



CAPTURING SYNCHRONOUS COLLABORATIVE DESIGN ACTIVITIES: A STATE-OF-THE-ART TECHNOLOGY REVIEW

M. Hall, C. A. McMahon, P. Bermell-Garcia, A. Johansson and R. Ravindranath

Abstract

This paper provides a technology-focused state-of-the-art review applied to capturing, processing and reviewing collaborative design activities. It presents a descriptive study (DS-I), which explores the landscape of relevant technologies and future trends that can be applied to improve the situation within engineering organisations.

Keywords: knowledge management, design informatics, technology development

1. Introduction

Engineering organisations invest significant time and resources into design work. Collaborative design activities are described as a 'knowledge process in which rational action is taken by a design agent to achieve a design goal' (Sim and Duffy, 2003). Valuable information and knowledge that result from these activities are often lost or are challenging to retrieve. Capturing information and knowledge from collaborative design activities has the potential to 'aid the design process and through-life support of the developed systems' (McMahon et al., 2005; Ball et al., 2006). It is within the context of collaborative, synchronous and critical design situations that this review was focused on, capturing design activities at the micro-scale (Vergison, 2001) on the project level (Gasik, 2011).

Conway et al. undertook a series of case studies and test scenarios related to capturing information and knowledge that was generated during engineering design meetings (Conway et al., 2007). They claimed that throughout the design process, a wealth of information is generated that can be extremely useful at a later instance for design re-use and maintenance. They highlighted the issue of such important design information either being lost, or requiring considerable effort from the designers to document, with little incentive to do so. They presented the challenge that capturing this information and knowledge for future re-use is notoriously difficult, stating that any attempts should be simple, efficient and remain unobtrusive to design engineers. There was significant work around 1998 to 2008 that focused on 'aiding the capture of information within collaborative situations' (Conway et al., 2008). However, the development of new technologies from a range of fields in recent years presents new opportunities to apply in addressing this issue. The purpose of undertaking this technology review was to understand existing and emerging technologies and how they can be applied to this challenge both now and in the future. This review is not intended to be an in-depth analysis for each technology, but instead a synthesis of several key technology areas that may contribute to improving the capture of information and knowledge within collaborative design activities. The review focused on synchronous collaboration involving both face-to-face meetings and remote interactions, and does not include asynchronous activities which were considered out of scope.

This paper firstly presents an overview of the relevant literature, followed by a review of relevant key technologies. Finally, conclusions are drawn from this review to steer further development in this area.

2. Literature review

There are a variety of techniques presented within the literature in which collaborative design activities have been captured. These methods range from informal records such as engineer's notes in a logbook (McAlpine et al., 2006) to rich and detailed multi-media recordings (Torlind et al., 2009). Giess et al. observed that the capture of this type of work can be broadly classified into the following two approaches (Giess et al., 2008): (i) 'Record everything and analyse later' approach, which results in masses of information being captured, much of which can be regarded as 'noise', and requires significant time for post-processing, and (ii) 'Analyse and document at the time' approach, which documents the decisions made and the issues resolved using structured arguments. This does not record much of the dialogue associated with the decisions, resulting in lost information and knowledge.

The challenge with the latter approach is that the effort is emphasised during the work itself, resulting in an additional overhead to the designers. Conversely, the challenge with the former is that the records require sense-making after the event with limited context.

A widely-used method of capturing activities within industry is meeting minutes, which follows the second approach. The observation from a study conducted by Huet (2006) within the context of engineering design in the aerospace industry, suggests that actions and decisions in meetings are usually well recorded but rationale and experience leading to these are not generally captured well in the minutes (Huet et al., 2007). Meeting minutes, when used to formally document the record of meetings are often limited in the extent to which they capture the information exchange that took place. An evaluation of written descriptions of design activities was undertaken from an information loss perspective. Huet (2006) took a number of cases within Airbus and demonstrated a measurable information loss between meeting minutes and corresponding transcriptions. Similarly, McAlpine (2010) demonstrated significant information and knowledge loss from engineer's logbooks that typically does not get formally written or reported. This capture of collaborative design activities is captured in the form of personal notes, which follows the 'analyse and document at the time' approach. This form of information capture involves participants making notes in either electronic format or with paper and pen. The type of information contained within logbooks, tend to be informal and often semi-structured, as opposed to formal records such as written technical reports (McAlpine et al., 2006; McAlpine, 2010).

Certain types of design activities may lend themselves to other forms of capture such as facilitated group model building (Buckingham Shum, 1996; Bracewell et al., 2009) and dialogue mapping (Conklin, 2005). Such forms of capture often prescribe a graphical framework which structures the meeting with, or without a facilitator or a documenting observer. These methods follow the 'analyse and document at the time' approach where the information and knowledge is captured as part of the process, which is contained within the resulting knowledge representation. Research into the use and effectiveness of tools such as DRed (Bracewell et al., 2009), Compendium (Okada and Buckingham Shum, 2006) and Glyma (Culmsee and Awati, 2014) were undertaken within the industrial environment. Such tools provide active capture, supporting the design activity itself, not used to passively capture in the background.

There are many technology-supported methods that record meetings in a rich way, for example, Instrumented Meeting Rooms (IMR) (Hain et al., 2012) and Design Observatories (DO) (Carrizosa et al., 2002; Torlind et al., 2009), which adopt the 'record now and analyse later' approach. These record multi-party meetings, enabling the recall of important information such as decisions and key discussion points (Al-Hames et al., 2006). Technologies such as these capture the discussion (discourse) and other information for synchronous collaborative working. Some systems such as those developed by the AMIDA consortium (Hain et al., 2012) focus on generating automated transcriptions, where the focus is on 'who said what and when', with detailed rich capture of the discourse with supporting video and audio. Whilst these technologies reduce, or even remove any information loss from the collaborative discussion, they introduce issues of information retrieval and accessibility for later re-use.

It is evident within the literature that there are a number of branches of capture methods, each with advantages and disadvantages. If more advanced methods are to be developed to improve the capture of essential design knowledge, then a review of key technologies that may contribute to a solution is important.

3. Technology review

In order to understand what existing technologies may offer, the following capabilities were derived from the literature and organised according to the level within which they sit in the process of capturing, processing and retrieving collaborative design activities. The relationship between each of these is illustrated in Figure 1, and explained below:

- Capture – the generation and storage of records resulting from collaborative design activities and designer interaction
- Analyse Content & Context – the analysis of the records resulting from design activities and designer interaction
- Analyse Human Behaviour & Meta-Information – the analysis of human behaviour from records of design activities and associated meta information
- Enable Review & Retrieval (Post-activity) – the means of presenting data and information related to the design activity and the ability to retrieve after the design activity
- Provide Assistance & Support (Real-time) – the means of supporting the designer in undertaking a design activity
- Enable Collaboration – the means of enabling interactions between designers and co-ordinating design activities

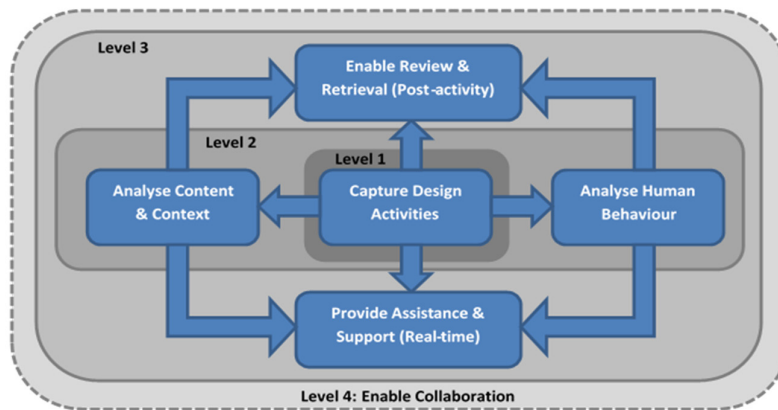


Figure 1. Four levels of capabilities

3.1. Devices

3.1.1. Visual - camera (Level 1)

Camera technologies used to capture collaborative design activities can be broadly categorised by quality, cost and their field of view. There are a variety of 360 degree cameras available that can be positioned at the centre of a table and used for meeting capture (Better, 2016). These cameras range from low-cost options with relatively low-quality imaging, to high-cost options offering high definition image capture. Some cameras comprise of multiple lenses within an enclosed body, such as the Nikon KeyMission. Whilst others are multiple cameras set up in an array, such as the GoPro Odyssey. The less expensive options often utilise a fisheye lens to capture a larger angle of view with fewer lenses, but typically compromise image quality. The higher quality 360 degree spherical professional production cameras, such as the Nokia OZO, are designed for very high quality replay in immersive virtual reality, but come with a very high price tag. However, there are now relatively inexpensive 360 degree cameras becoming more readily available, providing high quality recordings with depth perception from camera pair stereoscopic outputs. For example, the Insta360 and Samsung Gear 360 cameras are pocket-sized and can capture 4K video in 3D and 360 degrees. Whilst camera hardware is improving, key improvements can also be seen in the image processing, moving toward high-quality real-time image stitching for live streaming (Rastogi et al., 2017).

Other more general purpose devices can also be used to capture meetings such as webcams, smartphones or devices such as the Xbox Kinect (Zhang, 2012). These generally have a relatively narrow field of view and need to be mounted and positioned carefully. There are also cameras designed specifically for video-conferencing. These cameras typically have a narrow field of view compared to the 360 degree cameras, but have pan, tilt and zoom features. These cameras are usually built into a video conference suite such as those provided by Polycom, Cisco and Logitech.

In general, camera technology to capture collaborative engineering design activities is very mature for industrial use. The trend from the past few years has shown camera technology to be more affordable and smaller, with higher quality capture offering 360 degree, depth perception and capture for virtual reality applications. In many engineering design meetings, relatively low quality and inexpensive cameras may be sufficient to record design activities, where visual interaction between participants is low. However, in situations where interaction between individuals and digital artefacts is high such as in collaborative design, virtual reality design laboratories are showing how these technologies can improve the design process itself (University of Minnesota, 2017). So as high quality camera technologies become more affordable, this will provide opportunities for capturing a richer set of interactions for virtual reality in industrial design applications and not just a laboratory setting.

3.1.2. Audio - microphone (Level 1)

There are a variety of specialisations of microphones that are applicable in capturing engineering design activities. These can be desktop group-capture or personal recording devices, and can be broadly categorised according to their cost, quality and their direction of capture. In recent years, digital MEMS (micro electrical-mechanical system) microphones have become very popular since they are small, reliable and provide high sound quality. MEMS microphones are used in many mobile electronic devices and have been developed into microphone arrays (Zwyssig et al., 2012). Microphone arrays provide multiple microphones operating together to provide multi-channel signals. This enables acoustic source localisation and thus speaker diarisation. Recent popular microphone arrays that demonstrate this are the Amazon Echo and the Voice Tracker Array Microphone.

Microphones can also be attached to meeting participants such as wired clip-on options, for example Lavalier microphones which provide high quality, but can be quite cumbersome. Wireless alternatives tend to be more convenient, but generally provide lower quality, e.g. Bluetooth which has signal compression issues or WiFi wearable devices such as smartbadges, lanyards and headsets.

A great number of portable electronic devices are also available on the market that contain in-built microphones that can be used to record audio in meeting room environments. These include mobile devices such as smartphones, Dictaphones, tablet computers and webcams containing small in-built microphones. Devices such as smart pens enable audio recording, and also offer the possibility of automatically synchronising the audio to correspond to a user's pen strokes. These multi-functional portable devices tend to be non-specialist and useful for low quality personal records of a discussion. However, for group-use they typically offer a lower sound quality, and tend to be more susceptible to noise compared with dedicated group-capture devices.

Whilst there are continuous improvements to audio sensor hardware, there are also advances in audio signal processing. Sound quality enhancements are continuously being made to improve the signal-to-noise ratio (SNR). Such improvements are made through noise reduction and noise cancelling signal processing techniques. These issues can be particularly prevalent within a meeting room where ambient noise and reflected signals from the acoustic space provide a challenge to capture audio in sufficient quality.

In general, microphone arrays are best suited to capturing design-based meetings and collaborative working situations since they enable source localisation for speaker diarisation. They also provide a good quality signal whether from raw recording, from multi-channel or post-processing. Microphone and audio signal processing are relatively mature for an industrial setting, but they continue to become smaller, cheaper and provide better quality audio. Any audio signals that are used for speech recognition or speech-to-text rely upon a good quality audio capture, so the quality of the microphone is important to minimise the word error rate (WER) in a use such as natural language processing.

3.1.3. *Mobile and smart devices (Level 1)*

Portable electronic devices such as smart phones and tablets typically have many sensors and capabilities that can be used to capture design activities. With many manufacturers providing open application programming interfaces (APIs), it has meant that there have been a vast number of software applications being developed for different device platforms. Many of these ‘apps’ utilise combinations of the many sensors that are in-built, which can be used to undertake tasks such as note-taking and audio and video recording. Smart pen technologies such as Livescribe allow handwritten notes and sketches to be made in the traditional way, which are then easily digitised and can be synchronised with audio recordings for playback. Examples of more advanced note-taking methods include, NiCEBook which is a paper notebook that supports the taking, structuring and re-use of notes. Other commercial tools such as Evernote provide users with a collaborative workspace that enable personal notes to be stored electronically and shared with others.

Wearable technologies such as smart badges are electronic devices that have in-built sensors such as microphones, proximity sensors and location tracking. In recent years, generally these technologies have become smaller which has made them less cumbersome to wear. Health monitoring applications have driven the wide adoption of smart watches and activity trackers which contain sensors to capture a wide variety of data including heart rates, GPS location, and accelerometers for motion tracking. Research into brain activity such as electroencephalogram (EEG) also demonstrates how understanding human behaviour, such as within engineering design activities, may be possible in the future. Although, this is currently a relatively immature technology for application into engineering design organisations.

Whilst many of these wearable devices may be for personal use, there is also an application for business use to understand human behaviour within design activities. A big driver for many of these sensor technologies is the growing capability to handle extremely large datasets to analyse the data and extract meaningful information. With advanced analytics technologies becoming more readily available, it allows large datasets that are generated from mobile devices to be aggregated and analysed to understand the behaviour of the users such as designers. Sociometrics (Olguin et al., 2009) is a discipline that collects sensor data during human interactions to undertake analysis on the social interactions. There are of course privacy and ethical implications to consider with any use of personal data collected. But commercial technologies typically provide a sufficient benefit to the user in using the technology that outweighs any perceived drawback in this regard.

In general, mobile electronic devices and wearable technologies are becoming more connected, less cumbersome and are providing more features that can be used to capture design activities. Whether it is through audio-visual capture of discussions, photographs of flipcharts, or capturing detailed data about the state of designers- the technology in this area is providing the means of capturing design activities that are more readily accessible than was previously possible.

3.2. Natural language processing (NLP)

3.2.1. *Speech-based processing (Levels 2-3)*

Automatic speech recognition (ASR) technologies enable the conversion of human speech into a sequence of text-based words. This can be used to simply convert speech to text (STT) for transcription or note-taking purposes, or it can be used to form the basis of a command for a computer system to act upon such as with intelligent personal assistants (Shah and Shah, 2015). Research in ASR originates from the 1950s and has improved significantly over the years. The basic process of STT firstly involves the speech acquisition through audio using a microphone device or array, followed by a series of steps to extract features and apply predictive models to determine the text output. The measure of accuracy for STT technologies is the word error rate (WER). Continuous and spontaneous natural speech where collaborative design activities usually occur is the most challenging situation and typically produces high WERs. In recent years, the development of deep recurrent neural networks (DRNN) supported by computer memory advancements and increased processing power, have vastly improved the WER for STT applications.

Notable mentions of STT technologies include IBM Watson and Google Deepmind. The use of ASR for intelligent personal assistants (IPA) has also become very popular and has seen a significant increase

in commercial activity. ASR technologies that enable users to interface with smartphones in natural language are prevalent in systems such as Siri, Cortana and Google Now. Whilst these have become reasonably mature for personal use, a recent study suggests that using speech-based technology to actively assist meetings reliably is currently too challenging (McGregor and Tang, 2017). Developments in ASR are continually improving STT capabilities to generate automated transcripts from speech. Since IPAs tend to deal with more simple dictated sentences in the form of queries, the technology is currently well placed to provide a support function to designers in searching for information. However, in challenging environments such as those experienced when undertaking collaborative design activities, STT technologies are not yet mature enough to handle multiple speakers, in noisy environments, with continuous, spontaneous speech, for speaker independent situations. However, with on-going improvements in accuracy offered by neural networks for example, this could be possible in the near future, to provide fully automated and searchable transcriptions to capture collaborative design activities (Dai and Velde, 2017).

3.2.2. Text summarisation (Levels 2-3)

Automatic text summarisation (TS) systems produce a summary of written text that should contain the most relevant information and condense the original text. Research in this field originates from the 1950s, and has been identified as being desirable for analysing multiple meeting recordings (Tucker and Whittaker, 2005). However, this still proves to be a challenging task (Gambhir and Gupta, 2017). Text summarisation stems from the need for humans to digest growing volumes of text-based information which comes from content from repositories of corporate documents, electronic communications and many other sources. Text summarisation can be divided into several broad categories such as extractive or abstractive, indicative or informative, and generic or query-focused (Gambhir and Gupta, 2017). Gambhir and Gupta, along with Llotet and Palomar, provide comprehensive in-depth reviews into automatic text summarisation techniques. They identify a major challenging issue with TS is evaluation (Lloret and Palomar, 2012; Gambhir and Gupta, 2017). Gambhir and Gupta propose that future directions in TS include improving how information is presented, which is currently predominantly based on textual input and output. TS has been applied to the topic of information management within engineering, but it is typically used on e-mail corpus, patent documents or reports (Zhan et al., 2007). Gambhir and Gupta suggest that new approaches could also take input in the form of meetings, video, etc. and output in a format other than just text. However, they observe that performance of text summarisation is 'still moderate and summaries are far from perfect'. So whilst TS presents opportunities in this area and is ever improving in performance, it is evident that the capability is not yet at the level of maturity to effectively and robustly apply to industrial design-based meetings and collaborative working situations. However, if these issues were to be addressed to greatly improve TS quality, then applications in design activities such as providing discussion summaries or extracting the key points from written reports or transcriptions would be beneficial.

3.3. Audio-visual technologies

3.3.1. Video content analysis (Levels 2-3)

Video content analysis (VCA) is the ability to automatically analyse video to identify temporal and spatial events, which is a form of computer vision and artificial intelligence. Content-based video analysis for retrieval is about enabling the automatic identification of content features and meaningful structure (Dimitrova et al., 2002). Bridging the low-level audio-visual features and the high-level semantic representation of the video is a challenge (Garg and Ramsay, 2015). Automatic video analysis techniques can be used in the context of analysing and structuring video recordings of engineering design activities. The notion of creating searchable videos through the use of indexing, tagging and annotation provides video in a structured media. Technologies such as ClipMine (Ha, 2015) and Vosaic (2016) provide a metadata layer for video which predominantly relies upon manual user annotation and tagging through a crowdsourcing technique, although more automated methods are also being developed.

Automatic feature extraction techniques used within image processing include detection, recognition and tracking. Object detection and recognition is the ability to detect participants or artefacts within a

collaboration space in a video recording, such as identifying the number of people in attendance in a design review meeting, for example. Facial recognition techniques take this one step further by enabling participants to be identified in a video. These capabilities are relatively mature and industrialised, especially when there is a large set of training data and the video is of reasonable quality. Motion detection enables the automatic identification of features within a video whereby there is movement of participants or artefacts within the collaboration space which is identifiable as an event. Motion tracking techniques then provide the capability to automatically follow the participant or artefact around the video in space and time.

Structure analysis enables structural boundaries within the video to be detected before segmenting the video into sections. This is aimed at automatically generating a form of contents table from the video. This tends to be effective for produced videos such as films, which have an overarching story which is captured by different scenes which are further broken down into a series of shots. However, for raw videos captured from design meetings where there is one continuous shot and no post-production editing to ‘tell a story’, this becomes particularly challenging to automatically detect and structure the video into segments. This is especially difficult where there is limited movement or visual interaction of participants. Similarly, video abstraction or summarisation is where a subset of a video is extracted such as key frames that have been identified. Automatically extracting video highlights is a particularly challenging task in this context of engineering design activity video recordings. This is due to the problem of such an approach relying on determining what is important from visual or audio data rather than the meaningful information that the utterances or gestures convey.

3.3.2. Participant identification (Levels 2-3)

The technologies and patents in this area relate to the identification of participants involved within a meeting. The patented methods broadly fall into two categories, namely audio-based and visual-based.

Microsoft developed a system within which speaker clustering and spatial indexing is possible to improve a post-meeting review. Other organisations such as IBM, EADS (Airbus) and Verint Systems have also made progress in methods to diarise audio signals from speech to segment the audio into speaker channels (Espacenet, 2017). The results of this can later be used for speech-to-text and transcription purposes. Such segmentation methods may be trained or untrained, to identify the number of speakers and their spatial positioning. Anguera et al. undertook a review into the state-of-the-art in speaker diarisation, with a particular focus for an application for conference meetings (Anguera et al., 2012). The review highlighted the single biggest challenge is in handling overlapping speech to attribute to multiple speakers.

Whilst sound can be used in speaker identification, biometric characteristics from a visual approach are also apparent within the patent records. Speaker identification may describe identifying the spatial position of a speaker through either audio or image processing. However, speaker identification may also refer to the ability to actually recognise a speaker based on their characteristics such as voice or face. NTT Corporation developed a demonstrator meeting analyser system to recognise who said what in real-time and provide assistance to the discussion. The system captured the audio speaker utterances and visual facial poses. It then segmented the speech and tried to interpret the state of each participant, such as laughing or watching someone. In addition, it also attempted to detect and display the circumstances of the meeting such as topic or casualness of the discussion, and display with the transcript in a browser (Hori et al., 2012). These techniques involve approaches such as facial recognition and human state detection to not only segment a speaker’s utterances, but also to relate those utterances to a profile of an individual outside of the meeting capture system (Espacenet, 2017). The advantage of these speaker identification approaches is that participants can be identified automatically, which provides a richer capture of information that can be attributed to individuals. However, a major limitation with many of these techniques is the inability for the technology to adapt to changing circumstances such as the movement of participants around a room or the capture device being moved. Whilst these technological advancements provide significant benefits, their effectiveness relies heavily upon their robustness which is very challenging in real-world settings such as engineering design activities.

3.3.3. Event logging (Levels 1-3)

The patents in this area that are discussed below show how a variety of multimedia recordings from a meeting or conference can be synchronised for playback and review purposes. Such multimedia sources

include video, audio and whiteboard interaction. This form of indexing of multimedia content provides a means of bookmarking the points of interest within the recordings. This bookmarking may be automated or manual.

Microsoft have several patents in this area, for example, the ability to timestamp pen strokes on a whiteboard along with an audio recording, which aims to support users to view and manage the captured content by using this visual index. Xerox have also patented a system to capture and playback temporal data related to events that occurred during the course of collaborative activities such as a meeting. The advantage of this type of approach is that it provides a means of improving accessibility to the information contained within an audio recording, by synchronising the two media sources. However, the limitation of this approach is that the useful structuring of the audio depends heavily upon the ability of the user to sketch or write to provide an appropriate means of visual indexing. This approach of indexing multimedia data reduces information loss, but the knowledge encoded within the audio recording is not easily searchable and is likely to be difficult to find by anyone other than the producer of the material, supported by memory from the discussion.

There are also a number of smartphone and tablet applications that enable personal notes to be taken which are then synchronised with an audio recording. The advantage of this approach is that it provides a good reminder to the note-taker of the context of the discussion from which they wrote their notes. It also reduces the information loss, as note-takers have access to the audio recording, which is structured by the digital notes or sketches. This form of note-taking is an improvement to the traditional logbook, capturing a structured context to the notes taken. Its main benefit is in supporting the note-taker in remembering the context in which their informal personal notes were written. However, whilst this is useful for a person who is creating an informal record of the meeting, it is not necessarily usable for anyone other than the note-taker themselves. Examples include AudioNote, SoundNote, and Notability. Evernote also have a number of patents in this area of personal note-taking, but this feature of timestamping notes and using them to index audio recordings is something that is not implemented in their technology products.

Microsoft have patented the architecture to facilitate the ability to trigger the capture of meeting state/context by a manual user click operation. Similarly, Cisco have patented a method to bookmark a moment within a teleconference meeting, which can later be reviewed in audio playback. The benefit of this approach is that important events may be bookmarked, but it relies upon the user's ability to manually log this.

3.4. Collaboration technologies (Level 4)

Industry reliance on remote collaboration is evident in the number of patents related to remote conferencing and collaborative working. The main driver for this is in improving the efficiency and effectiveness for collaborative working. Technical developments in the area of computer-supported cooperative working (CSCW) improve support for remote collaborative design activities. The technology in the area of remote conferencing can be broadly classified into (i) audio teleconferencing (multiparty voice), (ii) video and web conferencing, and (iii) telepresence.

An audio teleconference is typically a three-way or more phone call that facilitates voice communication where there is no need to share visual information. Many phone conferencing systems use Voice over IP (VoIP) technology which enables conference calls to be made over a computer network. Many of the large teleconferencing companies also offer video and web conferencing facilities in addition to an audio call.

Typically video and web conferencing technologies offer the means of sharing live presentations and software applications, streaming multimedia, instant messaging and overlaid annotation. However, these technologies such as that which can be seen from the patents of Vidyo (Espacenet, 2017), are related to improving video conferencing itself, but usually not in recording or capturing information from the meeting aside from a coarse screen capture with audio. Other companies developing technologies in this area include Lifesize, Huawei, Avaya, and ZTE. Collaborative boards such as the Cisco Spark Board, Microsoft Surface Hub and Google's Jamboard provide a digital whiteboard that can be used for multiple collaborative purposes such as offering projector, whiteboard and audio-visual conferencing facilities. The interactions and outputs of these can be captured from collaborative activities.

Telepresence is the use of mixed reality technologies to create a sense of being in another location, known as 'presence'. Virtual reality (VR) and augmented reality (AR) technologies such as the Oculus

Rift and Microsoft HoloLens are aimed at improving the perception of presence through immersion, which can be used to improve collaboration within engineering design. Whilst each of these technologies supports collaboration itself, the use of them for design activities also provides new means of capturing interactions with others from an individual's unique perspective. For example, an engineer wearing a Microsoft HoloLens headset is able to collaboratively work in the real world with an overlay of holographic 3D representations (Kowalski et al., 2015). Mimesys is an example of a company that has developed a technology that provides a remote collaboration holographic platform. These representations could be people or digital artefacts augmented from elsewhere, and it is possible to record this augmented view that gives the unique perspective of the individual undertaking a collaborative design activity. Whilst the cost and hardware size for AR/VR devices has significantly reduced in recent years, the challenges with these technologies still remain in that they are still relatively large, heavy and cumbersome for sustained use for undertaking design activities. Whilst these technologies provide many advantages, their adoption into industrial settings is currently not particularly widespread due to usability issues.

These technologies are mainly aimed at improving collaborative design activities. However, when designers use these supporting technologies there are benefits to be gained in terms of capturing these design activities. The advantage of using such synchronous remote collaboration technologies is that it can be far easier to capture data from a digital platform than from a meeting room. This comes with its own challenges, but capturing data, and thus information and knowledge from such a platform is naturally simpler as it is already in digital form, unlike a face-to-face discussion, which requires conversion from the analog world to a digital form.

3.5. Big Data & advanced analytics (Levels 2-3)

The amount of data that is generated by individuals and engineering organisations is rapidly increasing. The challenge is to utilise these vast datasets to gain insights and understanding about the subject matter. The recent advancements in Big Data technologies have improved management and analysis to enable the handling of very large datasets efficiently. Big Data is a term that describes several enabling technologies to analyse vast datasets. It is characterised by the four Vs of data, namely, (i) volume, (ii) velocity, (iii) variety, and (iv) value. Whilst the volume of data relates to the large amount of data that is becoming available, the velocity describes the speed at which the data needs to be handled to extract insights. The variety relates to the different types of structured and unstructured data that is available, whilst the value represents the benefit and challenge of extracting real insights from the data (Rajeshwari, 2015).

As the cost of data storage continues to fall, if the amount of data that is captured for design-related activities increases, the more dependent the sense-making will rely upon Big Data technologies. The speed at which data needs to be processed will only increase as more automated systems offer to provide real-time intelligent assistance to designers. The variety of data is particularly apparent within engineering design from a number of different sources in different forms. The value of extracting useful information from datasets within engineering design can only be seen when this has a direct impact on the design or design processes. Technologies such as Apache Spark and Google BigQuery are enabling very large datasets to be handled and analysed efficiently. With the growing adoption of data warehouses and data lakes within engineering organisations, there is an opportunity to benefit from large scale analysis of data resulting from engineering design activities (Rajesh and Ramesh, 2016).

Advancements in Artificial Intelligence (AI) in recent years have shown resurgence in the use of neural networks with deep learning capabilities to handle large datasets. AI is being used in many applications in different domains and certainly has a place within engineering design. Advanced machine learning techniques enable complex datasets to be analysed to identify and interpret patterns, such as designer behaviour. Machine learning algorithms can be used to discover hidden patterns and unexpected relationships between elements within datasets. Such AI technologies enable knowledge extraction and representation from datasets resulting from design activities (Borana, 2016).

In the context of reviewing and retrieving design activities such as reviewing meetings, Human Computer Interaction (HCI) technologies are showing how users can interact with computers in new ways. Interactive dashboards are currently widespread in industry, which provide the means for a user to view a series of data or information visualisations and interact to explore the data and information.

Improvements in web-browser capabilities such as html5 and software visualisation libraries such as D3.js have enabled web-based interactive and exploratory visualisations to become readily available. Such developments in graphical user interfaces (GUI) have enabled multi-media playback to be integrated within meeting room browsers with other data visualisations to provide simpler means of reviewing design meeting activities (Chandrasegaran et al., 2017). Emerging technologies such as Fluid Interfaces (Fluid Interfaces Research Group, 2016) are challenging the ways that designers can interact with digital information through more natural interfaces. Such human-computer advancements can be applied to collaborative design to make the experience more intuitive, intelligent and better integrated into daily lives. For example, digital artefacts such as 3D CAD models can be viewed through augmented reality devices and manipulated using hand gestures. Such technologies currently tend to be laboratory-based, with limited adoption in engineering design organisations. However, if they are developed and adopted by engineers, then it presents an opportunity to capture rich interactions between individuals and digital artefacts which can be aggregated with other datasets and analysed.

4. Conclusion

This technology review paper has identified, synthesised and presented the landscape of the state-of-the-art in key technologies that currently contribute, or will contribute in the future to capturing, processing and reviewing collaborative design activities. Future improvement in capabilities in this context will be heavily influenced by the development of these technologies and advancements in related technological fields. New methods can be developed, drawing from advancing technologies from a range of fields. This review established that some of these are currently limited in their maturity, but may provide value in the future. Combinations of these technology blocks can be explored as they become more mature in their respective fields to provide value greater than the sum of their parts.

As technology advances, there will be more opportunities to automate elements of the capture and retrieval process, potentially eliminating the need to trade-off the detailed accessibility of information with manual effort. For example, if speech-to-text technology reaches a suitable maturity for meeting room conditions, capture of a discussion would be possible with no human up-front effort. It would then be possible to semantically extract meaning from the discussion using NLP techniques, and utilise AI algorithms to automatically ‘push’ this information to individuals at the appropriate time, to support the design activity itself. Improving natural language processing capabilities may aid support of capturing the activities, but will provide a limited benefit without a focus on information retrieval.

Future methods and tools should leverage the opportunities presented by mature and emerging technologies for application. The successful development and application of these technologies will reduce the information and knowledge loss within engineering organisations, to support re-use for re-design and maintenance purposes. But as the technologies mature, they will also provide a capability to improve the design process itself, through actively supporting engineers in their collaborative design activities.

Acknowledgements

This work was supported by the Systems Centre and the EPSRC funded Industrial Doctorate Centre in Systems (Grant EP/G037353/1) and Airbus.

References

- Al-Hames, M., Hornler, B., Scheuermann, C. and Rigoll, G. (2006), “Using Audio, Visual, and Lexical Features in a Multi-modal Virtual Meeting Director”, *Proceedings of MLMI 2006, Bethesda, USA, May 1-4, 2006*, Springer, Berlin, Heidelberg, pp. 63-74. https://doi.org/10.1007/11965152_6
- Anguera, X., Bozonnet, S., Evans, N., Fredouille, C., Friedland, G. and Vinyals, O. (2012), “Speaker Diarization: A Review of Recent Research”, *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 20 No. 2, pp. 356-370. <https://doi.org/10.1109/TASL.2011.2125954>
- Ball, A., Patel, M., McMahan, C., Culley, S., Green, S. and Clarkson, J. (2006), “A Grand Challenge: Immortal Information and Through-Life Knowledge Management (KIM)”, *The International Journal of Digital Curation*, Vol. 1 No. 1, pp. 53-59. <https://doi.org/10.2218/ijdc.v1i1.5>
- Bettors, E. (2016), *Best 360 Cameras*. [online] Pocket-lint. Available at: <http://www.pocket-lint.com/news/137301-best-360-cameras-the-best-vr-and-360-video-cameras-no-matter-your-budget> (accessed 13.02.2017).

- Borana, J. (2016), "Applications of Artificial Intelligence & Associated Technologies", *Proceedings of ETEBMS-2016, Jodhpur, Rajasthan, India, March 5-6, 2016*, pp. 64-67.
- Bracewell, R., Wallace, K., Moss, M. and Knott, D. (2009), "Capturing design rationale", *Computer Aided Design*, Vol. 41 No. 3, pp. 173-186. <https://doi.org/10.1016/j.cad.2008.10.005>
- Buckingham Shum, S. (1996), "Design Argumentation as Design Rationale", *The Encyclopedia of Computer Science and Technology*, Vol. 35 No. 20, pp. 95-128.
- Carrizosa, K., Eris, O., Milne, A. and Mabogunje, A. (2002), "Building the design observatory: a core instrument for design research", *Proceedings of the DESIGN 2002 / 7th International Design Conference, Dubrovnik, Croatia, May 14-17, 2002*, pp. 37-42.
- Chandrasegaran, S., Badam, S.K., Kisselburgh, L., Pepler, K., Elmqvist, N. et al. (2017), "VizScribe: A visual analytics approach to understand designer behaviour", *International Journal of Human-Computer Studies*, Vol. 100 No. 1, pp. 66-80. <https://doi.org/10.1016/j.ijhcs.2016.12.007>
- Conklin, J. (2005), *Dialogue mapping: building shared understanding of wicked problems*, Wiley, New York.
- Conway, A., Wodehouse, A., Ion, W. and Juster, N. (2007), "A study of information & knowledge generated during engineering design meetings", *Proceedings of ICED'07 / the International Conference on Engineering Design, Paris, France, July 28-31, 2007*.
- Conway, A.P., Giess, M.D., Lynn, A., Goh, Y.M., McMahon, C.A. and Ion, J. (2008), "Holistic Engineering Design: A Combined Synchronous and Asynchronous Approach", *Proceedings of the International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, New York, USA, August 3-6, 2008*, ASME, pp. 1227-1236. <https://doi.org/10.1115/DETC2008-49340>
- Culmsee, P. and Awati, K. (2014), "The Map and the Territory: A Practitioner Perspective on Knowledge Cartography", In: Okada, A., Buckingham Shum, S.J. and Sherborne, T. (Eds.), *Knowledge Cartography: Software Tools and Mapping Techniques*, Springer, New York City, pp. 261-292. https://doi.org/10.1007/978-1-4471-6470-8_12
- Dai, X. and Velde, F. (2017), "How explicit are we in a design meeting: Investigation on meeting knowledge structuring with design rationale", *Proceedings of ICED'17 / the International Conference on Engineering Design, Vol. 6, Vancouver, Canada, August 21-25, 2017*, The Design Society, Glasgow, pp. 337-344.
- Dimitrova, N., Zhang, H.-J., Shahraray, B., Sezan, I., Huang, T. and Zakhor, A. (2002), "Applications of Video-Content Analysis and Retrieval", *IEEE MultiMedia*, Vol. 9 No. 3, pp. 42-55. <https://doi.org/10.1109/MMUL.2002.1022858>
- Espacenet (2017), *Patent Search*. [online] European Patent Organisation. Available at: <http://worldwide.espacenet.com> (accessed 18.04.2017).
- Fluid Interfaces Research Group (2016), *Integrating digital interfaces more naturally into our physical lives, enabling insight, inspiration, and interpersonal connections*. [online] Massachusetts Institute of Technology. Available at: <https://www.media.mit.edu/groups/fluid-interfaces/overview/> (accessed 02.02.2017).
- Gambhir, M. and Gupta, V. (2017), "Recent automatic text summarization techniques: a survey", *Artificial Intelligence Review*, Vol. 47 No. 1, pp. 1-66. <https://doi.org/10.1007/s10462-016-9475-9>
- Garg, A. and Ramsay, A. (2015), "Semantic Content Analysis of Video: Issues and Trends", In: Lin, W., Tao, D., Kacprzyk, J., Li, Z., Izquierdo, E. and Wang, H. (Eds.), *Multimedia Analysis, Processing and Communications*, Springer, Berlin, Heidelberg, pp. 443-457. https://doi.org/10.1007/978-3-642-19551-8_16
- Gasik, S. (2011), "A model of project knowledge management", *Project Management Journal*, Vol. 42 No. 3, pp. 23-44. <https://doi.org/10.1002/pmj.20239>
- Giess, M.D., Conway, A.P., McMahon, C.A. and Ion, W.J. (2008), "The integration of synchronous and asynchronous design activity records", *Proceedings of the DESIGN 2008 / 10th International Design Conference, Dubrovnik, Croatia, May 19-22, 2008*, pp. 785-792.
- Ha, A. (2015), *ClipMine Launch*. [online] Oath Tech Network. Available at: <https://techcrunch.com/2015/07/07/clipmine-launch/> (accessed 15.02.2017).
- Hain, T., Burget, L., Dines, J., Garner, P.N., Grezl, F. et al. (2012), "Transcribing meetings with the AMIDA systems", *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 20 No. 2, pp. 486-498. <https://doi.org/10.1109/TASL.2011.2163395>
- Hori, T., Araki, S., Yoshioka, T., Fujimoto, M., Watanabe, S. et al. (2012), "Low-Latency Real-Time Meeting Recognition and Understanding Using Distant Microphones and Omni-Directional Camera", *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 20 No. 2, pp. 499-513. <https://doi.org/10.1109/TASL.2011.2164527>
- Huet, G. (2006), *Design Transaction Monitoring: Understanding Design Reviews for Extended Knowledge Capture*, PhD thesis, University of Bath.
- Huet, G., McMahon, C.A., Sellini, F., Culley, S.J. and Fortin, C. (2007), "Knowledge Loss in Design Reviews", In: Tichkiewitch, S. (Ed.), *Advances in Integrated Design and Manufacturing in Mechanical Engineering II*, Springer, Grenoble, pp. 277-291.

- Kowalski, M., Naruniec, J. and Daniluk, M. (2015), "LiveScan3D: A Fast and Inexpensive 3D Data Acquisition System for Multiple Kinect v2 Sensors", *Proceedings of 3DV 2015 / International Conference on 3D Vision, Lyon, France, October 19-22, 2015*, IEEE, pp. 318-325. <https://doi.org/10.1109/3DV.2015.43>
- Lloret, E. and Palomar, M. (2012), "Text summarisation in progress: a literature review", *Artificial Intelligence Review*, Vol. 37 No. 1, pp. 1-41. <https://doi.org/10.1007/s10462-011-9216-z>
- McAlpine, H. (2010), Improving the management of informal engineering information through electronic logbooks, PhD thesis, University of Bath.
- McAlpine, H., Hicks, B., Huet, G. and Culley, S. (2006), "An investigation into the use and content of the engineer's logbook", *Design Studies*, Vol. 27 No. 4, pp. 481-504. <https://doi.org/10.1016/j.destud.2005.12.001>
- McGregor, M. and Tang, J.C. (2017), "More to Meetings: Challenges in Using Speech-Based Technology to Support Meetings", *Proceedings in CSCW, Portland, USA*, ACM. <https://doi.org/10.1145/2998181.2998335>
- McMahon, C., Giess, M. and Culley, S. (2005), "Information Management for Through Life Product Support: The Curation of Digital Engineering Data", *International Journal of Product Lifecycle Management*, Vol. 1 No. 1, pp. 26-42. <https://doi.org/10.1504/IJPLM.2005.007343>
- Okada, A. and Buckingham Shum, S. (2006), "Knowledge mapping with Compendium in academic research and online education", *Proceedings of International Council of Distance Education Conference, Rio de Janeiro, Brazil, September 3-6, 2006*.
- Olguin, D., Waber, B.N., Kim, T., Mohan, A., Ara, K. and Pentland, A. (2009), "Sensible Organizations: Technology and Methodology for Automatically Measuring Organizational Behavior", *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 39 No. 1, pp. 43-55. <https://doi.org/10.1109/TSMCB.2008.2006638>
- Rajesh, K. and Ramesh, K. (2016), "An Introduction to Data Lake", *i-manager's Journal on Information Technology*, Vol. 5 No. 2, pp. 1-4.
- Rajeshwari, D. (2015), "State of the Art of Big Data Analytics: A Survey", *International Journal of Computer Applications*, Vol. 120 No. 22, pp. 39-46.
- Rastogi, U., Tiwari, R. and Suman, S. (2017), "StereoCam3D (An Android App That Lets You Capture Realtime 3D Pics And Videos)", *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, Vol. 2 No. 2, pp. 837-840.
- Shah, D.H. and Shah, T.V. (2015), "Speech Recognition: An Approach to Modernization", *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol. 5 No. 8, pp. 1-6.
- Sim, S.K. and Duffy, A.H.B. (2003), "Towards an ontology of generic engineering design activities", *Research in Engineering Design*, Vol. 14 No. 4, pp. 200-223. <https://doi.org/10.1007/s00163-003-0037-1>
- Torlind, P., Sonalkar, N., Bergstorm, M., Eric, B. and McAlpine, H. (2009), "Lessons Learned and Future Challenges for Design Observatory Research", *Proceedings of ICED'09 / the International Conference on Engineering Design, Vol. 2, Palo Alto, USA, August 24-27, 2009*, The Design Society, pp. 371-382.
- Tucker, S. and Whittaker, S. (2005), "Reviewing Multimedia Meeting Records: Current Approaches", *Proceedings of the 2005 (ICMI) International Workshop on Multimodal Multiparty Meeting Processing, Trento, Italy*.
- University of Minnesota (2017), *Virtual Reality Design Lab*. [online] University of Minnesota. Available at: <http://vr.design.umn.edu/> (accessed 13.02.2017).
- Vergison, E. (2001), "Micro-scale Knowledge Management: A Necessary Condition for Getting Corporate Knowledge Properly Implemented", In: Roy R., (Ed.), *Industrial Knowledge Management: A Micro-level Approach*, Springer, London, pp. 3-15. https://doi.org/10.1007/978-1-4471-0351-6_1
- Vosaic (2016), *Homepage*. [online] Vosaic. Available at: <http://www.vosaic.com> (accessed 06.02.2017).
- Zhan, J., Loh, H.T., Liu, Y. and Sun, A. (2007), "Automatic Text Summarization in Engineering Information Management", *Proceedings of the International Conference on Asian Digital Libraries, Hanoi, Vietnam, December 10-13, 2007*, Springer, Berlin, pp. 347-350. https://doi.org/10.1007/978-3-540-77094-7_44
- Zhang, Z. (2012), "Microsoft kinect sensor and its effect", *IEEE multimedia*, Vol. 19 No. 2, pp. 4-10. <https://doi.org/10.1109/MMUL.2012.24>
- Zwyszig, E., Renals, S. and Lincoln, M. (2012), "Determining the number of speakers in a meeting using microphone array features", *Proceedings of ICASSP 2012, Kyoto, Japan, March 25-30, 2012*, IEEE, pp. 4765-4768. <https://doi.org/10.1109/ICASSP.2012.6288984>

Mark Hall, Research Engineer
 Airbus, Data-Driven Technologies
 C2 Quadrant House, NP10 8FZ Newport, United Kingdom
 Email: mark.hall@airbus.com