project AVRiL

Advisors:
Dr. Sohaib Khan
Dr. Umar Saif

BSc Sproj '07 - Group 7:
Ahmad Humayun
Ozair Muazzam
Tayyab Javed
Yahya Cheema

avril.sproj.com

Note:
This is version 1.0 of the requirements specification.
Since we are following an iterative SE process this document is subject
to revision as we progress through the project.

Section 1: Introduction

Purpose of this section: General background and reference information

1.1 Purpose Of This Document

This SRS provides a description of the functions and constraints of project AVRiL, developed at the Lahore University of Management Sciences. The expected audiences of this document are our Senior Project TA's, Amjad Luna, and our advisors Dr. Sohaib Khan and Dr. Umar Saif. It will also help us, the group members, to maintain focus throughout the course of the project.

1.2 Scope Of The Development Project

AVRiL will aim to automatically record, using PTZ cameras, university lectures in order to assist distance learning programs. We are assuming that the system would be used inside auditoriums only, in which only one instructor is teaching at a time. We also assume that the instructor will be standing and in motion most of the time, teaching a group of seated students who face away from the camera.

1.3 Definitions, Acronyms And Abbreviations

AVRiL: Automated Video Recording of Lectures (*project name*)
TA: Teaching assistant
PTZ: Point-tilt-zoom
SRS: Software Requirements Specification

1.4 References

SRS Outline (University of Texas at Austin)
[http://www.cs.utexas.edu/~almstrum/cs373/sp05/doc-stds/SRS-outline.html]

IEEE SRS Format
[Local copy at http://suraj.lums.edu.pk/~cs405y07/SRS.pdf]

1.5 Overview Of Document

The rest of the document describes the problem we are trying to solve and what exactly we expect upon completing the project.

Section 2 of the document will discuss the different requirements of the system, and how the system is expected to be used, while Section 3 will discuss in detail the functional, performance and quality requirements of the system.

Section 2: General Description

Purpose of this section: an "executive overview", very client-oriented

2.1 User Personas and Characteristics

This section will describe the intended users of the system, and their personal characteristics that have been kept in mind when developing the system. It also looks at the goals of the intended users, when they use the system. The main intended users of the system that we envision at present are the Lecturer, the Lecture Audience, the Viewer of the Recorded Lectures, and the System Handler:

1. **The Lecturer**
   The lecturer is the person whose actions the system will mainly try to record. We take into consideration that the Lecturer might have no knowledge of how to operate the system. She/He might as well be anxious using computer systems; hence the system will try to minimize his explicit role of the Lecturer in operating the system. It is a goal of the system that the Lecturer can freely use the classroom space, as normal, and doesn't have to act in certain ways to get a better output from the system. The system for her or him should be as unobtrusive as possible. She/He should be able to act as naturally as possible.

   The Lecturer's goal when using the system would be to act as naturally as possible, while the system's role would be to capture his actions with good fidelity.

2. **The Lecture Audience**
   The audience are the people who are interacting and listening to the lecturer at the time of the lecture. Yet again, we take into consideration that none of the audience are literate in using the system effectively. They might as well be unaware of the system's presence. The system will try to capture the interactions of the audience with the lecturer as clearly as possible. The system should be designed in a manner that the audience doesn't have to act in a certain way to get their interactions with the lecturer captured by the system. The audience should be able to act as freely as possible, without the system seeming obtrusive.

   The Lecture audience's goal would be very similar to the Lecturer, in the sense that they should be allowed to act as naturally as possible while the system's role would be to record the audience, lecturer interaction in the best way.

3. **The Viewer of the Recorded Lectures**
   The Viewer of the Recorded Lecture is going to be the main active user interacting with the final output of the system. The Viewer would be the person aiming to watch and gain the best possible feel of the lecture environment. The system will require the Viewer to have basic knowledge of interacting with computers. The viewer should be able to grasp the setting of the lecture in a way to gain not only what the lecturer is

trying to convey, but also the nature of the discussion that takes place between the lecturer and the lecture audience.

The goal of the Viewer would be to as closely understand the happenings in the lecture environment, while the goal of the system will be to replay the lecture environment in a way that closely mimics an academic (or lecture) atmosphere.

4. System Handler

The System Hander would be in some ways the maintainer of the whole system. As the project is in its very early stage, it is not possible to clearly define this persona, but there are some aspects we can still describe. The goal of the system itself will be to minimize his role. Yet, the system will consider that the System Handler is not only well versed in handling and troubleshooting the system, but also computer literate. Though his role cannot be clearly defined at this point, but one of her/his *responsibilities might be* to turn on the system for recording, activate the initial setup sequences of the system before the lecture starts, and make the lecture available for the viewer (See Section 2.1 – 3). The system will not require the System Handler to know the inner workings of the system i.e. e.g. how the lecturer is actually tracked for recording. *It is clearly realized in the present situation that some of the tasks of the System Handler might as well be handled by the Lecturer (Section 2.1 – 1) herself/himself, like when to start and stop the lecture recording.*

The goal of the System Handler would be to give important cues to the system (e.g. when does the lecture start and when does it end), and to make the recorded lecture available for the Viewer, while the goal of the system would be to automate as many tasks as possible and be easy to troubleshoot.

2.2 Product Perspective

Since the system is envisioned to be a standalone, self-contained system, its relationship with other products will be minimal. When the system is fully developed, the System Handler (See Section 3.1 – 4) should be able to set it up in minimal time. It should be as unobtrusive in the lecture environment as possible, so that the lecture atmosphere is relaxed and natural. *The system's overall perspective is to record not only the actions of the lecturer but also the audience in order to make the recorded environment not only as truly representative of the lecture atmosphere. The system also lies in the perspective that it gives an interface to the Viewer which clearly conveys the lecture's milieu.*

External Interfaces with environment of the system would be a set of cameras and a set of microphones (see the **Design Specification** for detailed information). There will also be a system set-up which will govern the actions of these external interfaces. For the Lecturer and the Lecture Audience (see Sections 2.1) the external interfaces should be irrelevant. Only the actual Computing System developed should be of relevance to the System Handler. As mentioned before, the External Interfaces have to be unobtrusive to the Lecturer and the Lecture Audience.

The only meaningful hardware would be for the Viewer of the Lecture. The system should be designed in a way that there are no strict rules governing the system from which the recorded lecture can be viewed, i.e. a Viewer should be able to view from any standard computer with the installed software for the system or maybe just a browser (see the **Design Specification** for detailed information).

2.3 Overview of Functional Requirements
What functions the product should perform

This section will list the main functions the system should perform:

1. **Record the lecture environment**
   This requirement requires the system to record (for later replay) the lecture. This involves two main requirements; the system should be able to record the actions of the lecturer, and the interactions of the audience with the lecturer

   a. **Record the lecturer's actions**
      The system considers that a lecturer might do all the actions which a lecturer normally does in an academic environment. Hence the system considers the possibility that the lecturer will move, speak, do hand gestures, change slides, write on the board or even ask questions from the audience. Thus the system would be required to track all actions of the lecturer in an effective manner.

2. **Record the audience, lecturer interaction**
      The system considers that at some point in the lecture the audience will interact with the lecturer. They might be in form of questions posed to the lecturer or answers given by the audience to questions posed by the lecturer. Hence the function of the system at this point will be to record the actions of the audience itself.

3. **To interface usefully with the Viewer**
   This requires the system to create an interface for the Viewer where she/he can view and experience the lecture. This will not only require the system to show the edited (useful) part of the video of the lecture environment but maybe even give the slides (if present) corresponding to the lecture video. This also poses a requirement on the system that the Software / Interface the user uses is as user-friendly as possible.

4. **Interface for system maintenance**
   This puts a functional requirement on the system that it should have an interface which can be easily used by the System Handler (See Section 2.1 – 4). It should provision the transfer and set up of the recorded lecture to an environment which can be easily accessed by the Viewer. The system should also incorporate

functionality that it can be notified of cues like when to start or stop the lecture recording, or when to initialize the system for a lecture environment.

2.4 General Constraints, Assumptions, Dependencies, Guidelines

This part of the document is premature. We feel that over the months it can go through further change

There are some constraints and dependencies that the system holds for proper functioning. Some major ones are as follows, but by no means is the list exhaustive:

1. **The lecturer's actions should be clearly discernable**
   For proper functioning of the system, and recording of the lecturer's actions, it is important that the lecturer is in the clear discernable view of the system. For example, the system should not be expected to record the lecturer's actions if he hides behind the rostrum or there is an occlusion between the external recording interface of the system and the lecturer standing on the dais.

2. **A single lecturer on the dais**
   This limitation is posed by the system since it might get confused on which lecturer's actions to capture if there are more than one lecturer on the dais at a time. A parallel limitation would be that every object on the dais should be as static (not moving) as possible. Any movements of any other living or non-living object in the vicinity of the lecturer is bound to confuse the system.

3. **Only one person from the audience can speak**
   For proper capture of the audience lecturer interactions, the system would require that only one person from the audience speaks at some time. This limitation is posed since if there are multiple people in the audience interacting at a given moment, this might confuse the system of whose actions to capture.

4. **Recorded lecturer's movements would be limited to a given area**
   This limitation is posed because there would be some limitation to the field of view of the external recording interface. For example, if the lecturer stands in the extreme back corner of the lecture hall, the system shouldn't be expected to capture it. To aid this, the system should clearly specify the captured field of view.

5. **Limited portability of the system**
   Because the system as a whole requires cameras, microphones and a whole controlling system (see the **Design Specification** for detailed information), it might not be an easily portable system. It should not be expected that the system will be able to work outside a lecture hall environment.

2.5 User View of Product Use

This section will describe a common scenario in which the system might be used.

First the System Handler (see Section 2.1 – 4) would come in the lecture hall, where the system would be present, and activate the system and make it run its initializing scripts. At this point it might be important that no one is standing on the dais obstructing the view of the system's external cameras. It might also be a requirement (will be decided when the system is developed) that there is silence in the lecture hall, in order to calibrate the microphones. When the Lecturer (see Section 2.1 - 1) arrives and the audience (see Section 2.1 – 2) is seated the lecture recording of the system might be started either by the System Handler, or if a very easy process, by the Lecturer herself/himself. The recording of the lecture by the system would proceed normally, but if there is a break in the lecture, the system might have to be notified by the System Handler. Of course at the end of the lecture, the system would be explicitly told about it.

Once the lecture recording has finished, and the system has prepared the output, the System Handler will take necessary steps to make it viewable to the Viewer (see Section 2.1 - 3). The Viewer (whoever he may be) might use a software to access the recorded lecture whenever she/he wants. He might be given an interface close to the one given below:

* Courtesy Liu, Qiong et al., Microsoft Research "Automating Camera Management for Lecture Room Environments"

As you can see the Viewer should be able to see both the edited video and maybe even slides (if present). In short the interface given to the Viewer should add as much to the learning experience of the user as possible, and the system should be able to give an environment as close to the recorded lecture itself, while retaining the information conveyed.

Section 3: Specific Requirements
Purpose of this section: Technical information needed to design the software

3.1 Detailed Description of Functional Requirements

3.1.1 Requirement Template
This first section would give the **presentation of the template** that we will use in defining the functional requirements of the system.

Purpose: Description of the functional requirement and its reason(s)
Inputs: Which inputs; in what form/format will inputs arrive; what sources will be used and how can the inputs vary
Outputs: The form, shape, destination, and volume of the output; output timing; process by which the output is stored

For looking at the processes involved see the Design Specification

(for reference see Section 2.3)

3.1.2 Record the lecture environment
Purpose: This functional Requirement will cater for the actual capture by the system of the lecture environment.
Inputs: The inputs to the system would be a video of the lecture and the sounds coming from the lecture hall.
Outputs: The output of the system should be edited recorded lecture environment in a way that it somehow portrays the actual lecture environment. This might be an interleaved video of the lecturer and the audience.

a. Record the lecturer's actions
Purpose: This functional Requirement will record the actions of the lecturer.
Inputs: The input here will be the video view of the lecturer. From it the actions of the Lecturer should be perceived by the system, like his movement, interacting with the audience, writing on the board, hand gestures, etc.
Outputs: The output of this part of the system will be the edited video recording the actions taken by the lecturer, in a way that might be useful for the Viewer. It will also contain times for slide transitions if the lecturer uses some.

b. Record the audience, lecturer interaction
Purpose: This functional Requirement will record the interactions of the audience in the lecture hall.

Inputs: The input here will be both the video feed of the system and the sound from the microphones. From these inputs the system should be able to discern where the person in the audience is.

Outputs: The output of this part of the system will be the overall interactions of the audience with the lecturer.

3.1.3 Interface with the Viewer

Purpose: This functional Requirement will cater for the final output of the system, which the Viewer can use to gain feel of the lecture environment.

Inputs: The input of the system would be the edited video of the system.

Outputs: The output of this functional requirement would be an interface which can easily aid a viewer in experiencing the lecture hall's environment. This might include output in the form of video, slides or any other notes as the need might be.

3.1.4 Interface for system maintenance

Purpose: This functional Requirement will cater for the maintenance and proper functioning of the system. It will provide an interface for taking care of the system's functioning (like giving it cues) and also transfer of edited output of Section 3.2.2

Functioning: This requirement, in some ways, acts as the middle man between requirements of Section 3.2.2 and Section 3.2.4. The output from 3.2.2 comes in as an input to 3.2.4 with the help of this system. Other things that this functionality will incorporate would be the explicit cues for proper set-up, and running of the system.

3.2 Performance Requirements

- One system will only be able to handle one lecture at a time.
- One system should be capable of handling video input from a maximum of 5 cameras at one point in time.
- The compiled video for a lecture should be ready within 24 hours of the end of the lecture.
- System's recording start/stop time should be negligible, i.e. less than 2 seconds.

3.3 Quality Attributes

- The number of dropped frames should not exceed 5% of the frames from the input stream from the camera.
- The output video should be of acceptable resolution for a distance learning program (at least 640 x 480)

- The production/direction quality should be comparable to one that has been recorded and directed by humans. This includes panning, zoom levels, scene changes, camera selection, etc.
- The system should be able to correctly track the instructor and make suitable camera selections 97% of the time.

3.4 Other Requirements

- None at this time