TeMoCo-Doc

TeMoCo-Doc: A visualization for supporting temporal and contextual analysis of dialogues and associated documents

Shane Sheehan
University of Edinburgh
United Kingdom
shane.sheehan@ed.ac.uk

Saturnino Luz
University of Edinburgh
United Kingdom
s.luz@ed.ac.uk

Pierre Albert
University of Edinburgh
United Kingdom
pierre.albert@ed.ac.uk

Masood Masoodian
School of Arts, Design and Architecture
Aalto University
Finland
masood.masoodian@aalto.fi

ABSTRACT

A common task in a number of application areas is to create textual documents based on recorded audio data. Visualizations designed to support such tasks require linking temporal audio data with contextual data contained in the resulting documents. In this paper, we present a tool for the visualization of temporal and contextual links between recorded dialogues and their summary documents.

CCS CONCEPTS

• **Human-centered computing** → *Visualization design and evaluation methods*; *Information visualization.*

KEYWORDS

Assessment of medical communication; temporal visualization; temporal mosaics; speech visualization; contextual visualization

ACM Reference Format:
Shane Sheehan, Saturnino Luz, Pierre Albert, and Masood Masoodian. 2020. TeMoCo-Doc: A visualization for supporting temporal and contextual analysis of dialogues and associated documents. In *International Conference on Advanced Visual Interfaces (AVI '20), September 28-October 2, 2020, Salerno, Italy.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3399715.3399956

1 INTRODUCTION

Written reports of events or actions which have related recorded audio or video data are widespread across a variety of domains. For example, medical contexts where recording and reporting is commonplace, audio and video recordings are often made during clinician-patient consultations, multidisciplinary medical team meetings and training, and so on. In these medical contexts, effective verbal communication is crucial to the success of clinical encounters, and as such, several frameworks have been developed to help standardise the analysis of medical communication (such as the widely used Roter Interaction Analysis System (RIAS) [6]).

AVI '20, September 28-October 2, 2020, Salerno, Italy

ACM ISBN 978-1-4503-7535-1/20/09.
https://doi.org/10.1145/3399715.3399956

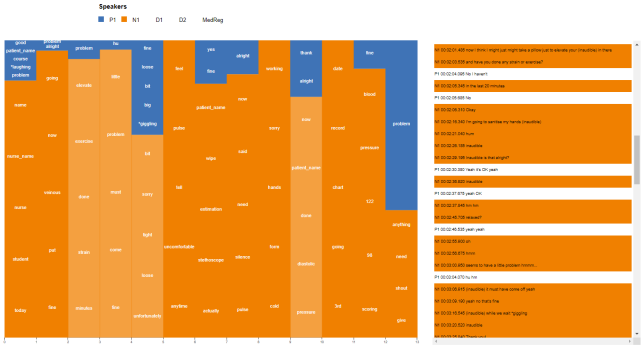

Figure 1: Temporal mosaic (left) and transcript (right) in the *TeMoCo-Doc* visualization.

Such analyses can be very time-consuming and not well supported by existing visualization tools. For instance, the link between an audio recording and a textual document is rarely made explicit, thus making it difficult to quickly switch between the textual document and the recording. To identify which spans of text in a report are linked to particular times in a recording requires either a full examination of both sources, or some other linking mechanism.

In this paper we present an interface which maps a textual document via contextual links to a temporal visualization of the contents of a recording and individual speaker's contributions. In addition, both the document and temporal visualization are implicitly linked to a transcript of the recording (which could easily be replaced by a highlighted timeline with accompanying video or audio). This interface is an extension of the *TeMoCo* visualization [7] which visualizes the speakers involved in each time-slot (without showing the amount of their speech contributions) in a recording, and enables exploration of the transcript via interaction with a temporal mosaic visualization. While this prior work has focused on identifying the temporal content, the work presented here focuses on identifying the links between the contents of textual documents and the temporal speech segments of recorded audio. In this paper we describe this visualization tool (called *TeMoCo-Doc*).

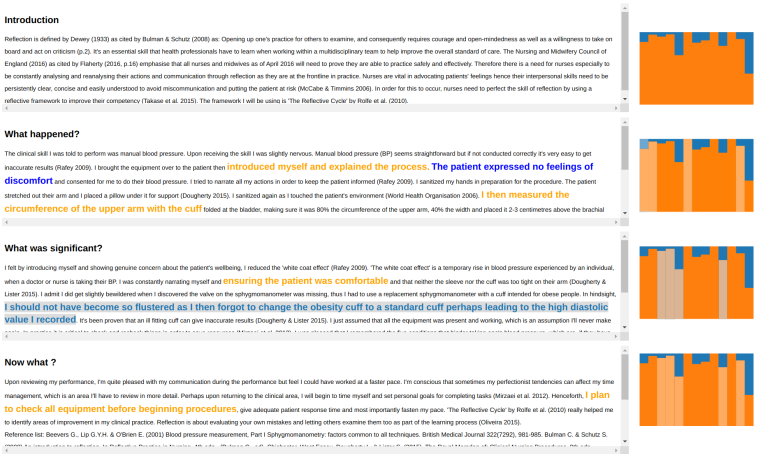

Figure 2: Textual document with contextual links in the *TeMoCo-Doc* visualization.

2 THE *TEMOCO-DOC* VISUALIZATION

The proposed *TeMoCo-Doc* visualisation combines three juxtaposed views [2]: the temporal mosaic (Figure 1, left), the transcript (Figure 1, right), and the view containing the report (Figure 2) which links all three views via manually encoded contextual edges. The interface is rendered across two web pages: one contains the report, the other contains the temporal and transcript views. These pages communicate via the JavaScript broadcast channels API.

The temporal aspect of the visualization, seen in the prototype Figure 1, uses temporal mosaic visualizations [4] to render the speaker contribution per time-slot, and displays the salient word for each speaker in the corresponding slot. This choice of encoding is informed by Mackinlay's ranking [5] of visual variables [1]. The order of the time-slots is mapped to horizontal position, the quantity of speaker contribution is mapped to length, and the categorical speaker information is mapped to color hue. The transcript is placed next to the temporal mosaic, and the speech segments are placed in a scroll box.

When used as an interactive visualization [3] each segment of a temporal mosaics visualization can be linked to the corresponding part of the data-stream it represents – thus supporting access to media content, both temporally as well as contextually. In Figure 1 one speaker has been selected in four time-slots. This causes the transcript to scroll to the beginning of the first selected time-slot, and the corresponding speech segments are highlighted. This follows the well-known visual information seeking mantra [8]. This combination of temporal mosaic and transcript is based on the *TeMoCo* visualization [7] which displays speakers for each time-slot evenly (i.e. not based on speech contribution) and does not allow for contextual linking across multiple time-slots.

Mini temporal mosaics (with the salient words removed) are rendered beside each section of the report view. Contextually linked spans of text in each section are colored according to the speaker they map to. By hovering over any of these contextual spans the user can see the related time-slots on the mini mosaic, and the highlight is maintained if the span is clicked (as can be seen in the third mosaic in Figure 2), which gives a preview of the slots that will be selected by clicking on the span. Each of the document sections is presented in a fixed height scroll-box, so that even large documents can be explored in a single screen.

The example shown across figures 1 and 2 represents a reflexive report based on a medical communication training exercise. In Figure 2, the user can see the span "I should not have become so flustered as I then forgot to change the obesity cuff to a standard cuff perhaps leading to the high diastolic value I recorded" is clicked, outlining the related time-slots on the mini mosaic. Looking at the large mosaic in Figure 1, the user can see the same time-slots are again selected, and the transcript is highlighted accordingly. This allows the user to explore the transcript, and quickly see the time-slots and content in the conversation which relate to "I plan to check all check all equipment before beginning procedures".

The design of the *TeMoCo-Doc* visualization was informed by collaboration with colleges involved in medical education. Preliminary evaluation of the prototype led to some adjustments after experts' feedback, issues with highlighting and color were resolved. Some of the medical domain experts had difficulty with the representation of time using different axis: top to bottom for the transcripts, left to right for the mosaic. While the use of 3 panels for the visualization was received positively, the number of words rendered in the temporal view were seen as potentially overloading, thus needing further explanation(the salient words do not necessarily reflect medical issues or problems).

Overall, the experts found the tool not sufficiently developed to be used as a primary assessment tool for scenario-based medical training. A major limitation for the use in assessment is due to the diversity and complexity of aspects used to evaluate communication and technical skills which cannot be reflected by text alone. As such, the experts stated that the prototype could not be used to replace watching the full session. The visualization tool was considered suitable for its primary purpose comparing students. The potential of this tool for distant learning in medical training was suggested by the domain experts.

In the future we will extend the interface, investigate other application areas and carry out formal evaluations of the visualisation tool in application specific areas.

ACKNOWLEDGMENTS

This paper is supported by European Union's Horizon 2020 research and innovation programme under grant agreement No 825153, project EMBEDDIA (Cross-Lingual Embeddings for Less-Represented Languages in European News Media). The results of this paper reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains. Pierre Albert has been funded though the INCA project. We thank the INCA project members in Ireland for granting us access to the trainee data.

REFERENCES

[1] Jacques Bertin. 1983. *Semiology of Graphics.* University of Wisconsin Press.

[2] Waqas Javed and Niklas Elmqvist. 2012. Exploring the design space of composite visualization. In *Pacific Visualization Symposium (PacificVis), 2012 IEEE.* 1–8. https://doi.org/10.1109/PacificVis.2012.6183556

[3] Saturnino Luz and Masood Masoodian. 2005. A Model for Meeting Content Storage and Retrieval. In *Proceedings of the 11th International Multimedia Modelling Conference (MMM '05).* 392–398. https://doi.org/10.1109/MMMC.2005.12

[4] Saturnino Luz and Masood Masoodian. 2007. Visualisation of Parallel Data Streams with Temporal Mosaics. In *Proceedings of the 11th International Conference Information Visualization (IV '07).* 197–202. https://doi.org/10.1109/IV.2007.127

[5] Jock Mackinlay. 1986. Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics* 5, 2 (1986), 110–141. https://doi.org/10.1145/22949.22950

[6] Debra Roter and Susan Larson. 2002. The Roter interaction analysis system (RIAS): utility and flexibility for analysis of medical interactions. *Patient education and counseling* 46, 4 (2002), 243–251.

[7] Shane Sheehan, Pierre Albert, Masood Masoodian, and Saturnino Luz. 2019. TeMoCo: A Visualization tool for Temporal Analysis of Multi-Party Dialogues in Clinical Settings. In *2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS).* IEEE, 690–695.

[8] Ben Shneiderman. 1996. The eyes have it: a task by data type taxonomy for information visualizations. In *Proceedings the IEEE Symposium on Visual Languages.* 336–343. https://doi.org/10.1109/VL.1996.545307