ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization) Vol.2, Special Issue 1, March 2014

Proceedings of International Conference On Global Innovations In Computing Technology (ICGICT'14)

Organized by

Department of CSE, JayShriram Group of Institutions, Tirupur, Tamilnadu, India on 6th & 7th March 2014

Discovering Patterns of Human Interaction in Meetings Using Tree-Based Mining

S.Abarna[1]

PG Scholar, Department of CSE, Dr. Sivanthi Aditanar College of Engineering, Tiruchendur-628215, Tamilnadu, India[1]

ABSTRACT: Mining frequent tree patterns is an important research problem with broad applications in bioinformatics, digital library, e-commerce, and so on. Previous studies highly suggested that pattern-growth methods are efficient in frequent pattern mining. Multimedia records of meetings contain a rich amount of project information. However, finding detailed information in a meeting record can be difficult because there is no structural information other than time to aid navigation. In this paper we survey and discuss various ways of indexing meeting records by categorizing existing approaches along multiple dimensions. In particular, a survey of users who watch meeting recordings on a regular basis revealed that it is also useful to provide (1) speaker-related information, including who the speaker is talking to, looking at, and being interrupted by, and (2) more control of the interface, including changing the relative sizes of the speaker and context views and navigating within the context view. A novel probabilistic framework is proposed for analyzing cross-modal nonverbal interactions in multiparty face-to-face conversations.

KEYWORDS: Multimedia meeting record, indexing, Face-to-face multiparty conversation

I. INTRODUCTION

Face-to-face conversation is one of the most basic forms of communication in our life and is used for conveying/sharing information, understanding others' intention/emotion, and making decisions. To enhance our communication capability beyond conversations on the spot, the automatic analysis of

a conversation scene is a basic technical requisite to enable effective teleconferencing, archiving/summarizing meetings, and to realize communication via social agents and robots. The conversation scene analysis targets various aspects of conversations, from individual/group behaviors such as "*who is speaking now?*" and "*who is talking/listening to whom?*", to context/mental status such as "*who made him angry?*"and "*why is she laughing?*". In the face-to-face setting, the messages include not only verbal but also nonverbal messages. The nonverbal messages are expressed by nonverbal behaviors in multimodal channels such as eye gaze, facial expressions, head motion, hand gesture, body posture and prosody; psychologists have elucidated its importance in human communications. Therefore, it is expected that conversation scenes can be largely understood by observing people's nonverbal behaviors with sensing devices such as cameras and microphones.

As such, meetings contain a large amount of rich project information that is often not formally documented. Capturing all of this informal meeting information has been a topic of research in several communities over the past decade. The most common way to capture meeting information is through note-taking. However, fully writing down the content of a meeting is a difficult task, and can result in an inability to both take notes and participate in the meeting. The prospective benefits of having a meeting record on the one hand and the problems with traditional meeting recording on the other hand have triggered the use of technology to create meeting records. While technology automatically captures meeting activities, humans are left free to actively engage in discussions and synthesize what is going on around them, without worrying about tediously preserving details for later memory. The method of the choice for recording meetings has been audio and video, which can provide a comprehensive meeting record that allows people to see who was present and what was discussed.

Moreover, this recording technology is unobtrusive and ideally does not require any further interaction during the meeting once recording has been started. The greatest problem with audio/video recordings is that they are sequential and do not provide any structural information but time to navigate. A team member will not exhaustingly watch one hour of a meeting when she can go ask a colleague who attended the meeting for a quick summary. Multimedia records of meetings will only be generally useful if tools or technologies exists that help users avoid replaying much of what has been recorded.

Existing annotation schemes for human movement can be classified according to the amount of detail they capture, where high detail seems to be proportional to high annotation cost and a low level of abstraction.

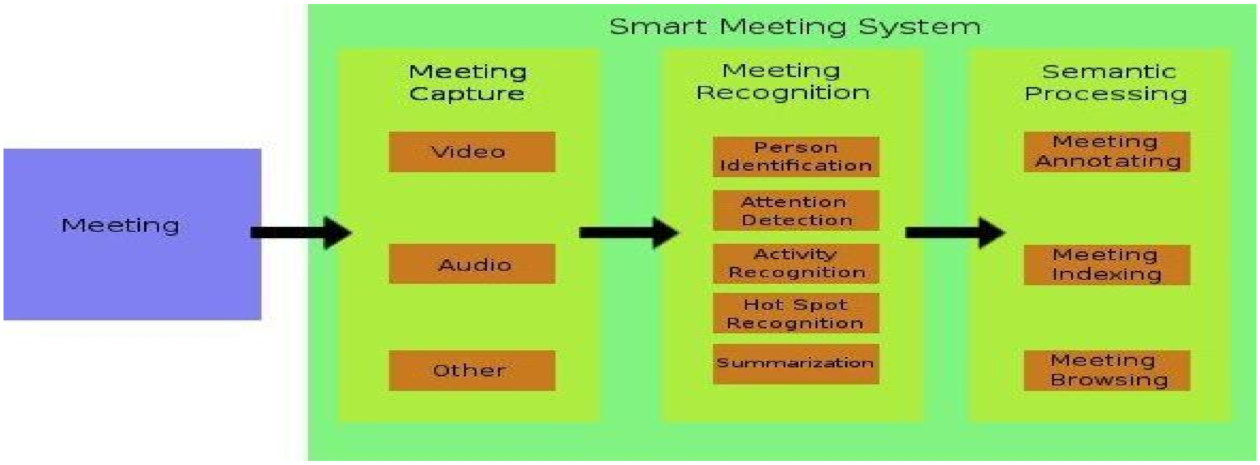
Related works:

The method of visualizing hand-annotated data with an animated virtual character can also be employed for the systematic study of the perception of nonverbal behavior (cf. Kramer *et al.*, 2003). Buisine *et al.* (2006) have analyzed blended emotions in perception tests by replaying them from manual annotations. A virtual character allows one to switch on or off particular aspects of the annotation and thus explore which parts of the annotation carry the biggest effect, an approach that the authors call copy-synthesis. The need for a standardization of gesture form has recently been formulated by an international initiative who are developing a unified XML language called BML (Behavior Markup Language) for sending descriptions of form and timing of nonverbal behaviors to an animation engine that controls a virtual character (Vilhjalmsson *et al.*, 2007; Kopp *et al.*, 2006). Our scheme shares with BML the insight that transition points between movement phases represent a key abstraction for animation. However, we impose a stronger restriction by choosing to focus on the stroke phase, showing that this is sufficient to re-create the observed gesture. In terms of gesture form description, BML is currently under development, but is aiming for a complete set of descriptors where the question remains which components are best suitable for annotation and re-creation.

Before launching into the design of our system, it is useful to situate our work in terms of related research. As is presented in the next section, Second Messenger uses visualizations of speech data to express the presence and participation of individuals in a conversation. Within the domain of technology that aims to *influence* individual behavior and participation, we are aware of two projects that utilized concepts from social psychology to influence involvement in community discussion: a project at Georgia Tech (Hudson & Bruckman, 2004) applied the concept of the "bystander effect" to analyze the cause of low contribution to discussions within educational communities; the movie recommendation Web site Movie Lens has been used as a research platform to evaluate whether theories of social loafing and goal setting can be applied to interface design to encourage more or less participation (Beenen et al., 2004). These projects have demonstrated that concepts from social psychology can be applied to online community development.

Benefits of meeting records:

The purpose of meetings is to move group activities towards a common goal. Meetings help team members coordinate their work, come to a shared understanding of their work, and focus on their task . Team members present information to others and collaborate with each other through reviewing, evaluating, discussing, problem solving, and deciding. There are also social reasons to meet such as the need to belong, to achieve, and to make an impact, or the desire to communicate, build, and share a common reality.. Good meeting practices, which are mainly based on heuristics, emphasize the importance of preparing meeting records and suggest capturing at least all the decisions made, all the action items assigned, and all the open issues. By doing so, the meeting record helps not only build a shared group memory but also makes meetings more efficient by decreasing the need to revisit decisions made, by making it possible to easily recall open issues and deferred items, and by having increased confidence that action items will be done.



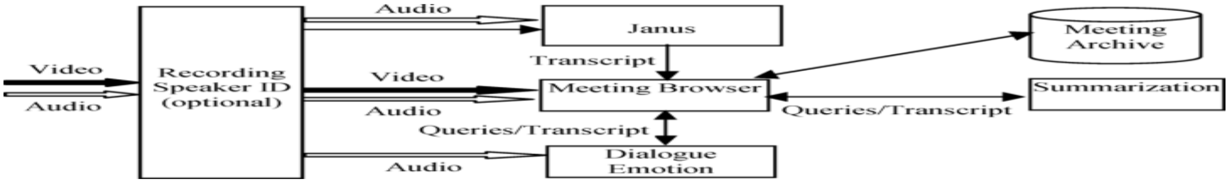
Finally, there are strong arguments for keeping meeting records from a knowledge management perspective also to locate experts in a company similar to the email mining approach in Discovery Server.

Problems of Meeting Captures:

Private note taking is common practice in meetings and often a designated person prepares meeting minutes, which are usually a narrative meeting record. Khan and Whittaker et al have studied note taking as the primary way of recording what occurred during a meeting . The studies show a general satisfaction with note taking; about 33% reviewed their notes regularly. However, 70% reported that there had been occasions when they wished they had written better notes. The major difficulties encountered with note taking are

- Failure to note facts that turned out to be crucial later,
- Illegible names,
- Not enough time to write everything,
- reduced ability to participate, and
- Inadequacy of notes for later detailed

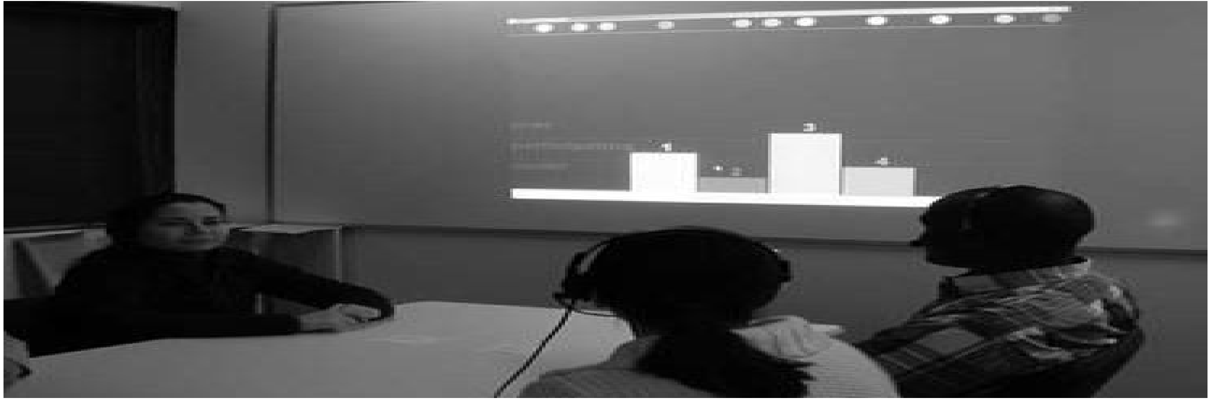
In contrast, meeting minutes are a narrative public record of participants, decisions and issues. Minutes are usually prepared by a dedicated person (recorder) during the meeting and can provide a very good summary. However, a dedicated recorder is required, the ability of the recorder to participate is reduced, and the minutes do not capture individual information needs and are subject to the interest and the contextual knowledge of the recorder.



A more technologically oriented approach is to video-tape, or digitally records the audio and video of a meeting. This recording technology is unobtrusive, provides a comprehensive record, and ideally does not require any further interaction during the meeting once recording has been started. These techniques allow for a maximum time savings factor of 1.5 to 2.5. This may be acceptable for people who were not at a meeting, and thus want to review the entire event. However, this is time-consuming if one instead wants to focus on specific topics or details of the meeting.

Human Interaction:

Human interactions in a meeting discussion are defined as social behaviors or communicative actions taken by meeting participants corresponding to the current topic. Various interactions imply different user roles, attitudes, and intentions about a topic during a discussion. The definition of interaction types naturally varies according to usage. In this paper, we mainly focus on the task-oriented interactions that address task-related aspect. The other communicative actions that concern the meeting and the group itself (e.g., when someone invited another participant to take the floor) are not included. For generalizability, we create a set of interaction types based on a standard utterance-unit tagging scheme propose, comment, acknowledgement, request Info, ask Opinion, pos Opinion, and neg Opinion. The detailed meanings are as follows: propose a user proposes an idea with respect to a topic; comment—a User comments on a proposal, or answers a question; acknowledgement—a user confirms someone else's comment or explanation, e.g., "yeah," "uh huh," and "OK;" request Info— a user requests unknown information about a topic; ask Opinion-a user asks someone else's opinion about a proposal; pos Opinion—a user expresses a positive opinion, i.e., supports a proposal; and neg Opinion—a user expresses a negative opinion, i.e., disagrees with a proposal.

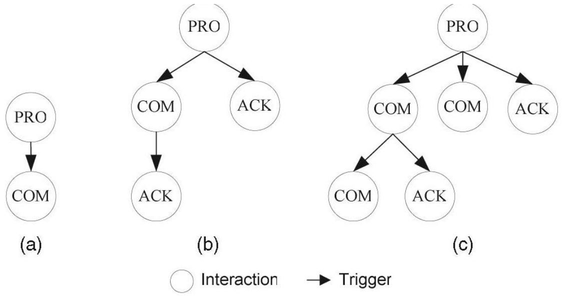


However, this is time-consuming if one instead wants to focus on specific topics or details of the meeting. Computer support for meetings often provides additional meeting functionality that can be recorded, such as electronic whiteboards in conference rooms, or application sharing for distributed meetings. IBM's Same time conferencing as well as WebEx's meeting software feature recording of all media involved in a distributed meeting. However, these additional media are either only statically presented, such as the image of the whiteboard, or in the same VCR-like playback. Thus, while more information is being recorded, that record is still difficult to review.

Interaction Flow Construction:

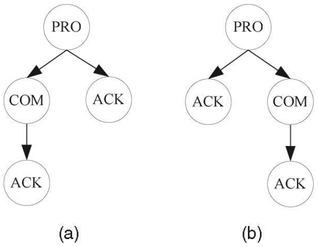
Based on the interaction defined and recognized, we now describe the notion of interaction flow and its construction. An interaction flow is a list of all interactions in a discussion session with triggering relationship between them. We first give the definition of a session in a meeting discussion.

Session: A *session* is a unit of a meeting that begins with a spontaneous interaction and concludes with an interaction that is not followed by any reactive interactions.

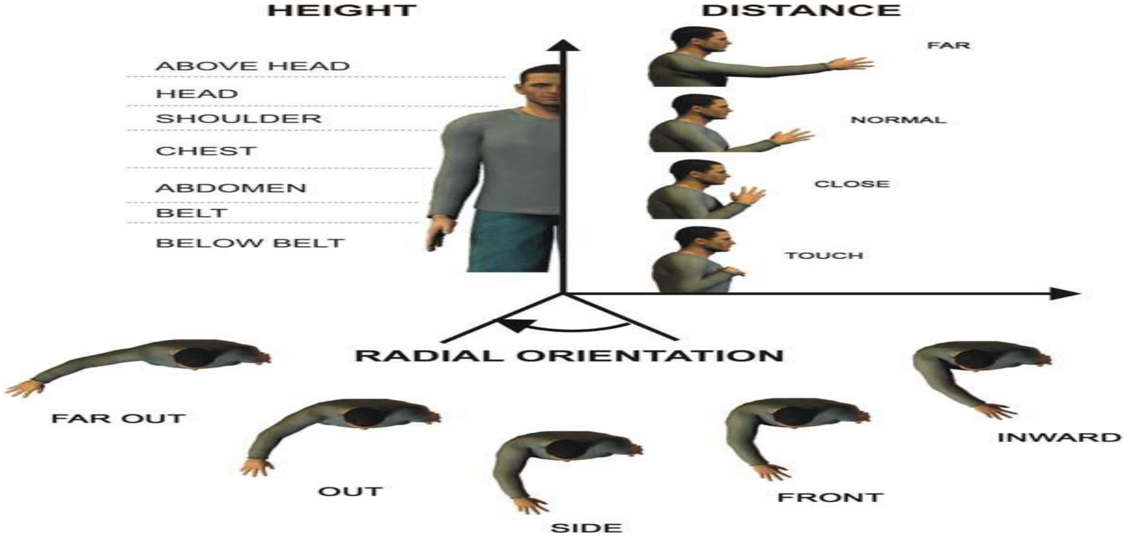


Here, spontaneous interactions are those that are initiated by a person spontaneously, and reactive interactions are triggered in response to another interaction. For instance, propose and ask Opinion is usually spontaneous interactions, while acknowledgement is always a reactive interaction. Whether an interaction is spontaneous or reactive is not determined by its type (e.g., propose, ask Opinion, or acknowledgement), but labeled by the annotator manually. Hence, a session contains at least one interaction (i.e., a spontaneous interaction). A meeting discussion consists of a sequence of sessions, in which participants discuss topics continuously.

Isomorphic Tree: Given two trees, T1=(V1,E1) and T2 =(V2, E2), if tps(T1)#tps(T2)and through exchanging the places of siblings on T1 or T2 (i.e., commutation processing), tps(T1)= tps(T2), we call T1 and T2 isomorphic trees.



The purpose of the isomorphic tree definition is to find the same tree structure by exploiting temporal independence in the original interaction trees. For instance, two trees are isomorphic because although their tree preorder sequences are different (PRO-COMACK-ACK and PRO-ACK-COM-ACK), through commutation processing, their preorder sequences become the same.



In summary, for each stroke based gesture we encode 2 positions where each position is expressed by 5 attributes. Adding handedness and trajectory gives us 12 attributes to code for the spatial form of a gesture. Hold gestures only require 1 position, for the beginning of the hold.

Algorithms for Pattern Discovery:

With the representation model and annotated interaction flows, we generate a tree for each interaction flow and thus build a tree data set. For the purpose of pattern discovery, we first provide the definitions of a pattern and support for determining patterns.

Pattern: Patterns are frequent trees or sub trees in the tree database.

Support: Given a tree or sub tree T and a data set of trees TD, the support of T is defined as

Supp (T) = number of occurrences of T/total number of trees in TD

Algorithm 1: fitm (TD,σ) (Frequent interaction tree pattern mining)

Input: a tree database TD and a support threshold σ

Output: all frequent tree patterns with respect to σ

Procedure:

(1) Scan database TD, generate its full set of isomorphic trees, ITD

(2) Scan database ITD, count the number of occurrences for each tree t

(3) Calculate the support of each tree

(4) Select the trees whose supports are larger than σ and detect isomorphic trees; if m trees are isomorphic, select one of them and discard the others

(5) Output the frequent trees

In developing our frequent sub tree discovery algorithm, we decided to follow the structure of the algorithm Agrawal used for finding frequent item sets, because it achieves the most effective pruning compared with other algorithms. The high-level structure of our frequent Sub tree mining algorithm is shown in Algorithm 2. It first calculates the support of each node and selects the nodes whose supports are larger than σ to form the set of frequent nodes, F1 (Steps 2-3). It then adds a frequent node to existing frequent *i* - sub trees to generate the set of candidates with *i* + 1 nodes (Steps 4-8).

Algorithm 2: fistm (TD,σ) (Frequent interaction sub tree pattern mining)

Input: a tree database TD and a support threshold σ

Output: all frequent sub tree patterns with respect to σ

Procedure:

```
(1)  i← 0
(2) Scan database TD, calculate the support of each node
(3) Select the nodes whose supports are larger than σ to form F1
(4) i← i + 1
(5) For each tree ti in Fi, do
(6) For each node t1 in F1, do
(7) Join ti and t1 to generate Ciþ1
(8) Sub tree Support Calculating (TD; tiþ1)
//calculate the support of each tree in Ciþ1
(9) If there are any trees whose supports are larger than σ, then select them to form Fi+1 and return to Step (4)
(10) Else output the frequent sub trees whose supports are larger than σ
```

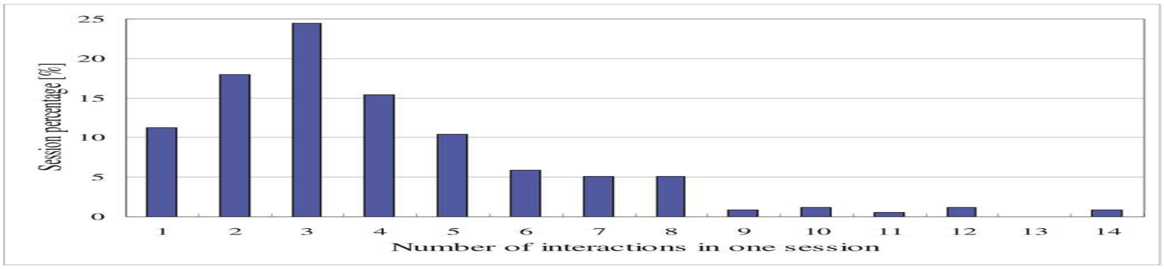
The sub procedure, Sub tree _ Support _Calculating, first creates sub trees of each tree t in TD with the size of sub trees the same as that of *st* (Steps 3-4). Then, for each sub tree of t, it generates its isomorphic trees and compares their string codes with that of st. If it matches, the number of occurrences of *st* is increased by 1 (Steps 5-14). It finally calculates and returns the support of st (Steps 15-16).

Sub procedure: Sub tree Support Calculating (TD, st)

```
(1) Count← 0
(2) Supp (st) ← 0
(3) For each tree t £ TD do
(4) Create sub trees S of t with any item s ∑ S; |s| =|st|
(5) Flag← false
(6) For each item is s ∑ S do
(7) Generate isomorphic trees IS of s
(8) For each item is is∑ IS do
(9) If tsc (is) = tsc (st) then
(10) Count ←count + 1
(11) Flag← true
(12) break
(13) If flag = true then
(14) break
(15) supp (st)← count=/|TD|
(16) return supp (st)
```

Data Set:

Our studies involve four real meetings, lasting 20 minutes on average. Multiple devices, such as video cameras, microphones, and motion sensors, were used for capturing the meetings. The four meetings include one PC purchase meeting (26 min, discussing PCs to be ordered for the laboratory, such as types, configuration, size, weight, manufacturer, etc.), one trip-planning meeting (18 min, discussing time, place, activities, and transportation for a summer trip), one soccer preparation meeting (23 min, talking about the players and their roles and positions in an upcoming match), and one job selection meeting (10 min, talking about factors that will be considered in seeking a job, such as salary, working place, employer, position, interest, etc.). Each meeting had four participants seated around a table. Human interactions were detected by using a multimodal approach. In order to use a correct data for mining, we tuned the interaction types manually after applying the recognition method.



CONCLUSION

We proposed a tree-based mining method for discovering frequent patterns of human interaction in meeting discussions. The mining results would be useful for summarization, indexing, and comparison of meeting records. They also can be used for interpretation of human interaction in meetings. In the future, we will develop several applications based on the discovered patterns. We also plan to explore embedded tree mining for hidden interaction pattern discovery. Embedded sub trees are a generalization of induced sub trees, which allow not only direct parent child branches, but ancestor-descendant branches.

REFERENCES

[1] C. Wang, M. Hong, J. Pei, H. Zhou, W. Wang, and B. Shi, "Efficient Pattern-Growth Methods for Frequent Tree Pattern Mining," Proc. Pacific-Asia Conf. Knowledge Discovery and Data Mining (PAKDD '04), pp. 441-451, 2004.

[2] W. Geyer, H. Richter, and G.D. Abowd, "Towards a Smarter Meeting Record—Capture and Access of Meetings Revisited, "Multimedia Tools and Applications, vol. 27, no. 3, pp. 393-410, 2005.

[3] S. Junuzovic, R. Hegde, Z. Zhang, P. Chou, Z. Liu, and C. Zhang, "Requirements and Recommendations for an Enhanced Meeting Viewing Experience," Proc. ACM Int'l Conf. Multimedia, pp. 539-548, 2008.

[4] K. Otsuka, H. Sawada, and J. Yamato, "Automatic Inference of Cross-Modal Nonverbal Interactions in Multiparty Conversations,"Proc. Int'l Conf. Multimodal Interfaces (ICMI '07), pp. 255-262, 2007.

[5] J.M. DiMicco, K.J. Hollenbach, A. Pandolfo, and W. Bender, "The Impact of Increased Awareness while Face-to-Face," Human-Computer Interaction, vol. 22, no. 1, pp. 47-96, 2007.

[6] R. Bakeman and J.M. Gottman, Observing Interaction: An Introduction to Sequential Analysis. Cambridge Univ. Press, 1997.

[7] M.S. Magnusson, "Discovering Hidden Time Patterns in Behavior: T-Patterns and Their Detection," Behavior Research Methods, Instruments and Computers, vol. 32, no. 1, pp. 93-110, 2000.