Universidad de Alcalá

TEACHING GUIDE

Voice and Audio Processing

Degree in
Telecommunication Systems Engineering (GIST)

Universidad de Alcalá

Academic Year 2023/2024

4th Year - 1st Semester (GIST)

Approved by the School Board on June 12th, 2023

TEACHING GUIDE

Course Name:	**Voice and Audio Processing**
Code:	**390007 (GIST)**
Degree in:	**Telecommunication Systems Engineering (GIST)**
Department and area:	**Teoría de la Señal y Comunicaciones** **Signal Theory and Communications**
Type:	**Optional (Oriented) (GIST)**
ECTS Credits:	**6.0**
Year and semester:	**4th Year - 1st Semester (GIST)**
Teachers:	Roberto Gil Pita (coordinator), Manuel Rosa Zurera, Manuel Utrilla Manso
Tutoring schedule:	Consultar al comienzo de la asignatura
Language:	Spanish/ English Friendly

1. COURSE SUMMARY

Since the invention of the telephone by Alexander Graham Bell, engineers and scientists have studied the phenomenon of oral communication, with the aim of creating more efficient systems of man-man and man-machine communication.These studies were strongly promoted from the decade of the 60s of the twentieth century, thanks to the development of digital signal processing, which assumed a central role in this issue. Studies on digital voice processing were the basis in the 80s of the twentieth century to create new audio signal processing systems, on which important commercial developments in the world of Telecommunications have arisen.

Parallel developments in the field of electronic technology and computer architecture, and advances in the theory of digital signal processing have been aligned to create a technological environment without virtual limits to create novel applications in telecommunication with voice and audio signals.

This subject focuses on the role of digital signal processing in the development of this type of applications. Basic concepts are presented about the nature of the signals to be treated, the acoustic models of production of these signals, the existing models for the hearing organ, and how these models can be used in applications such as voice and audio coding, synthesis audio, and spatial filtering.

In the subject, other techniques will also be studied, such as methods of analysis by synthesis, time-frequency analysis, homomorphic signal analysis, adaptive filtering, array processing, which find applications in many other problems related to telecommunication technologies.

Prerequisites and Recommendations:

For a good understanding of the subject, it is advisable to have studied the course Digital Signal Processing.

2. SKILLS

Basic, Generic and Cross Curricular Skills.

This course contributes to acquire the following generic skills, which are defined in the Section 3 of the Annex to the Orden CIN/352/2009:

en_TR1 - Knowledge, understanding and ability to apply the necessary legislation during the development of the profession of Technical Engineer of Telecommunication and ease of handling specifications, regulations and mandatory rules.

en_TR3 - Aptitude to solve problems with initiative, decision making, creativity, and to communicate and to transmit knowledge, skills and workmanship, comprising the ethical and professional responsibility of the activity of the Technical Engineer of Telecommunication.

en_TR8 - Capacity of working in a multidisciplinary and multilingual team and of communicating, both in spoken and written language, knowledge, procedures, results and ideas related to telecommunications and electronics.

Professional Skills

This course contributes to acquire the following professional skills, which are defined in the Section 5 of the Annex to the Orden CIN/352/2009:

en_CST1 - Ability to build, operate and manage telecommunications networks, services, processes and applications, understood as systems for capturing, transporting, representing, processing, storing, managing and presenting multimedia information, from the point of view of transmission systems .

en_CST6 - Ability to analyze, encode, process and transmit multimedia information using analog and digital signal processing techniques.

Learning Outcomes

RA1. Know the characteristics of the voice signal and its modelling, and the techniques of signal processing used for coding, synthesis and filtering of voice and audio.

RA2. Know how to implement specific digital processing tools in the field of voice and audio processing.

RA3. Know the main standards of voice and audio coding.

RA4. Know how to implement voice and audio coding systems, as well as audio synthesis systems and spatial filtering.

RA5. Know how to solve problems with initiative, working both individually and in work teams. Know how to present results effectively, both orally and in written reports.

3. CONTENTS

Contents Blocks	Total number of hours
1. The voice signal Phonology and phonetics. Characteristics of the voice: pitch and formants. Voice production models.	4 hours
2. Auditory perception The sense of hearing. Perception of sound volume. Critical bands Perception of the pitch. Temporal and auditory masking.	3 hours
3. Speech and audio signal processing techniques Linear prediction. Localized analysis (localized energy,l ocalized autocorrelation, localized Fourier transform). Homomorphic analysis. Adaptive filtering.	10 hours
4. Voice and audio coding. Sampling and uniform and non-uniform quantification. Vector quantification. Coding without losses. Open-loop and closed-loop encoders: predictive coding, analysis by synthesis encoders. Encoders in the frequency domain: subband and by transforms.	12 hours
5. Improvement of voice quality. Measurement of voice quality. Pathologies of the auditory system. Acoustic losses Hearing aid systems: digital hearing aids.	10 hours
6. Spatial hearing. Spatial perception of sound. Head-torsorelated response transfer function(HRTF). Reverberation and distance. Acoustic modeling of reverberation. Mirror source technique.	10 hours
7. Microphone arrays Application of array processing techniques. Improved quality with arrays of microphones: spatial filtering.	7 hours

4. TEACHING - LEARNING METHODOLOGIES. FORMATIVE ACTIVITIES.

4.1. Credits Distribution

Number of on-site hours:	58 hours (56 hours on-site +2 exams hours)
Number of hours of student work:	92
Total hours	150

4.2. Methodological strategies, teaching materials and resources

The teaching strategy of the subject is developed in three differentiated learning strategies but whose interrelation will allow to address the objectives set by the competences described above, namely: theoretical learning in the classroom, learning in work sessions in the laboratory and self-learning through the elaboration of essays.

Theoretical learning in the classroom:

The work sessions in the classroom, in large groups, will consist of master classes, where the main concepts of the subject under study will be presented. The objective is to introduce the student to the theoretical foundations of the subject in a guided, sequential and reflective way. The assimilation of these concepts will culminate with the implementation of them in the laboratory groups. The support with teaching materials will be fundamental to create reflective learning environments, where students and teachers can undertake a critical analysis that allows the student to relate concepts autonomously.

The order of presentation of the contents will evolve from the simplest to the most complex, with the aim of avoiding a high degree of abstraction that could cause the student lack of interest in the subject. In any case, it is very convenient during the work sessions in the classroom, to establish links with other subjects of the curriculum, and to provide possible experiences on the contents, which will help to attract the attention of the student and encourage their interest in the subject.

Learning in work sessions in the laboratory:

The practices in the laboratory make up another of the learning scenarios. The work sessions will be carried out in small groups, in which the student must work as a team. The objective is that the student deepens on the theoretical knowledge of the subject and explore, with the help of a manual of practices designed for the subject, the applicability of said knowledge.

Within the laboratory, the students will carry out the practices in groups of 2 students and at the end they will give the teacher a report that includes the work and the conclusions obtained with the realization of the practice.

Self-learning through the completion of work:

The last stage of learning is set by the realization of works related to the techniques and concepts developed in the classroom. The objective is for the student to develops kills related to the search for information, the management of bibliography and the making of reports on the topics proposed to them. In the same way, the aim is to encourage teamwork-groups of 2 students, coinciding with the work teams in the laboratory practices.

5. ASSESSMENT: procedures, evaluation and grading criteria

Preferably, students will be offered a continuous assessment model that has characteristics of formative assessment in a way that serves as feedback in the teaching-learning process.

5.1. PROCEDURES

The evaluation must be inspired by the criteria of continuous evaluation (Regulations for the Regulation of Teaching Learning Processes, NRPEA, art 3). However, in compliance with the regulations of the University of Alcalá, an alternative process of final evaluation is made available to the student in accordance with the Regulations for the Evaluation of Apprenticeships (approved by the Governing Council on March 24, 2011 and modified in the Board of Directors). Government of May 5, 2016) as indicated in Article 10, students will have a period of fifteen days from the start of the course to request in writing to the Director of the Polytechnic School their intention to take the non-continuous evaluation model adducing the reasons that they deem convenient. The evaluation of the learning process of all students who do not apply for it or are denied it will be done, by default, according to the continuous assessment model. The student has two calls to pass the subject, one ordinary and one extraordinary.

Ordinary Call

Continuous Assessment:

The content and timing of the Continuous Evaluation will be detailed at the beginning of each course in the Work Plan of the subject, which will include:

1. Work in small groups (E1). Each group will be responsible for the realization and presentation of works related to the contents of the subject. In these works, the student will deepen in topics explained in class, or that suppose tendencies to explore within the scope of the processing of voice and audio.The subjects on which they will treat the works will be presented in class, but the student will also be able to propose a topic of their interest, that will have to be submitted to the academic criterion of the teaching staff.
2. Performance of laboratory practices and delivery of the corresponding reports(E2). The evaluation will consider the systematic observation, where the teacher will record the main difficulties and skills observed in each student, through individual interviews carried out periodically, and the realization of a single memory by practice, by each of the groups of students who have done it. Students must attend the laboratory sessions on a regular basis and submit the corresponding reports to all laboratory practices. Recovery sessions will be enabled for those students who have not attended any of the sessions and justify it documentally. The students, as a group, will deliver the reports of the laboratory practices following the calendar established in the Work Plan of the subject. These practices will be evaluated by the professor responsible for the laboratory group, to check if the objectives indicated in the script of the same have been fulfilled.

3. Theoretical-practical and laboratory written test (PEF-T+L). There will be a final exam for the evaluation of the knowledge acquired in the subject.

In the process of continuous evaluation, the attendance to the theoretical and practical classes is mandatory, having to justify the absences.

Evaluation through final exam:

In the case of evaluation by means of a final exam, the evaluation elements to be used will be the following:

- Practical laboratory test (PEF-L).
- Theoretical-practical written test (PEF-T).

It is recommended that students perform laboratory practices during the semester, thus replacing the practical laboratory test by the evaluation of the reports corresponding to the different practices.

Extraordinary Call

The procedure will be the same as that described for the evaluation by means of a final exam in the ordinary call.

5.2. EVALUATION

EVALUATION CRITERIA

The procedures described above for the ordinary and extraordinary calls and the two evaluation methods have the objective of evaluating if the student has acquired the knowledge, procedures and skills professions that are listed below:

CE1. The student knows the properties of voice and audio signals.

CE2. The student knows the voice production model, its utility and how to calculate its parameters.

CE3. The student knows the main techniques of voice and audio coding.

CE4. The student knows the fundamentals of voice and audio synthesis.

CE5. The student knows the problem of spatial filtering and arrays of microphones, the main techniques used, and how to implement them.

CE6. The student is able to design and implement voice and audio coding systems.

CE7. The student is able to design and perform speech synthesis systems.

CE8. The student is able to design and perform spatial filtering systems and arrays of microphones.

CE9. The student knows the regulations and basic legislation in this discipline.

CE10. The student is able to work in a group, in multilingual environments, and to communicate orally and in writing, ideas, knowledge, conclusions.

Rating Instruments.

This section specifies the assessment instruments that will be applied to each of the evaluation criteria.

Ordinary call. continuous assessment:

-Internship (E2) and delivery of reports on the following topics:

- Voice coding systems
- Spatial filtering systems
- Audio processing systems

-Works of the subject (E1): Works of a theoretical-practical nature will be carried out,where the specific aspects of the audio and voice systems described in the subject will be deepened.

-Theoretical-practical and laboratory written test (PEF-T+L)

GRADING CRITERIA

Ordinary call: continuous evaluation

In the ordinary call-continuous assessment the relationship between the criteria, instruments and qualification is as follows.

Competences	Learning result	Assessment criteria	Assessment instrument	Grade weighting
CST1, CST6, TR1, TR3	RA1-RA5	CE1-CE11	E2	50%
CST1, CST6, TR1, TR3, TR8	RA1-RA5	CE1-CE11	E1	20%
CST1, CST6, TR1, TR3	RA1-RA5	CE1-CE10	PEF-T+L	30%

Taking this into account, in the case of continuous evaluation, the qualification will be made taking into account the following percentages and recitals:

- Work in small groups: 20%
- Laboratory practices: 50%.
  - Individual interviews during the development of the practices:20%
  - Practice reports (average score of all of them): 30%
- Final evaluation test: 30%

The qualification of "Not Submitted" will be awarded to the student who, having opted for the continuous assessment procedure, meets any of the following requirements:

- When the student has broken the attendance at least 60% of the classes in small groups.oWhen the student has not delivered, at least 60% of the jobs requested.When the student has exceeded the limits of attendance or delivery of work mentioned in the previous paragraph, regardless of their participation in the final exam, they will not be eligible for the grade of "Not Submitted".

Ordinary call: evaluation through final exam

In the ordinary call-final evaluation, the relationship between the criteria, instruments and qualification is as follows.

Competences	Learning result	Assessment criteria	Assessment instrument	Grade weighting
CST1, CST6, TR1, TR3, TR8	RA1-RA5	CE1-CE11	PEF-L	40%
CST1, CST6, TR1, TR3	RA1-RA5	CE1-CE10	PEF-T	60%

In the case of evaluation through a final exam, the grade will be taken taking into account the following percentages and recitals:

- Practical laboratory test: 40%.The exam will consist in the realization of a practical test in the laboratory, where the acquisition of the practical competences of the subject is demonstrated.
- If the students had made the laboratory practices and delivered the corresponding reports, 30% of the final grade will be the corresponding to the average grade of the reports and 10% to the skills in the development of the same.
- Theoretical-practical written exam: 60%

Extraordinary call

In the case of extraordinary calls, the same percentages that have been established in the case of the evaluation will be maintained by means of a final exam in the ordinary call, giving the option of carrying out the practical laboratory test or maintaining the grade obtained in laboratory practices. (continuous evaluation) or in the practical exam (final evaluation of the ordinary call), according to the student's decision.

Competences	Learning result	Assessment criteria	Assessment instrument	Grade weighting
CST1, CST6, TR1, TR3, TR8	RA1-RA5	CE1-CE11	PEF-L	40%
CST1, CST6, TR1, TR3	RA1-RA5	CE1-CE10	PEF-T	60%

6. BIBLIOGRAPHY

6.1. Basic Bibliography

- DigitalProcessing of Speech Signals, L. R. Rabiner and R. W. Schafer,Prentice-Hall Inc., 1978.
- Speech Communication, Human and Machine, D. O'Shaughnessy, Addison-Wesley, 1987.
- Discrete-Time Processing of Speech Signals, J. Deller, Jr., J. H. L. Hansen, and J. G.Proakis, Wiley-IEEE Press, Classic Reissue, 1999.
- Springer Handbook of Speech Processing and Speech Communication, J. Benesty, M. M. Sondhi and Y Huang (eds.), Springer, 2008.
- Theory and Application of Digital Speech Processing, L. R. Rabiner and R. W.Schafer, Prentice Hall Inc., 2009.
- Audio Signal Processing and Coding, A. Spanias, T. Painter and V. Atti, JohnWiley and Sons, 2007.

6.2. Additional Bibliography

- Digital Coding of Waveforms, N. S. Jayant and P. Noll, Prentice Hall Inc.,1984.
- Vector Quantization and Signal Compression, A. Gersho and R. M. Gray, Kluwer Academic Publishers, 1992.
- A Practical Handbook of Speech Coders, R. Goldberg and L. Riek, CRCPress, 2000.

Disclosure Note

During the evaluation tests, the guidelines set out in the Regulations establishing the Rules of Coexistence of the University of Alcalá must be followed, as well as the possible implications of the irregularities committed during said tests, including the consequences for committing academic fraud according to the Regulation of Disciplinary Regime of the Students of the University of Alcalá.