

Interaction Model and Respect of Rules to Enhance Collaborative Brainstorming Results

Thierry Gidel^{*1}, Andrea Tucker², Shigeru Fujita³, Claude Moulin⁴, Kenji Sugawara³, Takuo Suganuma⁵, Yuki Kaeri⁶, Norio Shiratori⁷

¹Sorbonne Universités, Université de Technologie de Compiègne, COSTECH EA 2223, 60200, France

²Interuniversity Center for Research in Education, Université de Lille, CIREL EA 4354, France

³Faculty of Information and Computer Science, Chiba Institute of Technology, Narashino-shi, Chiba-ken, Japan

⁴Sorbonne Universités, Université de Technologie de Compiègne, JR Unit CNRS 7253, Heudiasyc, 60200, France

⁵Cyberscience Center, Tohoku University, Sendai, Miyagi, Japan

⁶Faculty of Media Studies, Department of Media Studies, Mejiro University, 4-31-1 Nakaochiai, Shinjuku-ku, Tokyo, Japan

⁷Research and Development Initiative, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan

ARTICLE INFO

Article history:

Received: 14 January, 2020

Accepted: 21 March, 2020

Online: 09 April, 2020

Keywords:

Creative problem solving

Interaction model

Multi-agent system

Multi-touch device

Methodological rules

ABSTRACT

This paper presents a collaborative interaction model (CIAO) and several behaviour rules that could enhance brainstorming results. The model is composed of five elements which may be used during brainstorming sessions, consisting of activities performed by participants which are characteristic of different modes of interaction. Some sequences of these interactions may be considered more or less adequate than others. In particular, in order to get better results during brainstorming activities participants must respect certain rules when they write their ideas and when they consider notes written by somebody else. We argue that a multi-agent system can recognize different interaction modes and verify the respect of these rules by analyzing videos and notes produced by the participants in real time. Such a system must be trained as a machine learning system before being used during actual meetings. This system can simplify the role of the meeting facilitator. It can send a summary of the meeting situation, such as the proportion of each mode of interaction and identify behaviors that may need to be addressed. We present how feedback could be sent individually or addressed to the entire team. We will begin by presenting the interaction model, then propose an automatic recognition of these modes from video recordings and log analysis. We also address the necessity of rules and the structure of a multi-agent system which is able to verify whether or not the participants are respecting them. Finally, we propose how experiments could show the level of acceptance of such a system by users. The goal of this prospective research is to define a non-intrusive system that can be used during brainstorming sessions, based on an interaction model to enhance the quality of meeting results.

1 Introduction

This paper is an extension of work originally presented in the 23rd IEEE International Conference on Computer Supported Cooperative Work in Design [1]. Several research projects have proposed various architectures for supporting activities, in particular creative problem solving, in the context of collaboration between teams situated in different locations. Their front-ends typically involve rich user

interfaces and some of them use large multi-touch surfaces[2]-[4]. The user interface of such tools should be synchronized between sites to ensure a continuity of experience. Their back-end is mainly based on a managing multi-agent system situated in the cloud that supports the functions of the applications (such as storing data and synchronization). To facilitate the collaboration, provide a sense of presence, and ensure awareness of other team members' actions and oral conversations, the sites must have specific equipment (cameras

*Corresponding Author, +33 3 44 23 52 10 & thierry.gidel@utc.fr

and microphones).

Figure 1 shows two people working during brainstorming activities around a large multi-touch table in France, at the Université de Technologie de Compiègne (UTC) while communicating with a remote team situated in Japan, at the Chiba Institute of Technology (CIT). In this image, we can see that people use virtual keyboards on the table for creating digital notes. At the same time, they can see the other team (on the right section of the board), from several angles. The team in France can also observe the similar multi-touch device used by the other team (left part of the board).

The initial stage of a project meeting is often brainstorming, which requires a facilitator to manage member contributions.

From a technological point of view, these types of collaboration environments reach a high level of quality that should allow people to obtain good results during their collaboration. However, it appears that is not always true due to some practical, methodological and behavioural reasons:

Practical reasons: people do not know how to use these systems and, for example, cannot find the right way to activate a menu, a virtual keyboard, etc. in order to perform a task. Beyond the software, the hardware may also be slow or not as responsive as the user might expect.

Methodological reasons: collaboration is a complex process with multiple steps, which are not necessarily completed in a specific order. Participants do not necessarily know which is the best step to choose, when to move to the next step or repeat previous one in order to make their collaborative activity successful.

Behavioural reasons: people do not necessarily know how to conduct themselves when they are in a group meeting. It is the role of the meeting facilitator to help, to encourage or to set limits to, for instance, ensure equal participation. Everybody should be able to participate and use competencies at their best. This can be particularly complex in multi-cultural teams with potential language barriers. Moreover, it is not always easy to find a meeting facilitator with enough expertise to manage collaborative sessions and produce interesting results.

Our long term objective is to bring an AI based support to facilitators dynamically in order to make the meeting management easier especially when several distant teams need to collaborate. We claim that such a support could be useful to tackle methodological and behavioural issues. The first step was to define an interaction model between users, presented in Section 3. The second step will be the real-time AI based analysis of all the digital footprints left by meeting participants. This will allow us to get as many indications as possible on what is happening during the meeting. Digital footprints must be understood globally: audio/video recording of sessions with several cameras and several points of view, event logs of people performing activities on the large multi-touch devices. Then, the digital footprint data needs to be analysed and we need to define how the results of this analysis are communicated to the meeting facilitator and participants.

In the following sections we present some behaviour rules concerning brainstorming session, the collaborative interaction model (CIAO), and the architecture of a framework able to deliver advice to a meeting facilitator. The last section describes how the facilitator could communicate tips and hints to meeting participants.

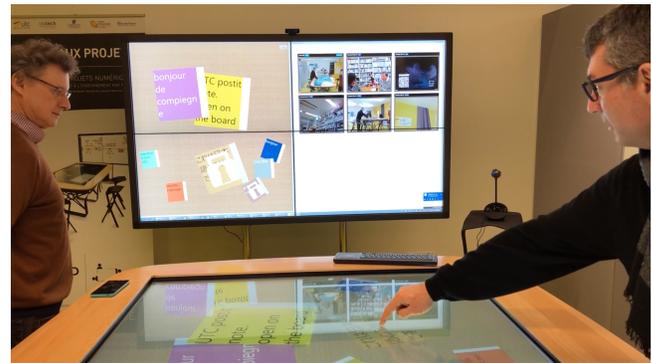


Figure 1: Environment with large multi-touch tabletop

2 Creative Problem Solving

2.1 Brainstorming

Creative problem solving is a key activity of teamwork that leads to novel ideas with great value [5]. Multi-touch and multi-user systems are expected to support this type of collaborative work by allowing cognitive stimulation through shared space [6, 7], common views, joint attention, awareness [8, 9] and social interactions[10].

Group brainstorming is one of the most popular creative problem solving methods [11] but it is often misused and the role of the group facilitator is often underestimated[12].

This seems especially critical when it comes to multi-disciplinary, multi-cultural distributed teams[13]. That is why we propose in this article to develop a complete system that could capitalize on the positive impact of multi-touch, multi-user devices based on cognitive simulation and social comparison processes[14, 15], simultaneous written and oral interactions [16, 17], AI and multi-agent based systems that would facilitate the team-work during creative problem solving and, more specifically, brainstorming sessions.

2.2 Rules

In order to encourage better results during brainstorming activities participants must follow some rules. Osborn[18] outlined four guidelines for brainstorming: (1) criticism is excluded, (2) free-wheeling and wild ideas are encouraged, (3) the greater the number of ideas, the greater the likelihood of useful ideas, (4) combination and building on others' ideas is welcome.

2.3 Barriers and propositions for effective implementation

Based on an extensive literature review, the main barriers to effective brainstorming are summarized by Isaksen [12] as (1) judgments during generation, (2) members giving up on the group and (3) an inadequate structure of interaction. In section 3 we first propose a model of interaction between users during general meetings and then we show how to analyze people's behaviour according to this model. Currently, analysis is done manually, looking at video recordings. A first objective of our research is to enable the automatic analysis of

videos, based on deep learning techniques. A second objective is to define acceptable and unacceptable structures of interaction during meetings.

To avoid judgment, which leads to inhibition due to fear of critical evaluation, voice and video analysis will track those attitudes and generate appropriate responses, either directed to one person (in a semi-private way), to a sub group (a local team on a specific site) or to the full team (simultaneously on all remote sites).

In order to avoid cognitive inertia or uniformity, wild ideas could be inserted within the group production by a computer support system in such a way as to be unable to identify the idea's origin.

One of the reasons for members giving up on the group is social loafing or freeriding. To avoid this, the setting will allow participants to view, through peripheral vision, others' production of ideas. This use of a matching of effort effect is supposed to work if there is a good group dynamic. To initiate this dynamic, the system could set challenges to stimulate the global team or the local sub-teams in terms of numbers of ideas generated, by displaying and announcing when each sub-group has produced ten more ideas. This could be associated with encouragement to extend their effort, which is known to lead to increased participation[19]. This kind of tool would require an experienced facilitator to avoid potentially harmful competition between sub-groups, or even encourage a team vs. computer competition.

System architecture and front-end software are designed to avoid inadequate interaction with devices. However, the way to favour adequate interaction between people, is still a meeting management issue. We propose a first step towards an automatic recognition of adequate and inadequate sequences of interaction between people during meetings thanks to video analysis.

The fact that all participants can write their ideas simultaneously will help to avoid production blocking [20] and limited airtime issues that are a common problem in oral brainstorming, where one person talks at a time. Simultaneous processing could even be encouraged by the challenges proposed previously. As it is true that we often think of more ideas than we could write down, a particular attention should be granted to text entry interaction: keyboard, auto-completion, word suggestion, spelling and grammar check, etc. to speed up the writing process.

We propose to develop a system that handles those major problems by helping the facilitator in his/her managing role, not aiming to replace this role. It must be seen as a critical added value. Nevertheless, the group facilitator has no ubiquitous power and he/she can only be in one location. Therefore, the system should help complete repetitive tasks and act as a relay in each sub-group. The system's contribution would include reinforcing the guidelines (follow the rules) and encouraging the participants. The system could also have a more direct interaction with the facilitator, as a timekeeper, suggesting the right time to shift from the production to the sorting phase, for instance, to maintain the group energy and productivity. That is the objective we would like to pursue.

2.4 *Brainstorming process and rules*

Brainstorming is a phase where people try to express ideas about the topic of a new project. Everyone is free to propose new ideas

and there should be no limitations.

Generally people write ideas on post-it notes. In previous papers [2, 3] we showed that virtual post-it notes organized on large multi-touch devices present advantages over paper notes. That defines the context of our research: brainstorming activities using large multi-touch devices (tables and boards) during meetings recorded with video cameras. To respect privacy, videos are used for real time analysis as explained in the next sections. The research seeks to facilitate brainstorming by making these rules explicit.

As people are not always aware of the brainstorming rules, an automatic system supporting the respect of these social and process rules must be well integrated and as symbiotic as possible. Because it is difficult for the facilitator or group leader of a remote brainstorming session to ensure that the rules are followed in all sites, having a system supporting the animation role would be of great help. We define two types of rules: social/behavioural rules and process/note content rules.

2.5 *Rules regarding behavior*

During brainstorming activities, people are free to propose ideas. Nobody should feel constrained because it leads to a hesitation to express ideas. Ideas are to be analyzed in a later step, which will weed out those ideas that do not correspond to the topics developed in the project. That means that everybody must respect any proposal and not be critical of any idea during the brainstorming phase. There are at least three ways to criticize: by voice (e.g. someone saying "this is not compatible with our project"), by gesture (e.g. somebody raises an arm or turns the head or makes disapproving facial expressions) and deleting someone else's idea before discussion.

2.6 *Rules regarding note content*

Only the author of a note is really aware of its meaning. However, somebody else reading a note must make sense of it. So, the way to write a note must not be ambiguous. Some syntactic rules, even if they are not extensive, may avoid some misunderstandings. Notes must not contain only a few words, but complete sentences. It is a good way to ensure that people have really expressed their ideas in a way that is comprehensible for other team members. Following this rule may also help people to fully develop an idea before proposing it.

2.7 *Other phases of project management*

Brainstorming is usually only one step of the project or problem solving process. Ideas produced during brainstorming activities need to be organized, filtered, and sometimes repeated if the group feels that there are not enough ideas to move the project forward. For that reason, the system analyzing the respect of syntactic rules may be used for other, more semantic tasks. It should be possible for this system to recognize similar ideas, new ideas and bring attention to them. New ideas are ideas where the content has not be reused in other notes.

3 Collaborative Interaction Model

3.1 Analysis Model

Behaviors that can be observed during globally collaborative work include cooperation, individual work, presentation of that work etc. As such, we propose five modes of interaction which break globally collaborative work into its finer details: individual work, communication, coordination, cooperation and collaboration[21]. We define each of these modes as:

- **Individual Work:** moments when individuals retreat from the group in order to reflect and construct their ideas [22], as well as work performed on tasks with which they were entrusted by the group.
- **Communication:** this mode of interaction allows individuals to introduce new information into the group, creating the point of departure for a shared vision [22]. Communication can take the form of providing information orally, presentations or adding written notes into shared spaces.
- **Coordination:** denotes the organization of activities (events, behaviors and actions) that structure and organize tasks in order to facilitate cooperative work [23].
- **Cooperation:** is produced following individual work, often preceded by the division of tasks amongst group members [24]. It appears as the results are put back together. This pooling of individual work necessitates negotiation to synchronize each actors representations.
- **Collaboration:** designates the co-elaboration, co-evolution, or co-construction of tasks and ideas by participants in order to reach a common goal [23, 22]. The most fundamental difference between collaboration and cooperation relates to how the production is constructed: together (in the case of collaboration), to the point that it is difficult to determine who contributed what; separately (in the case of cooperation).

These modes of interactions are mobilized, in a non-linear manner, by participants during work sessions and over the course of long-term projects. The project environment, methods and tools used, intervention or instructions given by the moderator, influence the behavior of participants and as such, the emergence of these modes of interaction.

We go on to present the targeted production of the different modes as well as some elements regarding the complexity of the interactions that make them up.

3.2 Targeted Production

Targeted production leave traces either on tools or meeting video recordings. In this section we present the elements we can detect in the video.

- **Individual Work:** individual work is shown through reflection, aiming at the construction of ideas and meaning with the goal

of eventually re-introducing the elements they judge as useful for the group. Producing written notes can be observed on video recordings and also on logs produced by the tools, because notes are written using virtual keyboards on multi-touch devices. Parsing these logs, we can also analyze the content of notes and detect those which do not respect predetermined guidelines or are interesting because the idea they contain has not yet been produced/shared with the group.

- **Communication:** Each member of the group is likely to have different results and ideas based on their individual work and experiences. Those ideas need to be introduced to allow new information into the group's discussions. This can be observed on recordings through speech, especially immediately following individual work or through actions on the device, such as sending a note from one surface to another (moving the idea from an individual work space to a collective one).
- **Coordination:** In collective sessions, coordination discussions allow for the definition of tasks and identification of responsibilities. This can be observed through the content of the conversation, related to organization and planning or through the type of tool being used, such as a GANTT. It is also typical to see a change in coordination following the intervention of a coach, as they often challenge the work that is being done.
- **Cooperation:** The division of tasks necessitates a pooling of work that is completed individually (or in sub-groups). This combining requires the establishment of consensus after having considered the information, opinions and arguments of each member. This can be detected based in the content of the discussion in audio recordings. Typically few notes are produced during this phase, but we see the modification of existing digital objects, instead.
- **Collaboration:** The group works together to co-produce a shared vision of concepts, of solutions, strategies, which is materialized through writing, models, reports or presentations. This, again, can be detected based on the content of the discussion in audio recordings.

The concept of globally collaborative work is introduced in [21] which examines some aspects of physical-digital workspaces, focusing on multi-user, multi-touch technologies and shows how different workspaces impact collaboration.

3.3 Manual Video Annotation

At the UTC, students have courses about project management. We installed a class environment in which there are five cubicles, each one containing a large multi-touch table and a large multi-touch board. We used software developed by the Ubikey start-up¹. During these management courses, students work on several case studies, performing activities mediated by these new tools and experimental working methods. For research purposes, the Ubikey software allows each action on these devices to produce a digital event that

¹Ubikey proposes a collaborative visual management tool consisting of two large multitouch surfaces (a table and a board) as well as a suite of applications which makes their use simple by supporting many kind of activities, including brainstorming.

could be tracked in near real-time on a log file. Each session produce hundreds of events registered in the log.

Each cubicle in the class environments also contain a system of four cameras able to record students' activities, behaviour and discussions. With the agreement of the students, we have recorded more than fifty hours of videos that we have manually analyzed. Figure 2 shows the view from the different cameras in one cubicle. For privacy reasons, we show the devices before the students enter the room.



Figure 2: Views from cameras for video recording

For analyzing the videos, we divided each recording into thirty seconds sequences and annotated them with one or more tags corresponding to the interaction modes described previously. This was a huge but very interesting work that consolidated our model. Figure 3 shows an excerpt of an annotation video file. For each sequence, we indicated which interaction mode appeared. The number simply represents the associated mode of interaction (8 for collaboration, 7 for cooperation, etc..) This system allowed us to avoid coding errors and graph the sequences more easily, demonstrating when and how these modes of interaction overlapped.

A	B	C	D	E	F
Time	Collaboration	Cooperation	Coordination	Communication	Individual Work
00:00:00					
00:00:30					4
00:01:00					4
00:01:30					
00:02:00					
00:02:30					
00:03:00					4
00:03:30				5	4
00:04:00			6	5	4
00:04:30				5	4
00:05:00	8		6		4
00:05:30				5	4
00:06:00					4
00:06:30					4
00:07:00	8			5	4
00:07:30	8				4
00:08:00			6	5	4
00:08:30		7		5	4
00:09:00					4
00:09:30		7		5	4
00:10:00		7			4
00:10:30		7			4

Figure 3: Excerpt of an Excel sheet containing video annotation

3.4 Automatic Video Annotation

We claim that an AI based automatic video annotation, together with a log event analysis could automate the video annotation process. Machine learning or inductive learning is based on software that learns from previous experiences and builds a model for predicting intelligent results for newly fed inputs (see figure 4). Such a computer program improves performances as more and more examples are available. In our case, we benefit of a large video corpus manually annotated. We think we can feed enough data to this machinery software for it to learn patterns, with the final output being the annotations according to the five annotation criteria defined in section 3.

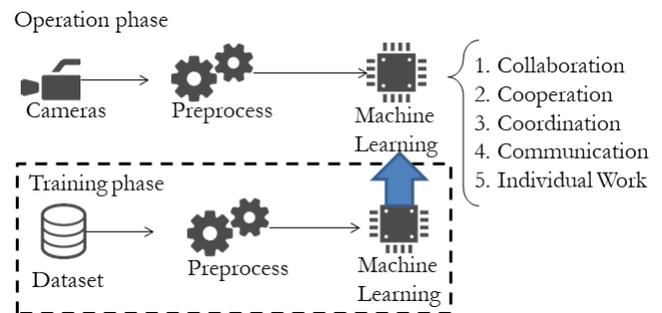


Figure 4: Video analysis, general training and prediction pipeline

Our first experiments were focused on using the already annotated video recordings to build and train a deep-learning action recognition model capable of directly recognizing each of these modes of interaction. We used the videos which were already manually annotated as training data. As of today, seven meetings have been manually annotated. Of those seven meetings, two also have the digital footprint logs associated with the devices used, and two more have been captured using four different points of view. We decided to discard the two videos taken from the ceiling viewpoint as they were too different from the rest of the dataset. This left a training dataset of around 8 GB and 12 hours. As the annotations were taken on 30 seconds intervals, we further cut the each video into the corresponding 30 seconds segments witch left us with 1473 annotated videos of 30 seconds. We finally doubled this number adding left-right flipped copies of each for a final training dataset of 2946 data points.

We based our work on the Inflated Inception model² described in [25] and used the TensorFlow framework³. This network has already been successfully used by their authors in [26] with weights trained on the Kinetics dataset [27] to improve the SOTA predictions on another action dataset performed by Google and described in [28] through transfer learning. In order to accommodate the weights, each video was first cropped to a frame size of 224 x 224 px and then had their RGB values normalized in [-1,1]. Our trials were run on a Tensorflow docker container with access to 64 Gb of RAM, Intel(R) Xeon(R) CPU E3-1230 v6 @ 3.50GHz and an Nvidia 2080 Ti with 11GB.

However, our first experiments met with some difficulties. Video sequences we annotated have a duration of thirty seconds, whereas

²See the following repository for the Keras implementation of this model we used: <https://github.com/dlpcb/keras-kinetics-i3d>

³TensorFlow initially developed by Google researchers is now an open source tool available in several programming languages. See <https://www.tensorflow.org/>

current action recognition datasets and models focus on shorter video sequences. Moreover, those datasets, and the models built upon them, focus on the recognition of atomic actions which are on a different semantic level versus the annotations we are trying to obtain. This implies the need for an intermediate semantic step in our models' architecture and the partial if not total retraining of said models. Finally, another difficulty we met was the relative limitations of our computing capabilities : the original network was trained on 500,000 iterations using 16 V100 GPUs each with at least 16GB of memory. This considerably slowed our trials with added memory and performance management, the original I3D in itself occupying 10 GB of video memory when accounting for the space needed for the input data. This is by no means a blocking factor, but hardware capabilities should not be ignored in future experimentations.

Moreover, the reason why humans choose a category when annotating is often not only based on video element recognition like a hand gesture or an interaction with an object but may also be based on vocal interactions and on their activity on the multi-touch surfaces. This led us to surmise the need of a multi modal approach combining video, audio and digital footprints. We especially believe our model will benefit from the event logs devices and thus, we plan to focus future experiments on leveraging this wealth of data to augment the video approach.

4 A Support Mechanism for Remote Brainstorming Analysis

In this section we study the questions presented above from a technological point of view and discuss the characteristics that an automatic (coaching) system must master to support brainstorming and the different tasks necessary to detect when a rule is not followed. We explain how this can be done by a multi-agent system. Each agent can detect a default and then produce an alert when something wrong appears. However some regulation is necessary in order to propose efficient feedback to participants. Also, when rules are correctly followed and the team is producing good results, we detail how this could be encouraged.

4.1 Architecture

We are able to make use of existing system architectures, which have been designed to support team activities (managing support system). Figure 5 shows a simplified architecture of a management system (left) and the added coaching system we plan to develop (right). Notes and video streams are sent by the local devices to the cloud managing system, where they are treated by its agents and dispatched to the other sites.

We will design another multi-agent system (coaching support system) which can be installed on the same cloud server that receives the streams from the different sites thanks to the managing support system agents; it will produce feedback that can be dispatched to the sites by the managing system agents. The new system can receive the streams because it can request them from the managing agents. It will be necessary to develop a new interface to provide feedback to each site, but, from a technical point of view, the work required

to design and implement it is negligible. A multi-agent system is suitable for developing a heterogeneous distributed system. Our multi-agent system is in compliance with the FIPA architecture and uses the FIPA-ACL communication protocol. Agents can make decisions according to a rule-based inference system. Rule descriptions are declarative and rules are associated with a dynamic priority level. Agents run procedures associated with the rules that are chosen. No machine learning method is implemented here.

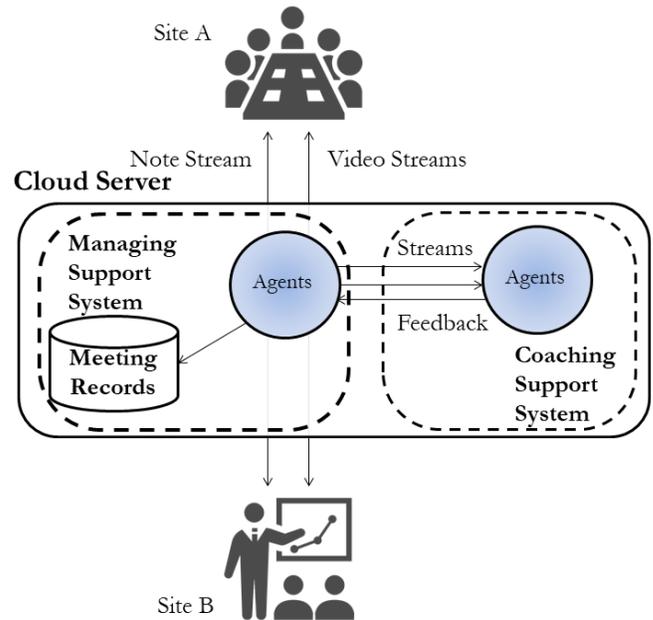


Figure 5: Multi-agent systems managing collaboration between remote teams

We believe that the coaching system must consist of two layers: a detection layer and a regulation layer. In the first layer, the agents detect the activities done by each person (production of an idea, criticism, facial expression...) and produces alerts after each detection. Other agents track the number of ideas produced during a session and produces alerts based on a comparison between sub-group metrics (i.e. returning to the notion of competition), standard production metrics (i.e. the systems default settings for what good production looks like), and on group goals and/or the specificity of the group according to the facilitator.

In the second layer, regulation agents analyze the different alerts (positive and negative) coming from the detection layer agents. According to pre-defined rules, these agents decide what feedback should be given, when and to whom (see figure 6).

4.2 Identification of participants

It is necessary to identify all the participants in a meeting. An identifier is sufficient, personal data should not be required. The system must give feedback either to one participant, to a sub-group at one specific site or to all the participants at all sites. It is when a message has to be sent to one individual or to identify the source of a particular idea that the identifier is necessary.

When using a large multi-touch surface, people use virtual keyboards to enter content into post-it notes. It is easy to add a unique

symbol on this keyboard and then a camera can associate a participant to a keyboard for identification. If necessary, facial recognition maybe used to add identifiers to participants. An piece of individual feedback, dedicated to a single participant, could be achieved thanks to an individual written notification, like a pop-up attached to the virtual keyboard he/she is using, making it discreet.

4.3 Behavioural Rules

According to the examples we proposed in the previous section, there are three ways to express disapprobation of an idea: by voice, by gesture or by deleting a note. Because of this, we believe that the installation must include different types of devices: i) cameras for video streams transfer, ii) microphones for better capturing individual voices and iii) multi-touch screens where the activity is performed. Several video streams are transferred to agents belonging to the detection layer. Agents use their own algorithm to analyze either gestures, voice or faces. Representations of notes are also transferred to detection layer agents.

4.4 Syntactic Rules

A note contains a text which expresses an idea about the brainstorming subject. The rule concerning an idea specifies that the text must be a complete and expressive phrase. This avoids any misunderstanding and ambiguity. The analysis of a note is described in section 4.6.

In the detection layer an agent is associated with each participant (i.e. with each identifier). It receives structures representing notes containing their content, the participants' identifiers and other meta-data (timestamp, etc.). The natural language processing support of each agent is able to detect acceptable notes and refuse others.

The system uses different kinds of rules. Rules will be checked through experimentation and inconsistencies, then adjusted manually.

manage syntactic alerts. All the alerts produced during the activities are compiled in shared memory (see figure 7). Globally, the regulation agents have different tasks: they determine the given situation of the meeting: is it the beginning, middle or end of the session?; they count the number of positive alerts (rules followed, production rate) and negative alerts (rules not followed); they check the production rate of each participant (a participant is considered inactive when the number of notes is low); they propose feedback.

A feedback agent is in charge of ordering and producing definitive feedback to participants. Feedback agreement from participants (a record of whether or not the feedback is acceptable or helpful) are also stored in the shared memory (however their exploitation will be performed in a future research).

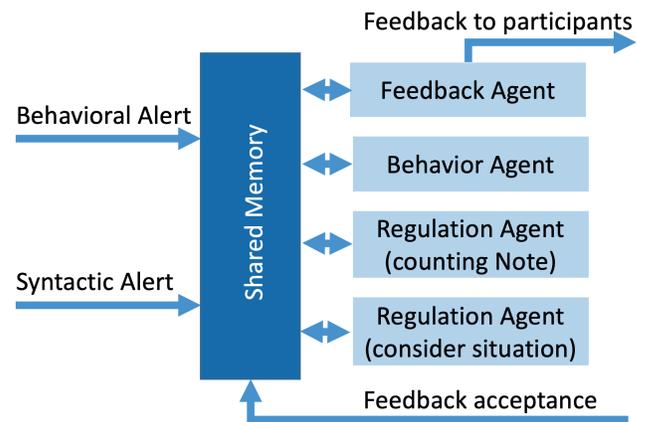


Figure 7: Regulation layer

4.6 Design of Text Processing System

This subsection describes the life cycle of a note within the agent layers (see figure 8). The note counting mechanism counts each participant's note. The number of notes may launch feedback that is designed to reinforce his/her enthusiasm.

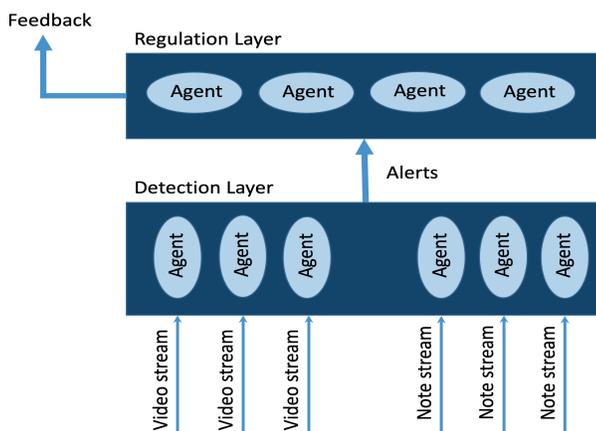


Figure 6: Multi-agents system layers

4.5 Regulation System

The regulation agents consist of two main categories: behaviour agents that manage behavioural alerts and regulation agents that

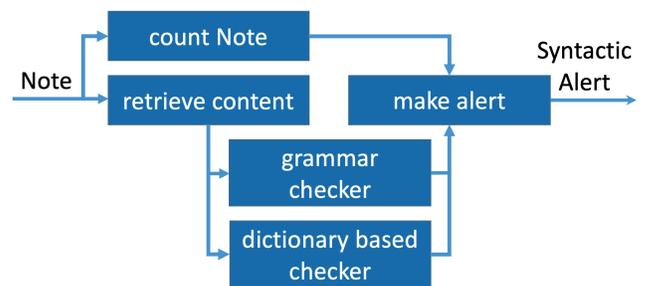


Figure 8: Text checker on note

The text of a note is checked by a dictionary and grammar checker. The dictionary based checker verifies if words are contained in predefined dictionaries and associates a value to each word. The grammar checker verifies syntactically the text of each note using a tool such as the LanguageTools. An alert is launched if this global evaluation is not high enough.

4.7 Design of Image and Voice Processing System

In this subsection we detail the main components of a checking mechanism for a video stream (see Figure9): The facial recognition mechanism identifies a participant; the gesture recognition mechanism categorizes a participant's actions; the emotion recognition mechanism associates a participant's face to a class of emotions; finally the tone of voice recognition mechanism associates a participant with a class of emotions. The voice recognition mechanism achieves a speech to text transformation.

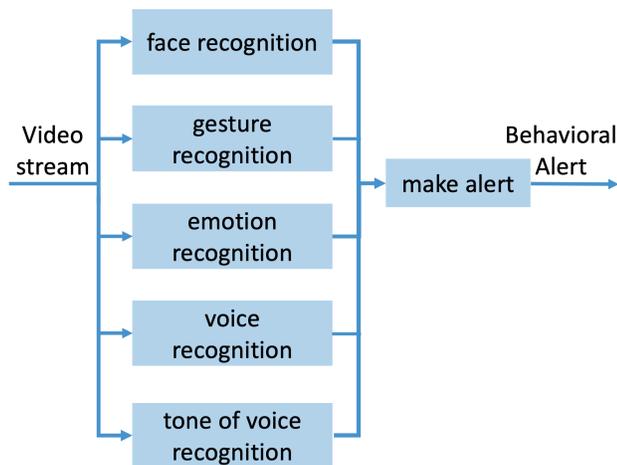


Figure 9: Video checker

These recognition systems require large data sets to learn to categorize faces, emotions and gestures. Alerts produced by this mechanism will have a relative score which indicates the level of reliability of the identification.

Different tools can be used for these mechanisms. Open source projects OpenCV[29] with YOLO[30] are suitable for identifying a participant. The OpenPose tool[31, 32] can detect human body, hands, facial and foot key-points on single images; the Microsoft Emotion engine[33] recognizes emotion categories from Kinect sensors.

Of course, non-verbal communication depends largely on participants' language, nation, region, age or personal experience and its analysis requires a good deal of testing.

5 Feedback For Rule Respect

In this section we detail the way an automatic system could give feedback in real time to the participants of a brainstorming meeting about the respect of rules. The feedback is based on the different types of rules we identified in previous sections.

The main questions are when, how, to whom to deliver the feedback and if hints are necessary. Another question is how to avoid interfering with the meeting by being as discreet as possible, while also being sensitive to participants' emotions to help create a safe psychological space.

Both individual and generic messages have to be integrated in the coaching system. Individual notifications are dedicated to a specific person, for instance because he/she does not follow one

rule consistently. This means that the system is able to refer to and recognize all the participants of a meeting. This could be achieved through an individual written notification, like a pop-up attached to a keyboard (see section 4.2).

Common notifications can be sent to a group of people, for instance, a local team (i.e. a distant team would not receive it). If the regulation layer of the coaching system considers that the same rule is not followed by several people in the same place, a local verbal or written message could be sent to the local team. A verbal message could be heard by the other team, so if we want to avoid this possibility, it is better to choose to send a message written on the large multi-touch surface.

Social interactions are important in brainstorming activities. The notifications should not divert participants from their work. Therefore, the number of notifications should be adjusted to the situation and should take into account the group dynamic and notification content.

The balance between feedback to motivate or correct behaviors should be taken into account. This means positive comments such as "good job, you are producing lots of ideas, keep it up!" are as important as constructive feedback such as "be careful, you should not criticize other people's ideas".

Finally, specificity of the team should be taken into account in the way the feedback is delivered. For instance, cultural bias could lead to different ways of producing and interpreting feedback.

6 Evaluation Scenario

In this section we explain how we will test the coaching system. It is necessary to evaluate the efficiency of the system at both a technical and acceptance level.

The first step is to test the quality of the coaching system at the detection layer with the production of alerts. Different scenarios creating notes automatically will be played and statistics about the number of alerts delivered by the system analyzed. Video with specific gestures and faces will be created and proposed to the agents. When the rate of recognition is satisfactory we can begin the second step of our evaluation, i.e. the test of the regulation layer.

Based on pre-recorded videos and notes from a session, we determine what feedback individuals and the team should receive. It is a test relying on subjective judgments made by researchers, but based on the rules outlined previously. Researchers will manually balance individual and global feedback delivery for timeliness and content. When a satisfactory rate of feedback is delivered by the coaching system we can begin the third step of the evaluation, i.e. experiments with real scenarios and real participants playing foreseen sequences.

The last step of the experiment will focus on the way people accept to respect constraints during brainstorming activities and receive feedback during their work. Prior to the experiment, participants will receive instructions about the rules to be followed during the session in order to make the conditions explicit. We will conduct experiments in two ways: i) brainstorming activities managed only by a human in charge of the facilitation to ensure the respect of the rules, ii) brainstorming activities where an automatic system gives feedback together with a human facilitator.

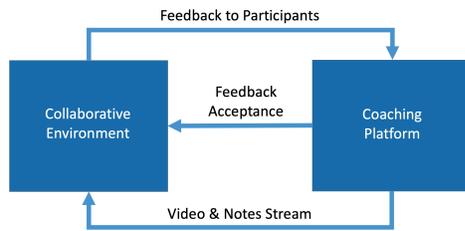


Figure 10: Exchanges between systems

The results of the last experiment will be evaluated by observers who monitor how people react when they receive feedback. We postulate that the user interface may allow participants to react when they are notified, with a very quick action such as pressing a button on the notification itself. In this way, we can consider a loop where the coaching system registers the level of feedback acceptance by the participants (see figure 10).

Our objective is not to cancel out the role of the facilitator because integrating a coaching system into pre-existing support system architecture can be difficult, but also because the role of the facilitator is fundamental. The coaching system is only an assistant which allows the facilitators to devote time to their main tasks, such as managing the group dynamic.

7 Research limitations

This position paper defines our research goals, details our propositions on how to achieve this goal and describes our three step research protocol for evaluation. As presented, we are already facing difficult technical challenges that we need to overcome in order to carry out this research, such as automatic video annotation issues. We also foresee some tricky problem to be solved with regards to fine tuning the rules implemented in the regulation system. Finally, once the technical issues are solved, acceptance of this type of system by both the participants and the coach needs to be thoroughly evaluated.

8 Conclusion

In this paper, we have presented how brainstorming results could be enhanced if meeting participants observed several simple practical, methodological and behavioural rules. We have proposed a model of interactions between people during collaborative meetings. We have described how an AI based analysis system could both summarize the state of a meeting according to this model and give advice to a meeting facilitator and the participants.

We hypothesise that a technological support system could help meeting facilitators with their group management tasks if it is able to analyze whether or not behavioural and methodological rules are followed and if it produces efficient and adapted feedback.

We proposed an architecture for such a system and explained the way we can conduct the necessary experiments for testing its implementation and the way people receive and agree with the feedback.

In the context several teams situated in different locations, this model and associated architecture could help supporting activities

and in particular creative problem solving. It therefore could be integrated into existing computer supported collaboration tools along with video-conferencing systems. In a future work, we would like to implement new technical functionalities such as an automatic learning system based on analysis of previously recorded sessions. Additionally, we would like to study the retroactive feedback acceptance by participants on the system itself. For instance, an agent would be able to understand and adjust its behavior in order to achieve the highest impact on the targeted participant. This means that the agent would need to learn from participant feedback acceptance and choose the appropriate way to give feedback to each participant.

References

- [1] T. Gidel, S. Fujita, C. Moulin, K. Sugawara, T. Sukanuma, Y. Kaeri, and N. Shiraori. Enforcing methodological rules during collaborative brainstorming to enhance results. In *2019 IEEE 23rd International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pages 356–361, 2019.
- [2] Y. Kaeri, K. Sugawara, C. Moulin, and T. Gidel. Agent-based design of iot applications for remote brainstorming support. *2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pages 820–825, 2018.
- [3] Y. Kaeri, C. Moulin, K. Sugawara, and Y. Manabe. Agent-based system architecture supporting remote collaboration via an internet of multimedia things approach. *IEEE Access*, 6:17067–17079, 2018.
- [4] Y. Kaeri, K. Sugawara, C. Moulin, and T. Gidel. Agent-based management of support systems for distributed brainstorming. *Advanced Engineering Informatics*, 44:101050, 2020.
- [5] *The International Handbook of Creativity*. Cambridge University Press, 2006.
- [6] P. Marshall, R. Morris, Y. Rogers, S. Kreitmayer, and M. Davies. Rethinking ‘multi-user’: An in-the-wild study of how groups approach a walk-up-and-use tabletop interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’11*, pages 3033–3042, New York, NY, USA, 2011. ACM.
- [7] Y. Rogers, Y. kyung Lim, W. R. H. PhD, and P. Marshall. Equal opportunities: Do shareable interfaces promote more group participation than single user displays? *Human Computer Interaction*, 24(1-2):79–116, 2009.
- [8] R. Fleck, Y. Rogers, N. Yuill, P. Marshall, A. Carr, J. Rick, and V. Bonnett. Actions speak loudly with words: unpacking collaboration around the table. In *Proceedings of the ACM international conference on interactive tabletops and surfaces*, pages 189–196. ACM, 2009.
- [9] S. Conversy, H. Gaspard-Boulin, S. Chatty, S. Valès, C. Dupré, and C. Ollagnon. Supporting air traffic control collaboration with a tabletop system. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work, CSCW ’11*, pages 425–434, New York, NY, USA, 2011. ACM.
- [10] O. Shaer, M. Strait, C. Valdes, T. Feng, M. Lintz, and H. Wang. Enhancing genomic learning through tabletop interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’11*, pages 2817–2826, New York, NY, USA, 2011. ACM.
- [11] K. L. Dugosh and P. B. Paulus. Cognitive and social comparison processes in brainstorming. *Journal of Experimental Social Psychology*, 41(3):313 – 320, 2005.
- [12] S. G. Isaksen and J. P. Gaulin. A reexamination of brainstorming research: Implications for research and practice. *Gifted Child Quarterly*, 49(4):315–329, 2005.
- [13] N. Michinov and C. Primois. Improving productivity and creativity in online groups through social comparison process: New evidence for asynchronous electronic brainstorming. *Computers in Human Behavior*, 21(1):11 – 28, 2005.
- [14] B. A. Nijstad, W. Stroebe, and H. F. Lodewijckx. Cognitive stimulation and interference in groups: Exposure effects in an idea generation task. *Journal of Experimental Social Psychology*, 38(6):535 – 544, 2002.
- [15] K. Dugosh, P. Paulus, E. Roland, and H.-C. Yang. Cognitive stimulation in brainstorming. *Journal of Personality and Social Psychology*, 79(5):722–735, 2000.

- [16] B. A. Nijstad, W. Stroebe, and H. F. Lodewijkx. Production blocking and idea generation: Does blocking interfere with cognitive processes? *Journal of Experimental Social Psychology*, 39(6):531 – 548, 2003.
- [17] M. Diehl and W. Stroebe. Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of Personality and Social Psychology*, pages 497–509, 1987.
- [18] A. F. Osborn. *Applied imagination; principles and procedures of creative thinking*. Scribner, 1953.
- [19] D. Johnson and R. Johnson. *Learning together and alone: cooperation, competition, and individualization*. Prentice-Hall, 1975.
- [20] C. M. Hymes and G. M. Olson. Unblocking brainstorming through the use of a simple group editor. In *Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work, CSCW 92*, page 99106, New York, NY, USA, 1992. Association for Computing Machinery.
- [21] A. Tucker, T. Gidel, and C. Fluckiger. Designing physical-digital workspaces to support globally collaborative work. In *Proceedings of the Design Society: International Conference on Engineering Design*, pages 109–118. Cambridge University Press, 2019.
- [22] S. D. Teasley and J. Roschelle. Constructing a joint problem space: The computer as a tool for sharing knowledge. In S. P. L. . S. J. Derry, editor, *Computers as Cognitive Tools*, pages 229–258. Hillsdale, NJ: Lawrence Erlbaum Associates, 1993.
- [23] M. J. Baker. Collaboration in collaborative learning. *Interaction Studies*, 16(3):451–473, 2015.
- [24] K. A. Bruffee. Sharing our toys: Cooperative learning versus collaborative learning. *Change: The Magazine of Higher Learning*, 27(1):12–18, 1995.
- [25] J. Carreira and A. Zisserman. Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset. *arXiv:1705.07750 [cs]*, Feb. 2018.
- [26] R. Girdhar, J. Carreira, C. Doersch, and A. Zisserman. A Better Baseline for AVA. *arXiv:1807.10066 [cs]*, July 2018.
- [27] W. Kay, J. Carreira, K. Simonyan, B. Zhang, C. Hillier, S. Vijayanarasimhan, F. Viola, T. Green, T. Back, P. Natsev, M. Suleyman, and A. Zisserman. The Kinetics Human Action Video Dataset. *arXiv:1705.06950 [cs]*, May 2017.
- [28] C. Gu, C. Sun, D. A. Ross, C. Vondrick, C. Pantofaru, Y. Li, S. Vijayanarasimhan, G. Toderici, S. Ricco, R. Sukthankar, C. Schmid, and J. Malik. AVA: A Video Dataset of Spatio-temporally Localized Atomic Visual Actions. *arXiv:1705.08421 [cs]*, May 2017.
- [29] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.
- [30] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 779–788, June 2016.
- [31] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *CVPR*, 2017.
- [32] T. Simon, H. Joo, I. Matthews, and Y. Sheikh. Hand keypoint detection in single images using multiview bootstrapping. In *CVPR*, 2017.
- [33] Ms emotion engine. <https://azure.microsoft.com/en-gb/services/cognitive-services/emotion/>.