of the magnitude of acceleration for the Time component of Laban's Effort, and the normalized maximum kinetic energy for the Weight component of Laban's Effort. The features for Time and Weight are normalized by passing them through a sigmoid function. For the sake of simplicity, the parameters of the sigmoid can be set empirically. The normalized features are then represented in a 3D feature space, i.e., in the region [0, 1] x [0, 1] x [0, 1], where Space is on the x axis (0 Direct, 1 Flexible), Time is on the y axis (0 Sustained, 1 Quick), and Weight is on the z axis (0 Heavy, 1 Light). This space represents Laban's cube of Effort. The system displays as output the label of the Effort type that is closest to the computed position in the feature space.

1. Identify the sensor devices you can use to develop the system.
2. Design the system and identify possible critical aspects in its implementation.
3. Build an initial prototype of this system in Python.

Project #09

You have been asked to develop an interactive system enabling amateur sportsmen to perform exercises with their arms. The exercise consists of a series of five leg extensions, with the left leg extended to the left and the right leg extended to the right. An extension of the leg in a direction is successfully performed when

the corresponding body landmark is far enough from the body's center of gravity. Moreover, for each extension, legs should remain stretched for at least 3 seconds. The user is rewarded when the complete series of extensions (i.e., five extensions of the left leg and five extensions of the right leg) is successfully completed. The reward consists of playing a victory sound.

1. Identify the sensor devices you can use to develop the system.
2. Design the system and identify possible critical aspects in its implementation.
3. Build an initial prototype of this system in Python.

Project #10

You have been asked to develop an interactive system for an art installation: when a visitor enters the room, two "walls" appear on the right and left sides of the visitor's bounding rectangle, attempting to crush their body. The visitor can push away the walls by enlarging the bounding rectangle, thus seeing the background of the room. If the visitor's energy is lower than a given threshold, the visitor disappears, and the two walls join at the center of the image. You can build the wall by drawing two rectangles: one from the left edge of the image to the left edge of the bounding rectangle, and one from the right edge of the bounding rectangle to the right edge of the image.

1. Identify the sensor devices you can use to develop the system.
2. Design the system and identify possible critical aspects in its implementation.
3. Build an initial prototype of this system in Python.

B. Critical reading of scientific papers

Results include:

- **A poster (a size-A0 PowerPoint slide) summarizing the paper you read (objectives, methodology, obtained results, discussion).**
- **A short presentation (10 minutes plus possible questions) to be done on the day of the final exam.**

Project #11

C. Bustos, A. Sole-Ribalta, N. Elhaouij, J. Borge-Holthoefer, A. Lapedriza and R. Picard, "Analyzing the Visual Road Scene for Driver Stress Estimation," in *IEEE Transactions on Affective Computing*, vol. 16, no. 3, pp. 1787-1801, July-Sept. 2025.

Project #12

A. Derington, H. Wierstorf, A. Özkil, F. Eyben, F. Burkhardt and B. W. Schuller, "Testing Correctness, Fairness, and Robustness of Speech Emotion Recognition Models," in *IEEE Transactions on Affective Computing*, vol. 16, no. 3, pp. 1929-1941, July-Sept. 2025.

Project #13

F. Boucaud, C. Pelachaud and I. Thouvenin, "Vicarious Evaluation of a Decision Model for Human-Agent Social Touch Interactions," in *IEEE Transactions on Affective Computing*, vol. 16, no. 1, pp. 277-289, Jan.-March 2025.

Project #14

H. Niknazar and S. C. Mednick, “A Multi-Level Interpretable Sleep Stage Scoring System by Infusing Experts’ Knowledge Into a Deep Network Architecture,” in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 7, pp. 5044-5061, July 2024.

Project #15

A. Singh, M. Hrobat, S. Gui, N. Bianchi-Berthouze, J. Ley-Flores, F. Bevilacqua, J. R. Díaz Durán, E. Márquez Segura and A. Tajadura-Jiménez, “Pushed by Sound: Effects of Sound and Movement Direction on Body Perception, Experience Quality, and Exercise Support,” in *ACM Transactions on Computer-Human Interaction* vol. 31, no. 4, article 53, August 2024.

Project #16

S. Gauthier, S. M. Anzalone, J. Xavier, S. Boucenna, A. Berthoz, M. Chetouani, D. Cohen, “Is there a specific motor signature of autism? Exploratory studies from human-machine interaction,” in *IEEE Transactions on Cognitive and Developmental Systems*, published online, 2025.

Project #17

J. Dong and Z. Cai, “Quantitative Estimation of Human Height and Weight Using Motion Data From Multiple Smart Devices,” in *IEEE Transactions on Computational Social Systems*, vol. 12, no. 2, pp. 708-724, April 2025.

Project #18

F. Mueller, M. Obrist, F. Altarriba Bertran, N. Makam, S. Kim, C. Dawes, P. Marti, M. Mancini, E. Ceccaldi, N. Pasumarthy, S. Claire, K. Seo Jung, J. Deng, J. Steimle, N. Krasteva, M. Schwalk, H. Reiterer, H. Wang, Y. Wang, “Grand challenges in human-food interaction”, in *International Journal of Human-Computer Studies*, vol. 183, article 103197, 2024.

Project #19

M.-H. Tayarani-N. and A. Vinciarelli, “Learning Classifier Performance in an Ensemble of Classifiers for Personality Prediction Using Laughter,” in *IEEE Transactions on Affective Computing*, published online, 2025.

Project #20

L. Maman, N. Lehmann-Willenbrock, M. Chetouani, L. Likforman-Sulem and G. Varni, “Modeling the Interplay Between Cohesion Dimensions: A Challenge for Group Affective Emergent States,” in *IEEE Transactions on Affective Computing*, vol. 15, no. 3, pp. 1526-1538, July-Sept. 2024.

C. Something that you like…

If you have any ideas or concepts related to the course content that you would like to develop, you can propose them to the teacher and discuss them with him.