



KTH Computer Science  
and Communication

# DT2118: Speech and Speaker Recognition

## Course Analysis VT2012

### 1 Course data

Course name	Speech and Speaker Recognition
Course number	DT2118
Credits	7.5
When the course took place	Period 4, VT 2012
Teachers (hours)	Giampiero Salvi (course responsible and lecturer, 8F, 2T, 4Ö, 8L, 4S) <sup>1</sup> Mats Blomberg (lecturer, 8F)
Number of registered students	10
Number of students attending	6
Number of students completing	6

### 2 Course objectives

The course objective is to give students insights in the signal processing and statistical methods used for speech and speaker recognition. After attending the course, the students should be able to:

1. discuss concepts related to speech signal processing and speech recognition
2. explain and implement simple parameter estimation methods for Gaussian models
3. explain and implement the main algorithms related to hidden Markov model training and decoding
4. use the software package KTH to build and evaluate a simple speech recogniser
5. carry out a small project related to speech or speaker recognition

### 3 How the course is designed to reach the objectives

The course gives a central space to the role of student activities as a means to learning. The eight lectures are meant to give the context and sufficient insights to the students in order to perform the different kinds of activities. Among these are: a set of computational exercises, a laboratory exercise and a final project.

The computational exercises require a mixture of theoretical and programming skills. The students are not only required to solve the exercises theoretically and numerically, but also to

<sup>1</sup>F: lectures (föreläsningar), T: tutorial, Ö: exercises (övningar), L: lab, S: final seminar

hand in the code they have written for the solution. Matlab is the preferred language, but the students are left free to choose their favourite programming language. A central part of this activity has been the discussion with the teacher after the solution has been handed in. Given the relatively low number of students, the course responsible was able to sit with each student and check their solutions in details, searching for eventual bugs in their code.

The laboratory is based on a software package for speech recognition (HTK) and on recordings made by the students in a previous lab (from DT2112). Students work in groups of about 2. Students that did not participate in DT2112 are required to go through the DT2112 lab before they start the DT2118 lab. The aim of this exercise is to compare the effects of different feature extraction methods on the recognition accuracy in a digit recognition task. As an optional task, the students are asked to perform Vocal Tract Length Normalisation and to discuss the results they obtain.

The final activity is the final project. I tried to keep the subject of the projects as open as possible for two reasons. The first is the choosing the right project on the base of the information given by the lectures and one's own skills is in my opinion a learning activity *per se*. The other reason is that this kind of project work is best performed when the students are highly motivated, and being able to choose the task is a strong prerequisite to motivation. Of course, the students need continuous input from the teachers in order to realise if their ideas are feasible with the limited time given by the course. The students are required to work in groups of about 2 students (but also single student projects have been allowed this year). The need to perform the experiments related to the task they have chosen and write a report. Each student is also supposed to review the work of another group. In the final meeting in the course, each group is required to present the project, and the reviewers are required to ask related questions. Also the rest of the students are stimulated to participate in the discussion. In order to help the reviewers prepare for the discussion, each groups has been asked to hand in a draft of the report already one week before the final meeting.

## 4 Course pedagogical development I

For this year implementation of the course, a new lecture has been added to introduce the students to advanced topics.

The computational exercises have been updated with new theoretical questions, and an exercise package has been developed including a number of Matlab prototype function to guide the implementation of the functions that are necessary to solve the exercises.

The laboratory instructions have also been updated.

## 5 Contact with the students during the course

### 5.1 Students in this year's committee

### 5.2 Results of the course meeting

There was no formal meeting during the course, but there was always close discussion with the students, also simplified by the low number of students.

## 6 Contact with the other teachers during the course

The communication with the other teachers was always very smooth, being us all colleagues from the same research group.

## 7 Questionnaire, the student's point of view

### 7.1 Period in which the questionnaire was active

After the last meeting of the course, until the final meeting with the course responsible (about two weeks)

### 7.2 Questions that were added to the standard

See the questionnaire

### 7.3 Statistics of answers

67% of the students that completed all the course requirements, and 40% of the total.

### 7.4 Changes compared to the last implementation

No modifications were made this year.

### 7.5 General impression

The general impression on the course is positive. Students appreciate especially the computational exercises that give insights in the details of the algorithms, and the fact that the teachers were always available for discussion (this is of course dependent on the number of students that enrolled this year).

### 7.6 Relevant web links

## 8 Interpretation of the questionnaire by the course responsible

### 8.1 Positive views

- All students but one had a clear idea of the course objectives from the start.
- All students this year are happy about the balance between standard lectures and practical activities.
- Students are very happy about the computational exercises. Also the choice of the course responsible to sit with them individually and check for mistakes and bugs was very much appreciated.
- Students liked in general the availability and engagement of the teachers
- Students are happy about the organisation of the course

## 8.2 Negative views

- Some students found difficulties using a unix-like environment in the lab.
- Half of the students state that their work load was mainly concentrated at the end of the course.
- Some students would prefer grades based on a 7 scale

## 8.3 Was the course relevant with respect to the objectives

Yes

## 8.4 View on prerequisites

They are appropriate, some students found the laboratories based on linux hard to complete without help from other students.

## 8.5 View on forms of teaching

It seems that the lecturers were considered positively from a pedagogical perspective

## 8.6 View on literature and course material

The book seems to be better appreciated this year than last year

## 8.7 Opinion on exam

Most students are positive about the project presentations

## 8.8 Especially interesting comments

I had considered to remove the possibility for the students to choose a literature study instead of the practical project, but this year they seemed to have enjoyed the activity: "Very interesting work! First time for me that i write a paper about other papers, and it was quite interesting to read a lot about a precise topic!". So I am reconsidering now.

# 9 Opinions from the other teachers after the end of the course

## 9.1 What worked well

Positive interest and response from the students

## 9.2 What worked less well

- The lecture schedule did not fit all students.
- One (or a few?) lecture(s) had to be given in a quite small lecture room

## 9.3 Suggestions for improvements

- Graded examination
- Practical exercises on other components than acoustic modeling

## **10 Results of the course commission meeting after the exam**

This year we did not perform a formal meeting, but I was in contact with the students during the whole course and collected their feedback

### **10.1 Summary from the students**

### **10.2 Suggestions for changes**

### **10.3 Link to the meeting protocol**

## **11 Summary of the course responsible report**

### **11.1 General impression**

The general impression of the course is positive.

### **11.2 Positive points of view**

There seems to be a good balance between lectures and practical assignments. The availability of the teachers seems to be highly appreciated by the students. The organisation of the different steps in the course seems to have worked well. The students were very motivated, they attended the lectures regularly (although it was not mandatory) and participated actively in the discussions. They respected all the deadlines, sometimes anticipating them.

### **11.3 Negative points of view**

The lectures are a bit too much concentrated at the beginning of the course. Although this was done on purpose, it seems to make it harder for the students to elaborate the concepts that are discussed in the theoretical part of the course. In VT2013 we have moved some of the theoretical lectures forward in time.

### **11.4 Opinion on prerequisites**

The prerequisite are satisfactory. Some details will be adjusted.

### **11.5 Opinion on forms of teaching**

The forms of teaching seem to be satisfactory (see also previous points)

### **11.6 Opinion on literature/course material**

The literature seems to be satisfactory.

### **11.7 Opinion on examination**

The examination seems to have worked fine.

## 12 Course pedagogical development II

### 12.1 How the changes to this course work

The changes to the previous years seem to have worked well. Especially the decision to check the exercise solutions individually with the students.

### 12.2 Changes that should be done for the next time

- some of the lectures will be moved towards the end of the course.
- some ready made material for the project will be made available in order to simplify the work
- the course results might be graded on a seven grade scale
- the lab description will be revised and clarified
- the lab scripts will be modified to make sure they can be run on standard Linux computers available at CSC, and that the mechanical steps involved are simplified and reduced.
- the lectures by Mats Blomberg (that is going to get retired in Spring 2013) will be given by Giampiero Salvi.
- the course web pages will be moved to KTH Social, to simplify interaction with the students

# DT2118 Speech and Speaker Recognition

## Results of course evaluation

---

**24 questions to be answered and a few comments to be given.**

---

1. Do you think the course is easy or difficult?

1. 0% (0 st) Very easy.
  2. 25% (1 st) Easy.
  3. 75% (3 st) Average.
  4. 0% (0 st) Rather difficult.
  5. 0% (0 st) Very difficult.
- 

2. Did you get a clear idea of the course objective at the course start?

1. 75% (3 st) Yes.
  2. 0% (0 st) Hesitant.
  3. 25% (1 st) No.
- 

3. Do you think the course is interesting?

1. 50% (2 st) Yes, very.
  2. 50% (2 st) Yes.
  3. 0% (0 st) Neutral.
  4. 0% (0 st) Not very.
  5. 0% (0 st) No.
- 

4. How did you get to know about this course?

1. 25% (1 st) Friend.
  2. 75% (3 st) KTH web pages.
  3. 0% (0 st) From DT2112.
  4. 0% (0 st) Other (please specify).
- 

Comments:

*looking for courses about speech processing*

---

5. The requirement on previous courses is the Speech Technology course DT2112 or equivalent and experience with unix. Do you regard your level as sufficient at the time of the course start?

1. 75% (3 st) Yes.
2. 25% (1 st) Hesitant.
3. 0% (0 st) No.

Comments:

*No problem with the technical contents, but the HTK lab would have been much harder without my partner, regarding the Unix commands.*

---

6. What is your opinion on the course book "Spoken Language Processing"?

1. 0% (0 st) Very good.
2. 50% (2 st) Good.
3. 25% (1 st) Acceptable.
4. 0% (0 st) Not so good.
5. 0% (0 st) Bad.
6. 25% (1 st) Did not use it.

Comments:

*I used the book, especially for implementing the computational exercises.*

---

7. What is your opinion on the extra material (papers, handouts, etc)?

1. 0% (0 st) Very good.
2. 100% (4 st) Good.
3. 0% (0 st) Acceptable.
4. 0% (0 st) Not so good.
5. 0% (0 st) Bad.
6. 0% (0 st) Did not use it.

Comments:

8. Was there a good balance between lectures and practical activities?

1. 0% (0 st) Too many lectures.
2. 100% (4 st) Good balance.
3. 0% (0 st) Too many practical activities.

Comments:

*Good, but maybe put the lab earlier and the lectures later*

---

*However, some theoretical classes were concentrated over too short a time frame (the space of a couple of weeks). I understand that the academic year dates must be respected, however it was a bit heavy to follow 2 theoretical 1 practical class (HTK tutorial) in a couple of days.*



---

9. What do you think of the lectures in a pedagogical way? (Are the concepts well described? Do the teachers speak and write clearly)

1. 50% (2 st) Very good.
2. 25% (1 st) Good.
3. 25% (1 st) Acceptable.
4. 0% (0 st) Not so good.
5. 0% (0 st) Bad.

Comments:

---

10. How was your work load distributed during the course?

1. 0% (0 st) Mainly at the beginning.
2. 50% (2 st) Evenly during the course.
3. 50% (2 st) Mainly at the end.

Comments:

*Mainly before the computational exercises deadline, and for the project (which I still have not completed as of 1 June 2012).*

---

11. The assessment in this course is based on a Fail/Pass grading scale. Do you think we should assign more specific grades? For example on a 7 level scale?

1. 50% (2 st) Yes.
2. 25% (1 st) Hesitant.
3. 25% (1 st) No.

Comments:

*I think that computational exercises and the lab should stay P/F. But the term paper was really a lot of work and it would be nice to get a real grade! But not necessary.*

---

*I don't have a definite preference on grading scales (fail/pass versus mark), both have pros and cons. I'm fine with fail/pass: the grade is not the most important thing at the end of the day/*

---

12. What is your opinion on the computational exercises?

1. 75% (3 st) Very good.
2. 25% (1 st) Good.
3. 0% (0 st) Acceptable.
4. 0% (0 st) Not so good.
5. 0% (0 st) Bad.
6. 0% (0 st) No opinion.

Comments:

*Very good to understand some parts of the lectures.*

---

*A little bit more about the background of this kind of algorithms could be better*

---

*It was very useful to implement HMM algorithms from scratch. The detail of guideline in the exercise text was just adequate: some of them were quite difficult and time consuming... but still doable and most importantly clear.*

---

13. What is your opinion on the practical exercise (lab)?

1. 25% (1 st) Very good.
2. 25% (1 st) Good.
3. 25% (1 st) Acceptable.
4. 0% (0 st) Not so good.
5. 0% (0 st) Bad.
6. 25% (1 st) No opinion.

Comment:

*The topic was very interesting. But it was long because of the copying of the results, etc... Maybe something could be changed so that we think more and copy less from the screen.*

---

*Even though this course is held in a Swedish university, and it's only good if foreign students learn some Swedish language words like numbers... it would be good to have option of writing an English language recognizer.*

---

14. How long time did you spend preparing the practical exercise (lab)?

1. 0% (0 st) Less than 6 hours.
2. 75% (3 st) 6-12 hours.
3. 0% (0 st) 12-24 hours.
4. 0% (0 st) More than 24 hours.

Comment:

*Two days at the lab. It was a good amount of work.*

---

15. How long time did you spend on writing the term paper and preparing the presentation:

1. 0% (0 st) Less than one day.
2. 0% (0 st) 2-3 days.
3. 75% (3 st) 4-6 days.
4. 0% (0 st) 7 days or more.

Comments:

*Very interesting work! First time for me that i write a paper about other papers, and it was quite interesting to read a lot about a precise topic!*

---

16. How many other courses did you follow in parallel to this one (period 4)?

1. 25% (1 st) None.
  2. 0% (0 st) One.
  3. 75% (3 st) Two.
  4. 0% (0 st) Three.
  5. 0% (0 st) Four or more.
- 

17. How large proportion of your studying time in period 4 did you spend on this course?

1. 0% (0 st) Less than 15%.
  2. 50% (2 st) 15-30%.
  3. 50% (2 st) 30-50%.
  4. 0% (0 st) 50-70%.
  5. 0% (0 st) More than 70%.
- 

18. The course is 7.5 hp. How do you regard that in comparison with other courses?

1. 75% (3 st) 7.5 hp is appropriate.
  2. 0% (0 st) Should be more than 7.5 hp.
  3. 25% (1 st) Should be less than 7.5 hp.
- 

19. What is your opinion on the administration of the course and exchange of information between teachers and students?

1. 100% (4 st) Very good.
2. 0% (0 st) Good.
3. 0% (0 st) Acceptable.
4. 0% (0 st) Not so good.
5. 0% (0 st) Bad.

Comments:

*So nice to be so few students!*

---

*The teachers were always available and helpful, both in person and by email, with quick replies.*

---

20. Were the requirements at every stage of the course always clear to you?

1. 75% (3 st) Yes.

2. 25% (1 st) Hesitant.
3. 0% (0 st) No.

Comments:

---

21. Do you feel that you have been discriminated in this course due to gender, sexuality, ethnicity or disability?

1. 0% (0 st) Yes.
2. 0% (0 st) Hesitant.
3. 100% (4 st) No.

If yes, in which way?

---

22. How do you regard the course from a gender perspective (e.g. with respect to course book, teachers, etc.)?

*I don't understand this question. Is it about gender \*balance\* in the book's authors list and teachers??*

*Whatever the meaning was, I regard the course well from all aspects, save for a couple of suggestions regarding theoretical classes (see last question).*

---

23. Suggestion for course improvements:

*More lectures, especially about recent topics and trends (advanced topics). The extra lectures can be both theoretical or practical or intertwined - it doesn't really matter, as long as they give more practical information to students (algorithms, tricks, computational complexity and bottlenecks, showing software and code).*

*The lecture about HMM and the Viterbi and Baum-Welch algorithms is very dense: it could be expanded to last more than two hours.*

---

24. Further comments:

*I enjoyed the course for several reasons:*

- it ticked my old curiosity in language and speech processing*
  - teachers were patient and easy-going*
  - there is freedom in the choice and form of the final project*
- 

[giampi@kth.se](mailto:giampi@kth.se)

Denna sammanställning har genererats med [ACE](#).