ELEC9723
Speech Processing

Course Outline – Semester 1, 2015

Never Stand Still | Faculty of Engineering | School of Electrical Engineering and Telecommunications

Course Staff

Course Convener/Lecturer: Dr. Vidhyasaharan Sethu, Room 304, v.sethu@unsw.edu.au

Lecturer: Nicholas Cummins, Room 301, n.p.cummins@unsw.edu.au
Laboratory Demonstrator: Nicholas Cummins, Room 301, n.p.cummins@unsw.edu.au

Consultations: There are no fixed consultation times, you may make an appointment by email for any time. You may also try your luck and just knock on my door.

Keeping Informed: Announcements may be made during classes, via email (to your student email address) and/or via online learning and teaching platforms – in this course, we will use Moodle https://moodle.telt.unsw.edu.au/login/index.php. Please note that you will be deemed to have received this information, so you should take careful note of all announcements.

Course Summary

Contact Hours

The course consists of 3 hours of per week, comprising lectures and/or laboratory (a typical class might be 2 hours of lecture followed by 1 hour of lab):

Lectures: Thursdays, 6pm–9pm, room EE214
Lab sessions: Thursdays, 6pm-9pm, room EE214
Laboratory classes start in week 1 (Introductory MATLAB)

Context and Aims

ELEC9723 Speech Processing builds directly on students' skills and knowledge in digital signal processing gained during ELEC3104 Signal Processing and ELEC4621 Advanced Digital Signal Processing. Speech processing has been one of the main application areas of digital signal processing for several decades now, and as new technologies like voice over IP, automated call centres, voice browsing and biometrics find commercial markets, speech seems set to drive a range of new digital signal processing techniques for some time to come. This course provides not only the technical details of ubiquitous techniques like linear predictive coding, Mel frequency cepstral coefficients, Gaussian mixture models and hidden Markov models, but the rationale behind their application to speech and an understanding of speech as a signal. Contemporary signal processing is almost entirely digital, hence only discrete-time theory is presented in this course.

Aims: This course aims to:

a. Familiarise you with modeling the vocal tract as a digital, linear time-invariant system.
b. Convey details of a range of commonly used speech feature extraction techniques.
c. Provide a basic understanding of multidimensional techniques for speech representation and classification methods.
d. Familiarise you with the practical aspects of speech processing, including robustness, and applications of speech processing, including speech enhancement, speaker recognition and speech recognition.
e. Give you practical experience with the implementation of several components of speech processing systems.

Indicative Lecture Schedule

Week	Lecture	Lecturer	Laboratory
1	Introduction to speech processing Time-frequency analysis-1	Dr. V Sethu	**Introductory MATLAB**
2	Time-frequency analysis-2	Dr. V Sethu	Lab 1: Spectral analysis
3	Speech Modeling	N. Cummins	Lab 2: Time-Frequency Analysis
4	Linear Predictive Analysis	N. Cummins	Lab 3: Pitch Extraction
5	Human Auditory System	Dr. V Sethu	Lab 4: Formant Tracking
6	Speech Enhancement	Dr. V Sethu	**Checkpoint (for labs 1-4)**
7	**Mid-session examination, (23rd Apr) Duration 1 hour 15 min**	Dr. V Sethu	Lab 5:Speech Enhancement
8	Clustering and Gaussian Mixture models	Dr. V Sethu	No Lab
9	Front-end processing	N. Cummins	No Lab
10	Speaker Recognition	Dr. V Sethu	Lab 6: Front-End
11	Hidden Markov models	Dr. V Sethu	Lab 7: Speaker recognition
12	Speech recognition	Dr. V Sethu	Lab 8: Speech recognition
13	Project Demonstration	Dr. V Sethu	**Checkpoint (for labs 6, 7 & 8)**

Assessment

Laboratory work	30%
Project	10%
Mid-Semester Exam	10%
Final Exam (3 hours)	50%

Course Details

Credits

This is a 6 UoC course and the expected workload is 10–12 hours per week throughout the 13 week semester.

Relationship to Other Courses

ELEC9723 Speech Processing is a postgraduate course in the School of Electrical Engineering and Telecommunications and is the most advanced course offered by the university on this topic, and serves as an excellent basis from which to commence research in the area. Various aspects of the course bring students up to date with the very latest

developments in the field, as seen in recent international conferences and journals, and some of the laboratory work is designed in the style of an empirical research investigation. ELEC9723 is well complemented by ELEC9722 Digital Image Processing, which gives an insight into two-dimensional signal processing and image signals. ELEC9721 Digital Signal Processing Theory and Applications provides an excellent basis for Speech Processing, however for students who have not already completed this course (or ELEC4621), it is recommended for future study.

Pre-requisites

The minimum pre-requisite for the course is ELEC3104, Signal Processing (or equivalent). Knowledge from either ELEC4621 or ELEC9721 is highly desirable.

Assumed Knowledge

It is essential that you are familiar with the sampling theorem, digital filter design, the discrete Fourier transform, basic probability theory, random processes and autocorrelation and frame-by-frame processing. Students who are not confident in their knowledge from previous signal processing courses (especially the topics mentioned) are strongly advised to revise their previous course materials as quickly as possible to avoid difficulties in this course.

Learning outcomes

After successful completion of this course, you should be able to:

1. Express the speech signal in terms of its time domain and frequency domain representations and the different ways in which it can be modelled;
2. Derive expressions for simple features used in speech classification applications;
3. Explain the operation of example algorithms covered in lectures, and discuss the effects of varying parameter values within these;
4. Synthesise block diagrams for speech applications, explain the purpose of the various blocks, and describe in detail algorithms that could be used to implement them;
5. Implement components of speech processing systems, including speech recognition and speaker recognition, in MATLAB.
6. Deduce the behaviour of previously unseen speech processing systems and hypothesise about their merits.

This course is designed to provide the above learning outcomes which arise from targeted graduate capabilities listed in ***Appendix A***. The targeted graduate capabilities broadly support the UNSW and Faculty of Engineering graduate capabilities (listed in ***Appendix B***). This course also addresses the Engineers Australia (National Accreditation Body) Stage I competency standard as outlined in ***Appendix C***.

Syllabus

Fundamentals of speech production; speech analysis: pitch and period extraction, formant estimation, voiced - unvoiced decision, Linear prediction, Inverse filtering; Auditory modelling, auditory masking; Speech enhancement; Clustering, Gaussian mixture modelling, hidden Markov modelling, Implementation of speech and speaker recognition systems.

Teaching Strategies

Delivery Mode

The course consists of the following elements: lectures, laboratory work, and homework comprising of self-guided study and a project.

The teaching in this course aims at establishing a good fundamental understanding of the areas covered using:

- Formal face-to-face lectures, which provide you with a focus on the core analytical material in the course, together with qualitative, alternative explanations to aid your understanding;
- Laboratory sessions, which support the formal lecture material and also provide you with practical construction, measurement and debugging skills;

Learning in this course

You are expected to attend all lectures, tutorials, labs, and mid-semester exams in order to maximise learning. You must prepare well for your laboratory classes and your lab work will be assessed. In addition to attending lectures, you should read relevant sections of the recommended text. Reading additional texts will further enhance your learning experience. Group learning is also encouraged. UNSW *assumes* that self-directed study of this kind is undertaken in addition to attending face-to-face classes throughout the course.

Lecture/Study Notes

You are encouraged to make notes during lectures, laboratory sessions and when you are undertaking self-directed study. The course aims to familiarise you with the fundamental principles and concepts underlying speech processing and NOT make you memorise 'facts', consequently you will be allowed to refer to your handwritten notes in all assessments. **You may bring one notebook with handwritten notes to your mid-semester exam, lab assessments and your final exam.** Please note that loose sheets of paper and printed or electronic material is NOT allowed.

Lectures

During the lectures, techniques for the analysis, modeling and processing of the digital speech signal will be presented. The lectures provide you with a focus on the core material in the course, together with qualitative, alternative explanations to aid your understanding. Various examples will be given, to enrich the analytical course content. The lectures will serve as a good guide to the course syllabus, but you will need to supplement them with additional reading, of the recommended text book and/or other materials recommended by the lecturing staff. In particular, you should not assume that attendance at all lectures (even with a glance or two through the notes), on its own, is sufficient to pass the course.

Laboratory program

The lecture and laboratory schedule is designed to give you practical, hands-on exposure to the concepts conveyed in lectures soon after they are conveyed in class. Generally there will be around one week between the introduction of a topic in lectures and a laboratory exercise on the same topic, sufficient time in which to revise the lecture, attempt related problems and prepare for the laboratory. The laboratory work provides you with hands-on design experience and exposure to simulation tools and algorithms used widely in speech processing. You must be pre-prepared for the laboratory sessions: the laboratory sessions are short, so this is only possible way to complete the given tasks.

Laboratory classes will start in week 1 of session, with the compulsory Introductory MATLAB laboratory. Regular laboratory classes will start in week 2*.* You will need to bring to the laboratories:

- A USB drive for storing MATLAB script files
- Your lecture notes, laboratory preparation and/or any other relevant course materials

Laboratory Exemption

There is no laboratory exemption for this course. Regardless of whether equivalent labs have been completed in previous courses, all students enrolled in this course for Semester 1, 2015 must take the labs. If, for medical reasons, (note that a valid medical certificate must be provided) you are unable to attend a lab, you will need to apply for a catch-up lab during another lab time, as agreed by the laboratory coordinator.

Homework and Problem sheets

The lectures can only cover the course material to a certain depth; you must read the textbook(s) and reflect on its content as preparation for the lectures to fully appreciate the course material. Home preparation for laboratory work provides you with the background knowledge you will need. The problem sheets aim to provide in-depth quantitative and qualitative understanding of speech processing theory and methods. Together with your attendance at classes, your self-directed reading, completion of problems from the problem sheet and reflection on course materials will form the basis of your understanding of this course.

Assessment

The assessment scheme in this course reflects the intention to assess your learning progress through the semester. Ongoing assessment occurs through the lab checkpoints and the mid-semester exam.

Laboratory Assessment (30%)

The laboratory work will be assessed in the weeks marked 'Checkpoint' in the course schedule in order to ensure that you are studying and that you understand the course material. The laboratory assessment is conducted live during the lab sessions, so it is essential that you arrive at each lab having revised lecture materials in advance of each laboratory, and having completed any requested preparation for the labs. Without preparation, marks above 50% may be difficult to obtain. No lab reports are required in this course. You should bring all the code you've written in previous laboratory sessions to the checkpoints.

Note that *laboratory assessment will be conducted individually, not on a per-group basis*. Please also note that *you must pass the laboratory component in order to pass the course*.

Project (10%)

A project will be given in week 7 which you are expected to complete by week 13. A project report will have to be submitted by week 13 and the project code demonstrated (also in week 13) as outlined in the course schedule. The project work is to be completed individually and NOT in groups. The project will involve some topics covered in lectures.

Mid-Semester Exam (10%)

The mid-session examination tests your general understanding of the course material, and is designed to give you feedback on your progress through the analytical components of the course. Questions may be drawn from any course material up to the end of week 6. It may contain questions requiring some (not extensive) knowledge of laboratory material, and will definitely contain numerical and analytical questions. Marks will be assigned according to the correctness of the responses. The examination will be held in week 7, on the 23rd of April, 2015.

Final Exam

The exam in this course is a 3 hour written examination. ***It is a closed-book examination except for one notebook of handwritten course notes which you may bring to the exam for your reference***. University approved calculators are allowed. The examination tests analytical and critical thinking and general understanding of the course material in a controlled fashion. Questions may be drawn from any aspect of the course (including laboratory), unless specifically indicated otherwise by the lecturer. Marks will be assigned according to the correctness of the responses. *Please note that you must pass the final exam in order to pass the course.*

Relationship of Assessment Methods to Learning Outcomes

	Learning outcomes					
Assessment	**1**	**2**	**3**	**4**	**5**	**6**
Laboratory checkpoints	✓	-	✓	-	✓	✓
Mid-semester exam	✓	-	✓	✓	-	✓
Project	✓	-	✓	✓	✓	✓
Final exam	✓	✓	✓	✓	-	✓

Course Resources

Textbooks

The following textbook is prescribed for the course:

[1] Quatieri, T. F. (2002). *Discrete-Time Speech Signal Processing*, Prentice-Hall, New Jersey.

You may want to check the coverage of this text before purchasing, as some topics in the syllabus are not featured. Unfortunately there is no single text that covers all topics in a satisfactory depth. Additional references, listed below and at the end of some lecture slides, will in combination provide complete coverage of the course. Lecture slides (if used) will be provided, however note that these do not treat each topic exhaustively and additional reading is required.

Reference books

The following books are good additional resources for speech processing topics:

[2] Rabiner, L. R., and Juang, B.-H. (1993). *Fundamentals of Speech Recognition*, Prentice-Hall, New Jersey.

Books covering assumed signal processing knowledge

The following books cover material which is assumed knowledge for the course:

[3] Mitra, S. K. (2010). Digital Signal Processing: A Computer-Based Approach, McGraw-Hill.

On-line resources

Moodle

As a part of the teaching component, Moodle will be used to disseminate teaching materials, host forums and occasionally quizzes. Assessment marks will also be made available via Moodle: https://moodle.telt.unsw.edu.au/login/index.php.

Mailing list

Announcements concerning course information will be given in the lectures and/or on Moodle and/or via email (which will be sent to your student email address).

VOICEBOX: Speech Processing Toolbox for MATLAB
`http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html`

Other Matters

Academic Honesty and Plagiarism

Plagiarism is the unacknowledged use of other people's work, including the copying of assignment works and laboratory results from other students. Plagiarism is considered a form of academic misconduct, and the University has very strict rules that include some severe penalties. For UNSW policies, penalties and information to help you avoid plagiarism, see http://www.lc.unsw.edu.au/plagiarism. To find out if you understand plagiarism correctly, try this short quiz: https://student.unsw.edu.au/plagiarism-quiz.

Student Responsibilities and Conduct

Students are expected to be familiar with and adhere to all UNSW policies (see https://my.unsw.edu.au/student/atoz/ABC.html), and particular attention is drawn to the following:

Workload

It is expected that you will spend at least **ten to twelve hours per week** studying a 6 UoC course, from Week 1 until the final assessment, including both face-to-face classes and *independent, self-directed study*. In periods where you need to need to complete assignments or prepare for examinations, the workload may be greater. Over-commitment has been a common source of failure for many students. You should take the required workload into account when planning how to balance study with employment and other activities.

Attendance

Regular and punctual attendance at all classes is expected. UNSW regulations state that if students attend less than 80% of scheduled classes they may be refused final assessment.

General Conduct and Behaviour

Consideration and respect for the needs of your fellow students and teaching staff is an expectation. Conduct which unduly disrupts or interferes with a class is not acceptable and students may be asked to leave the class.

Work Health and Safety

UNSW policy requires each person to work safely and responsibly, in order to avoid personal injury and to protect the safety of others.

Special Consideration and Supplementary Examinations

You must submit all assignments and attend all examinations scheduled for your course. You should seek assistance early if you suffer illness or misadventure which affects your course progress. All applications for special consideration must be **lodged online through myUNSW within 3 working days of the assessment**, not to course or school staff. For more detail, consult https://my.unsw.edu.au/student/atoz/SpecialConsideration.html.

Continual Course Improvement

This course is under constant revision in order to improve the learning outcomes for all students. Please forward any feedback (positive or negative) on the course to the course convener or via the Course and Teaching Evaluation and Improvement Process. You can also provide feedback to ELSOC who will raise your concerns at student focus group

meetings. As a result of previous feedback obtained for this course and in our efforts to provide a rich and meaningful learning experience, we have continued to evaluate and modify our delivery and assessment methods.

Administrative Matters

On issues and procedures regarding such matters as special needs, equity and diversity, occupational health and safety, enrolment, rights, and general expectations of students, please refer to the School and UNSW policies:
http://www.engineering.unsw.edu.au/electrical-engineering/policies-and-procedures
https://my.unsw.edu.au/student/atoz/ABC.html

Appendix A: Targeted Graduate Capabilities

Electrical Engineering and Telecommunications programs are designed to address the following targeted capabilities which were developed by the school in conjunction with the requirements of professional and industry bodies:

- The ability to apply knowledge of basic science and fundamental technologies;
- The skills to communicate effectively, not only with engineers but also with the wider community;
- The capability to undertake challenging analysis and design problems and find optimal solutions;
- Expertise in decomposing a problem into its constituent parts, and in defining the scope of each part;
- A working knowledge of how to locate required information and use information resources to their maximum advantage;
- Proficiency in developing and implementing project plans, investigating alternative solutions, and critically evaluating differing strategies;
- An understanding of the social, cultural and global responsibilities of the professional engineer;
- The ability to work effectively as an individual or in a team;
- An understanding of professional and ethical responsibilities;
- The ability to engage in lifelong independent and reflective learning.

Appendix B: UNSW Graduate Capabilities

The course delivery methods and course content directly or indirectly addresses a number of core UNSW graduate capabilities, as follows:

- Developing scholars who have a deep understanding of their discipline, through lectures and self-directed study and assessed by laboratory checkpoints and written examinations.
- Developing rigorous analysis, critique, and reflection, and ability to apply knowledge and skills to solving problems. These will be achieved by the laboratory experiments and interactive checkpoint assessments during the labs.

- Developing capable independent and collaborative enquiry, through tutorial exercises together with self-directed study.
- Developing digital and information literacy and lifelong learning skills through project work and self-directed preparation for lectures and labs.

Appendix C: Engineers Australia (EA) Professional Engineer Competency Standard

	Program Intended Learning Outcomes	
PE1: Knowledge and Skill Base	PE1.1 Comprehensive, theory-based understanding of underpinning fundamentals	✓
	PE1.2 Conceptual understanding of underpinning maths, analysis, statistics, computing	✓
	PE1.3 In-depth understanding of specialist bodies of knowledge	✓
	PE1.4 Discernment of knowledge development and research directions	
	PE1.5 Knowledge of engineering design practice	
	PE1.6 Understanding of scope, principles, norms, accountabilities of sustainable engineering practice	
PE2: Engineering Application Ability	PE2.1 Application of established engineering methods to complex problem solving	✓
	PE2.2 Fluent application of engineering techniques, tools and resources	✓
	PE2.3 Application of systematic engineering synthesis and design processes	
	PE2.4 Application of systematic approaches to the conduct and management of engineering projects	
PE3: Professional and Personal Attributes	PE3.1 Ethical conduct and professional accountability	
	PE3.2 Effective oral and written communication (professional and lay domains)	✓
	PE3.3 Creative, innovative and pro-active demeanour	✓
	PE3.4 Professional use and management of information	✓
	PE3.5 Orderly management of self, and professional conduct	
	PE3.6 Effective team membership and team leadership	