This paper reports on a study that explored the possibility of measuring distance education classes using the Media Sensor, a device designed to sample and record various electronic impulses generated during a distance education session. Research questions explored: (1) the ability of the Media Sensor to identify patterns in the audio and video data; (2) identification of specific types of patterns; (3) the correlation between the pattern and the actual instructional situation; and (4) prediction of instructional events by identifying specific patterns from Media Sensor data.

The research team defined the following categories of classroom events by analyzing videotapes of several distance education sessions: far-end interaction, i.e., dialogs between far- and near-end people; near-end interaction, i.e., dialogs between people at the near end; teacher talking/lecturing; and unknown/other. Furthermore, the team identified patterns by comparing the categories and the raw data recorded by the Media Sensor, testing the validity and reliability of the patterns, and applying them to construct the context within distance education. Findings indicated a high correlation between the Media Sensor data and the instructional situation. Figures illustrate the location of the Media Sensor on the television screen; sample data sheets; Decision Making Trees; video observation data; and coding/decoding stages. (DLS)
Distance Education: Analysis of Far- and Near-Site Interaction Patterns

Philip Hsi-Chang Shih
Noriko Hara
Shaoqing Li
Qing Yi
Indiana University - Bloomington

Abstract of the Study

Media Sensor is a device designed to sample and record various electrical impulses generated during a distance education session. These electrical impulses originate from a variety of visual sources at both the near and far ends of a session, as well as audio from both ends. From the changes of voltage recorded by the Media Sensor on either the visual or audio source, researchers are able to identify specific patterns and analyze the data by using a series of pattern analysis strategies to determine whether the records can present significant events occurred during a session. Accordingly, both the instruction and the classroom management of an instructor can be evaluated by the results of the pattern analysis.

In this study, the research team defined categories of classroom events by analyzing videotapes of several distance education sessions. The categories are: Teacher Talking (at near end), Interaction between Far-end and Near-end, Interaction at Near-end, and Unknown/Other. Furthermore, the team identified patterns by comparing the categories and the raw data recorded by the Media Sensor, testing the validity and reliability of the patterns, and applying them to construct the context within distance education.

Purpose of the Study

The purpose of this study was to explore the possibility of measuring distance education classes by using the Media Sensor.

Research Questions

The research questions for this study are:
- Can the Media Sensor identify pattern in the audio and video data recorded from distance education sessions?
- How to identify specific types of patterns?
- What correlation exists between the pattern and the actual instructional situation?
- How to predict instructional events by identifying specific pattern from data that the Media Sensor generates?
- Is it applicable to other types of distance education session?

Literature Review

Distance education is categorized into three generations according to Kaufman (cited in Bates, 1994, a single technology, multimedia, and two-way interactive technology. The first generation, single media, is mainly conducted by postal service based on paper. There is no interaction between students and the instructor. The second generation, multimedia, is represented by open university in UK. It is widely distributed, but still lacks interactions. The third generation, two-way interactive technology, includes video-conferencing and computer-mediated communication, and allows interactive communications. In the present paper, we will discuss distance education categorized as the third generation. One of the major concerns in distance education is on the learner, and on how instruction can effectively support or facilitate learning (Moore, 1990). There are various studies on different aspects of distance education. Classroom instruction process, students' satisfaction towards distance education, and interaction in distance education are frequently seen in distance education literature. For instance, on classroom instruction process and classroom management, Westbury and Bellack (1971) identified four general categories emerged in a classroom process. They elaborated on "Teaching Action" as one of the major instruction events in distance education sessions. Further more, they identified three teaching actions: "The actions of a teacher directed to the production of intellectual acts within the classroom, such as teacher talking, or making the students to talk; The actions of a teacher directed at making the students 'learningable', such as providing motivational factors; The actions of a teacher which are intended to contribute directly to the students' learning, such as providing practice of learned materials or techniques." (P. 243-245)
On student satisfaction towards instruction in distance education sessions, Pugh and Siantz (1995) found that "student satisfaction did improve over time." (P. 21) They also reported that "evidence from the observer comments and student comments tended to support that there was more interaction between the instructor and students and between students at the near end." (P. 22) This further supports that interaction is a major event that might lead to the level of student satisfaction or dissatisfaction.

On interaction in distance education, Moore (1990) presented three types of interaction: learner-content interaction, learner-instruction interaction, and learner-learner interaction. The author also discussed issues such as "what level of interaction is essential for effective learning, what is good interaction and how to achieve it, what the real-time interaction contribute, and whether it is worth the cost." Obviously, various types of interaction in distance education sessions are major categories of events that this study should identify and analyze.

There is no literature, however, on pattern analysis on the distance education session, and no study has been done on such device as media sensor since the device was first invented by Appelman. The study intended to fill in this gap.

A minor concern of this study is whether the tool used for audio and visual data recording -- the Media Sensor is effective in recording data. Kounin (1970) in his study on group management in classrooms listed out the deficiency of using human observer as data-gathering medium. This finding indirectly supports the use of mechanical device for an accurate and complete data analysis.

**Significance of the Study**

Findings from the study can be applied by other researchers in distance education. First, this study intends to reduce the time of observing each session in distance education by establishing a measurement of pattern to identify what is really happening in the session, without watching the video tapes. Second, this study can help other researchers to expand their study of evaluating instructor's pedagogy by examining outcomes, teaching styles, and classroom interaction. Third, this study provides same justification of the cost-effectiveness in using high-technology in distance education. For example, if there are only few interactions between far-end and near-end, it seems that we don't have to use such expensive technologies for distance education. Instead, we could just use a VCR to record the class, and then send the videotapes to students at remote sites.

**Methodology**

This study is to develop a measuring tool to predict what happens during a session in distance education without watching a video tape. The Media Sensor invented by Dr. Appelman, a professor at Indiana University, Bloomington, generates coded data sheet from video taped distance education sessions with time coding for both audio and video sources. The team analyzed the data to find specific patterns to identify significant events.

In order to identify the pattern of the audio and visual data collected from several distance education sessions taking place at the School of Education, Indiana University, Bloomington, an effective and efficient method of pattern identification, definition and analysis needed to be decided. Part of the literature review was conducted in search of such a critical tool for this study. Frick recommended the method of "Analysis of patterns in time (APT)" for analyzing observable phenomena so that the patterns of these phenomena can be recognized. APT measures temporal relations between variables by counting their occurrence. Frick found out that using proper sampling strategies could lead to the prediction of temporal patterns from APT results (Frick, 1990).

In the example of Classroom Observational Study, Frick used APT to investigate classroom events. "Highly trained observers collected observational data on paper-and-pencil coding forms. For illustration, only two classifications are discussed: available instruction (direct, nondirect, null), and student orientation to academic instruction (engaged, nonengaged, null). . . . The observers also coded the type of target student orientation to academic instruction that was occurring simultaneously with the type of available instruction." (1990, p.182)

After the data collection, grouping, and the calculation of its means, queries were made about these APT scores by researchers to look for recurring patterns or combinations of events in classroom. Researchers can also aggregate duration of certain kinds of events to see what proportion of the overall time they occupy. This data collection, grouping, and analysis method provides a foundation for the methodology of this study.
Media Sensor

The Media Sensor is a device that can capture both changes of brightness on a screen and voices from Far-end and Near-end sites. It is designed to sample and record the various electrical impulses generated during a distance education session. These electrical impulses originate from a variety of video and audio sources at both the near and far ends of the session. Video sources are sampled via photoelectric cells placed in a grid pattern over the screen of a television. The sensors for video are located in five areas on the screen (see Figure 1 below). Audio is sampled at the echo-canceller as "Far-end Receive" and "Near-end Send" sources (sensor 7 and 8 respectively, sensor 6 doesn't record any signal). Figure 2 is an example of the data sheet generated by Media Sensor.

Figure 1: Sensor location on a TV screen (visual)

Figure 2: An example of data sheet

Data

Visual and audio sources pass through an analogue to digital converter and are sampled at various rates by the Media Sensor. While data sampling is occurring, records of each sample are sent to a computer as ASCII text for recording and analysis. Typical patterns of a data sheet is shown at the end of this Methodology section (Figure 5).
Pattern Definition Process

The analysis on the first two sessions generated the first version of the pattern definition. The team tested the accuracy of this definition in the analysis of the next two sessions. The result of the second round analysis generated information for the revision of the first pattern definition.

First Pattern Definition

Two video-taped sessions were provided to the team. The team listed several events that occurred in class by brainstorming and watching a 30-minute segment of a session on the videotape. Then, the team decided how to code those events on the data sheet, as shown below.

- **L** = Lecturing
- **T** = Teacher talking (at Near-end)
- **FS** = Far-end student talking
- **NS** = Near-end student talking
- **N** = Noise
- **NA** = Lack of audio/voice
- **I** = Interaction
  - **FS** $\rightarrow$ **NS** (between Far-end students and Near-end students)
  - **FS** $\rightarrow$ **NS**
  - **FS** $\rightarrow$ **T** (between Far-end students and teacher)
  - **NS** $\rightarrow$ **NS**
  - **NS** $\rightarrow$ **T**
- **M** = Camera Movement
  - **T** $\rightarrow$ **S** (switch between teacher and students) **S** $\rightarrow$ **T**
  - **P** $\rightarrow$ **D** (switch between people and document) **D** $\rightarrow$ **P**
  - **S** $\rightarrow$ **S** (switch between student and student) **S** $\rightarrow$ **S**
  - ZI, ZO (zoom in, zoom out)

Although the team noticed that some of the above events were not predictable from the data sheet without knowing the instructional content, the team decided to observe two video taped sessions and code the entire session according to those categories. Then the team compared the coding results. The two persons who watched the same session checked with each other to detect any difference in the coding.

The next task was to identify patterns in order to predict events from data sheet. The team looked at the data sheet vertically from sector 1 through 8 and then separated voice parts (sector 7 and 8) from video parts (sector 1 through 5).

Throughout the entire data analysis process, the team defined Noise as single and short bars. The team understands that when counting the amount of time for interaction between near end students and the teacher, it is possible for the team to ignore some periods of teacher talking because it is hard to tell if the teacher is answering the questions, which is regarded as an interaction, or starting to lecture merely from looking at the data sheet.

The followings are our patterns for four categories, lecture, interaction between Near-end and Far-end, interaction at Near-end only and Break.

**Lecture:**
- **Audio:** mostly blank, randomly single signal appears in section 8;
  - Signals are in continuous chunks.
- **Visual:** More signals in sector 4 and 5;
  - Sector 1 has almost no signal;
  - Sector 2 has signals sporadically. It has more signals than sector 1;
  - 3 or 4 out of 5 sectors in 1-5 visual sectors have signals.

**Interaction (Near- and Far-end):**
- **Audio:** Signals appear in pair in sector 7 and sector 8.
- **Visual:** Signals appear 3 or 4 out of 5 sectors in 1-5 visual sectors.

**Interaction (Near-end only):**
- **Audio:** More blank among the signals in sector 7;
  - No signal in sector 8.
Visual: 4 or 5 out of 5 sectors in 1-5 visual sectors have signals more dense than those in sector 3 through 5.

Break:

Visual: More dense than any other events in class time.

To test the validity of the patterns, the team gained two more video-taped sessions and paired up. Each pair looked at one data sheet separately without watching the videotape, and compared how much time was spent for each event in a session. Then, the same pair looks at another session on videotape, coded its data sheet and made comparison. Therefore, each pair conducted two analyses: (1) watched a video taped session and coded its data sheet during the observation; (2) coded events on another data sheet for the second session without watching the video tape.

The team also decided to mark colors on the data sheet to make it easy to differentiate and compare each other's coding on the categories of events. For example, pink marker was used for interaction among Near-end students. The team compared the coding results from data sheet only with the results from video tape observations to see if any agreement (validity) can be achieved. Refer to Table 2 on Page 18 for the results of video tape observation and prediction from data sheet.

As a result, the team obtained fairly good observer agreement on the total amount of time for each category calculated by persons who watched the video tape for coding. However, the team had a variety of results between the pair partners who predicted events from the data sheet only, and between the pair who coded only the data sheet and the pair who coded the data sheet by watching the videotape. This indicated that our patterns lack strong reliability and validity.

As a result of the data analysis at this phase, the team obtained advice from the instructor and decided to include Unknown as a category. The team also realized that the original definitions of patterns were subjective and vague. The team redefined the patterns, and created the decision making tree for pattern recognition. Our team obtained a new session on video tape taught by an instructor who had not appeared in the previous sessions. This time, all the team members predicted events from data sheet without watching the video tape, compared the total amount of time on each categories of events, and then watched the tape together to see if the coding results from data sheet matched what actually happened. The result was very positive, which showed to the team that the new definition of pattern was applicable at least to the analysis of this session.

Decision Making Tree

The team developed Decision Making Tree to present the pattern definition graphically. It can also be used as a method for data interpretation. The first Decision Making tree was developed after completing the first round of data analysis on the first two sessions. After the data analysis on the next two sessions, the team modified the first version of the Decision Making Tree. The following diagrams are the first and second versions of the Decision Making Tree.
Figure 3: Decision Making Trees (version I)
Final Pattern Definition
The followings are our final definition of events and patterns:

- **Far-end interaction**: dialogues between far-end and near-end people.
- **Patterns**:
  1. In Sector 8, if there is a group of signals which have blanks among the signals that last longer than 1 minute, it is a far end interaction.
  2. In Sector 8, occasionally there are individual bars whose length is longer than 1/3 of the maximum length of a signal bar. These individual bars are also regarded as far end interactions. If a blank among single signals or chunks of signals is longer than 1 minute, go to Sector 7. If a blank in Sector 7 is smaller than 1 minute, it might be either teacher talking or near end interaction. If the blank is equal to or more than 1 minute, the signal is regarded as Unknown/Others, which can be break, noise, technical problems (e.g., with sensors or recording devices, etc.).

- **Near-end interaction**: dialogues between people at near end.
- **Patterns**:
  1. Near-end interaction starts at the point when there are signals in each sector of sector 1-5 simultaneously, and ends when there is a signal or group of signals starts at Sector 8, which is a beginning of Far-end interaction.
  2. Near-end interaction starts at the point when there are signals in each sector of sector 1-5 simultaneously, and ends when the same kinds of signals through sector 1-5 are identified. Thus, the interval between these two signal groups is regarded as the length of this near end interaction time.

- **Teacher talking**: teacher is either talking or lecturing.
- **Patterns**: If it is not a Near-end Interaction, and there are signals in 3 or 4 out of 5 sectors in sector 1- 5, it is a Teacher Talking or Lecturing.
Far-end Interaction

Near-end Interaction
Research Results

According to the definition of pattern, the team studied other two distance education program, video 3 and video 4 (video 1 & 2 were the first two video sessions studied by the team). The results are shown in Table 1 and 2. D1 and D2 are the observers who predicted the instructional events by only looking at the data sheet; DV1 and DV2 are the observers who worked on both data sheet and video. The figures under D1, D2, DV1 and DV2 are the amount of time each category of instructional event lasted in these two sessions.

Pearson correlation coefficient is calculated within a pair of observers and between the two pairs. The coefficient within a pair (D1 and D2, DV1 and DV2) can be regarded as an indicator of observer agreement, coefficient between the mean of the two pairs (the mean of D1 and D2 and the mean of DV1 and DV2) and is served as an indicator of validity.
The observer agreement between the people who worked only on data sheet separately is -0.4176 (p < .484) in video 3 and .5827 (p < .303) in video 4, which means there is very little agreement when the two observers predicted instructional events by applying the pattern definition. The observer agreement between the people who viewed both video and data sheet is much higher, .9873 (P < .002) in video 3, and .8962 (.04) in video 4, which indicates there is very high agreement between those who looked at both video and data sheet.

The validity of the two observation is different from each other, .1489 (p < .8111) in video 3, and .947 (p < .015) in video 4. The validity of the first one is low comparing to the desired correlation, which is .60-.70. The second one is extremely high. However, according to the observers' comments, the defined pattern didn't work effectively, because there was so much uncertainty when the observers started to decided which category an event belonged to. Therefore the result could be a random guess. Besides, even though there is very good match between the total amount of time of each category of instructional events as predicated by D1/D2 and traced by DV1/DV2, those events observed by different observers did not match in terms of starting point and ending point of time, and often times these were totally different. Based on the statistics result and the observer's comments, the team changed the specifications of pattern definition. The team then applied the updated definition to a new distance education session, video 5 (Table 3).
There were 4 observers working on video 5, D1, D2, D3, and D4. All observers analyzed the data sheet without viewing the video according to the updated definition of pattern and the decision making tree. Table 3 is the result of this analysis. Table 4 is the Pearson correlation coefficient between any two decoding. D2, D3 and D4 have very high correlation, .9962 between D2 and D3, .9986 between D2 and D4, .9991 between D3 and D4. The correlation between D1 and the rest ones, which is .8595 (D1 and D2), .8960 (D1 and D3) and .8768 (D1 and D4) respectively, is not as high as those between any two of D2, D3 and D4 because of some misunderstanding on the pattern definition of Near-end Interaction and Teacher Talking between this observer and the others. This explains why there is bigger difference in Teacher Talking time and Near-end Interaction time between D1 and the rest of the observers. Due to the time constraint, the team did not count the amount of time each event lasts in the video as the team did in video 4 and video 3. However, the team compared the analysis result with the video and found out almost all the events in video matched with the decoding results in terms of the categories of the events and starting point and ending point of time of each event.

Table 3: Observation on video 5 (80 minutes)

<table>
<thead>
<tr>
<th></th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>D4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher talking</td>
<td>8'00&quot;</td>
<td>25'40&quot;</td>
<td>21'53&quot;</td>
<td>23'31&quot;</td>
</tr>
<tr>
<td>Interaction (N)</td>
<td>17'00&quot;</td>
<td>1'00&quot;</td>
<td>2'40&quot;</td>
<td>1'08&quot;</td>
</tr>
<tr>
<td>Interaction (F)</td>
<td>52'40&quot;</td>
<td>52'31&quot;</td>
<td>52'10&quot;</td>
<td>52'57&quot;</td>
</tr>
<tr>
<td>Break</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Unknown</td>
<td>2'00&quot;</td>
<td>1'00&quot;</td>
<td>2'24&quot;</td>
<td>2'10&quot;</td>
</tr>
</tbody>
</table>

Table 4: Pearson correlation coefficient between pairs of observation on video 5

<table>
<thead>
<tr>
<th></th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>D4</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>1.000</td>
<td>(.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D2</td>
<td>.8595</td>
<td>(.0562)</td>
<td>1.000 (.000)</td>
<td></td>
</tr>
<tr>
<td>D3</td>
<td>.8960</td>
<td>(.040)</td>
<td>.9962 (.000)</td>
<td>1.000 (.000)</td>
</tr>
<tr>
<td>D4</td>
<td>.8768</td>
<td>(.051)</td>
<td>.9986 (.000)</td>
<td>.9991 (.000)</td>
</tr>
</tbody>
</table>

Conclusion /Discussion /Limitation

Conclusion

The team only checked one Distance Education session using the latest pattern definition and analysis method the team developed. The conclusion can be only applied to this session. It is concluded from this study that there are high correlation between the data generated by the Media Sensor and the instructional situation. The patterns are recognizable from the data sheet, and applicable in recognizing the instructional events without seeing the actual session.
Limitations:

- **Validity of the study:**
  The team only studied five sessions, totally 8 hours and 50 minutes. Four out of these five sessions were taught by the same instructor to the same group of students. The definition of patterns obtained from and tested in these sessions might not be valid.

- **Generalization:**
  Because of the small amount of sessions the team studied, and the unique teaching style each instructor has, the patterns identified from these sessions might not be applicable to other sessions. The team only viewed one session after the final revision to pattern definition. Thus, the conclusions we made are only applicable to this study.

- **Content-free:**
  There is no way to differentiate the technical/equipment testing from instruction related interaction at the beginning of each session and interaction.

  - The data depends on the experience and working style of the instructor who operates the camera. Different instructors could have different styles or habit when they operate the camera, and it affects the signals on the data sheet.

  - The data record partly depends on the sensitivity of the sensors. There are some differences in data records when rearranging the position of the sensors.

  - It is impossible to differentiate use of document camera and near-end interaction because both of them are accompanied by signals in all 5 video sectors.

  - There are some technical problems when recording the signals.

    One of the problems the team encountered was that it was uncertain whether the Media Sensor was reliable because there were some contradictions between the video tape session and the data record. Appelman explained that the electronic equipment could be more sensitive than human eyes, but on the other hand, if the background remains the same, media sensor can't detect the change even the camera moved. In this case, it happened when the team did not see any movement while watching the video, but there are signals on the data sheet, or vice versa.

**Discussion**

In the whole process, there are two stages: coding and decoding, as shown in Figure 5. Coding happens when patterns were defined from looking at the video and using the data sheet; decoding happens when applying these patterns to predict the actual situations. In either of coding and decoding process, some information is lost inevitably. There is also some distortion between the predicted patterns and the actual happenings when going through the coding and decoding stage. It is understandable that it is not 100 percent accurate to apply the pattern definition to predict the instructional events.
To generate more accurate and applicable pattern, the team suggests:
- Establish a standard of camera movement, for example, the camera is only focusing on the relevant objects, with only the speaker appears on the camera.
- Revise the computer program depicting the data with smaller scale of time.

**Future Research**

Further validity studies are needed in which a wide spectrum of distance education sessions are coded by trained observers who will use the same pattern definition and pattern recognition method by viewing the data sheet generated by the Media Sensor.

If such validity is established, then the next step would be to see if the APT program can identify the patterns seen by human eyes when looking at the sessions. The APT program would be looking at some form of the raw data produced by the Media Sensor.

A third step is to specify, define and identify the pattern of different events in unknown part, such as break, noise, etc.

If these goals can be achieved, these patterns (from Media Sensor and APT computer program) can be correlated with other measures of distance education, such as effectiveness (e.g., student learning achievement), efficiency (e.g., cost-benefit analysis).

**References**


NOTICE

REPRODUCTION BASIS

This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").