Exploring Digital Media and Robotic Performance for Use in Educational Development: How Might Different Forms of Media Foster Reflection by A Novice Teacher?

D'Arcy Norman

University of Calgary

ABSTRACT

Technical design and implementation of a graduate course project involving digital media (video, audio, text) and robotic performance is described, as well as a preliminary discussion of the findings and potential next steps.

KEYWORDS

Human-robot interaction, digital media, abstract media, performance

INTRODUCTION

This project was conducted from January – April, 2017, as part of my course activities as a PhD student in Computer Science 701.21: Advanced Topics in Human-Robot Interaction. The goal of the project was to explore human-robot interaction, within the context of my PhD research agenda – developing a framework to support critical reflection and educational development by novices who are learning the craft of teaching. This paper describes the technical design and implementation of the project. A fuller description of the theory and connections to practice will be made in a subsequent article.

METHODS

For the project, I compared traditional video recording with various processed media formats with different levels of abstraction, in order to see if there was any potential benefit to attempting to remove the degree of "closeness" felt by a novice teacher who is watching a recording of their teaching. 3 participants volunteered (myself, and 2 other students from the University of Calgary's Interactions Lab[1]) to participate in the study. Each of the three participants performed a brief "introduction" presentation or mini-lesson, which was recorded in

[1] University of Calgary Interactions Lab http://ilab.cpsc.ucalgary.ca

high definition video using a Canon Vixia HF G30 high definition camcorder (Canon, 2017), with a RØDE VideoMic Pro microphone (RØDE, 2017). That recording was considered the canonical "recorded video and audio" version of the performance, and was then used as reference material in the conversion into the various media formats. Once the raw MP4 video file was captured, the audio track was then isolated as a separate AAC file, using QuickTime Player's (Apple, 2017a) "Export… Audio only" function. A text transcript of that audio file was manually prepared by the researcher, and this transcript was then used as source material for text-to-speech conversion to be used as the audio track for the "synthetic audio" format. Text-to-speech conversion was performed using macOS Sierra's integrated Text to Speech function (Apple, 2017b), using the command line "say" command to generate AIFF audio files for each presentation transcript, such as the following:

```
say -v Alex -f transcript-1.txt -o output.aiff
```

The raw MP4 video file was also imported into Adobe Premier Pro CC (Adobe Software, 2017), and the "Edge detection" filter was applied before exporting a version of the video to be used as the "synthetic" format. The "synthetic audio" file was imported into Adobe Premier Pro CC, and the duration of the audio was adjusted to match the duration of the video track. Exact lip syncing with the recorded video was not performed. Versions with the synthetic audio track were then produced and exported from Adobe Premier, one for the "recorded video" and one for the "synthetic video" track with edge detection applied. Video files of each version were then exported from Premier in standard MP4 format for later review on any device.

The robot performance formats were created using a SoftBank Robotics Nao v5 robot (SoftBank Robotics, 2017a), with motions manually authored in the Choregraphe software (SoftBank Robotics, 2017b) on macOS. The original MP4 video file for each introduction

performance was used as reference material, and unique "poses" and "gestures" were documented for timing and sequencing data, and manually recreated using the Choregraphe interface. These poses and gestures were then sequenced in the flow diagram panel (figure 1), and tested in the robot simulator as displayed in the robot view within Choregraphe. Early tests showed the Nao robot had difficulty maintaining balance with some of the poses, so movement of the legs was minimized in order to prevent damaging the robot. Final review of the robotic performances was conducted in the University of Calgary's Collaboration Centre[2] (University of Calgary, 2017), with the Choregraphe sequences uploaded and performed on the Nao v5 robot.

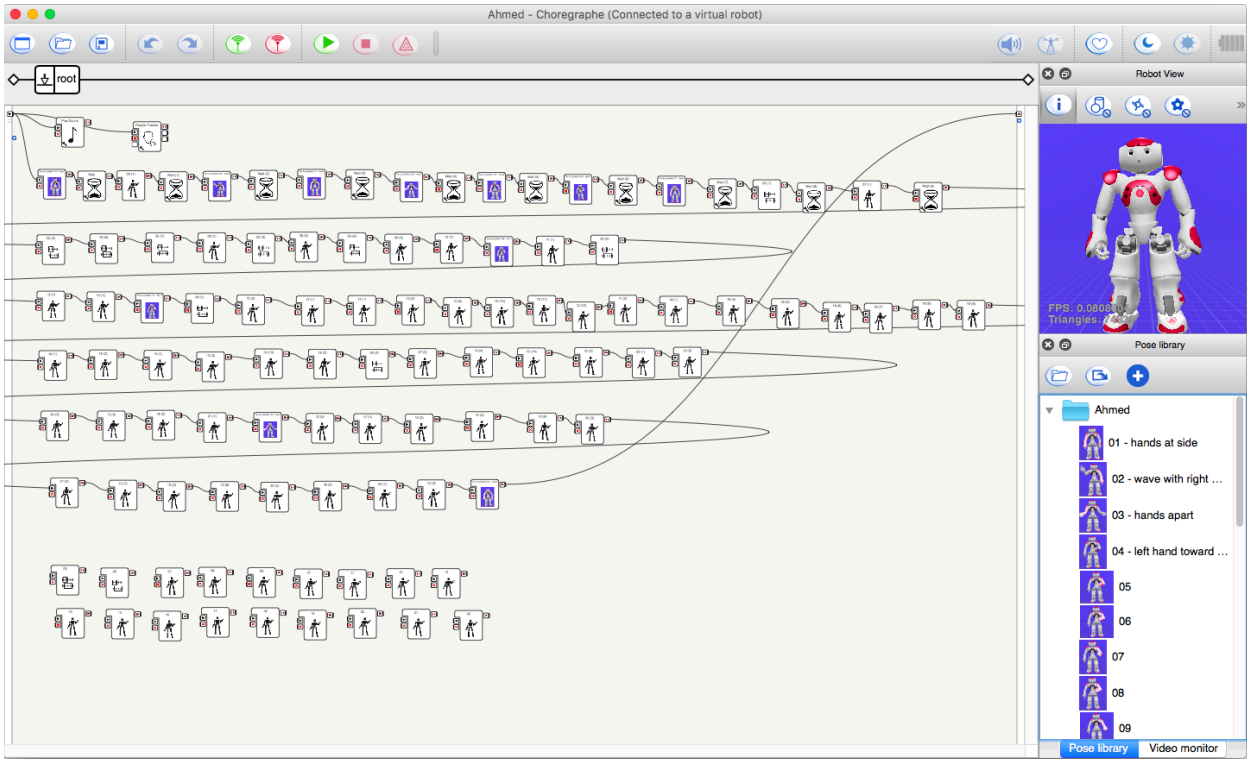

Figure 1. Choregraphe interface with sequenced poses, gestures and robot simulator.

[2] University of Calgary Collaboration Centre http://ucalgaryreservoirsimulation.ca/collaboration-centre

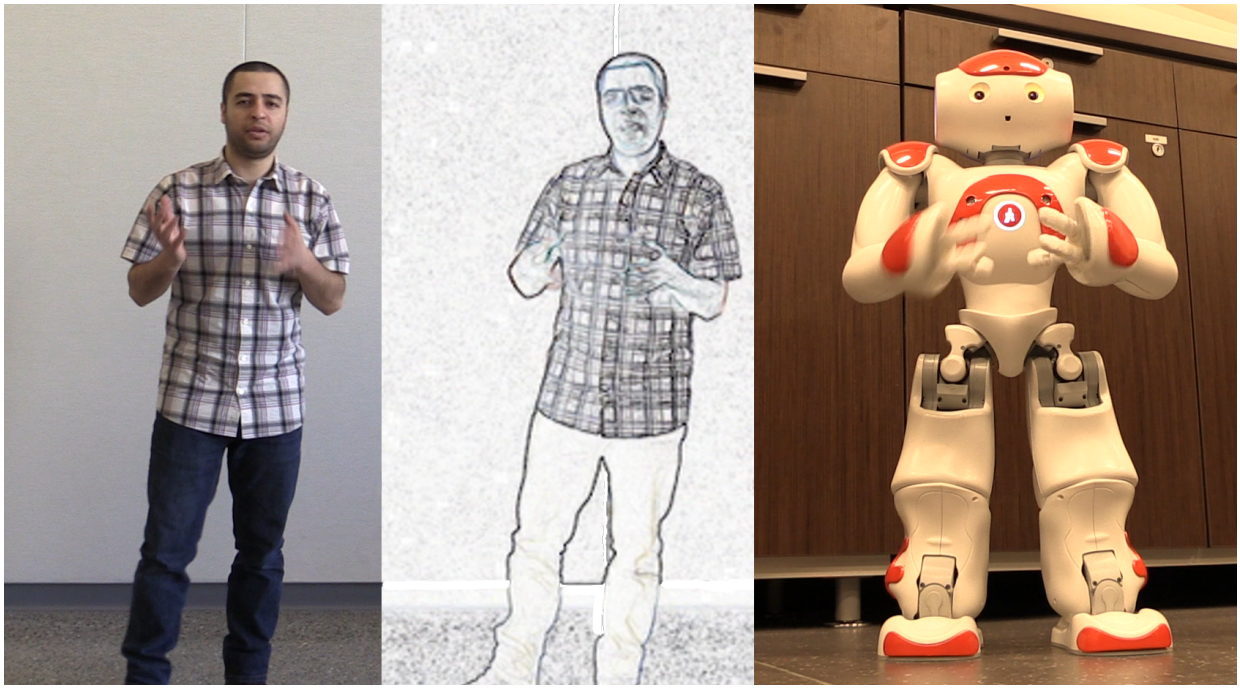

Figure 2. Three versions of media to review by the instructor: high definition video (left), edge-detection "synthetic" video (centre), and Nao robot performance (right).

Table 1 lists the 13 media formats that were recorded during the teaching session and created through various processing and conversion techniques. Figure 2 provides samples of three video files, representing different interpretations of the teaching session – the original, unprocessed, high definition video recording, the “synthetic” version created using the “edge detection” algorithm, and the robotic performance produced with the Nao humanoid robot.

Table 1. Media formats recorded or processed.

Media format	Audio	Video	Motion
Text-only	None	None	None
Synthetic audio	Synthetic	None	None
Recorded audio	Recorded	None	None
Recorded video + no audio	None	Recorded	None
Synthetic video + no audio	None	Synthetic	None
Recorded video + recorded audio (original recording)	Recorded	Recorded	None
Synthetic video + recorded audio	Recorded	Synthetic	None
Synthetic video + synthetic audio	Synthetic	Synthetic	None
Robot mimicked + no audio	None	None	Mimicked
Robot mimicked + recorded audio	Recorded	None	Mimicked

Robot mimicked + synthetic audio	Synthetic	None	Mimicked
Robot modified + recorded audio	Recorded	None	Modified
Robot modified + synthetic audio	Synthetic	None	Modified

RESULTS AND DISCUSSION

A full exploration of the 13 media formats of the recorded session was planned, but was unable to be completed due to scheduling conflicts, as the manual production of the robot performances took much longer than anticipated. Preliminary discussions with participants about the various media types showed interest in the "synthetic" and "robot" conditions. As the primary purpose of this project was to "test the testbed", important data about the recording and media conversion processes was gathered and will inform future projects.

All media types showed opportunities to support self-reflection by a teacher. Some had more potential to trigger deeper reflections and connections to theory and practice. The surprise reflection trigger was the stiff and poorly balanced Nao robot performance – exaggerating off-centre movements by the teacher, accompanied by audio feedback of servo motors working during the performance.

The Nao robotic performance provided two key surprises, both of which may be used as data or prompts for feedback and reflection. First, the noise of the servo motors as they adjusted the position of mechanical joints to move the plastic arms provided a surprising amount of information to observers who were reviewing the session. The whirring of the motors, sometimes with marked and pronounced changes in volume and intensity, prompted deeper observation and conversation about movements that were mostly overlooked during review of the video formats. Some of the movements also resulted in collisions between the robot's limb and body, producing a clicking sound and warnings in the Choregraphe software interface, providing further points of observation and reflection.

Since the Nao's performance was manually scripted, there were many opportunities to refine the performance to better mimic the recorded video of the teaching session. Most of these refinements were to the timing of motions, rather than to the motions themselves – which is something that was also described by Zeglin et al (2014) in their experimental robot theatre project, HERB's Sure Thing.

The more dramatic surprise involved the inability of the robot to dynamically adjust its balance in the same way that a human does. While initially perceived as an error condition or fatal limitation in the hardware, the observers quickly realized that this stiffness of balance provided an extremely important point of data. One repeated motion involved a shift in balance as one participant adjusted backward to put more weight on one leg. The movement was faithfully reproduced by the robot, who, lacking the dynamic balance adjustments, promptly fell backward onto the floor. While reviewing the original video of the teaching performance, this shift of balance was not explicitly noted, as it was a fluid movement that served as an anchor during transition point in the presentation. However, that off-balance posture is an important means of communication – in theatre, the blocking or nonverbal positioning and movement of a performer. This wasn't directly observed until made explicit as a data point during the failure of the Nao robot to reproduce the movement without falling.

CONCLUSIONS

While the various media formats all have potential to promote meaningful reflection and teacher development, the ongoing process of personal critical reflection is the key to an individual as they develop teaching skills. Technologies that enable an individual practitioner to record, review, and reflect on their teaching are essential, and these processes must be feasible and practical to conduct by non-technical people. This project is important in that it demonstrates

several media formats that can support meaningful reflection on teaching practice, but further work will need to be conducted in order to make these techniques available to teachers in the field.

FUTURE WORK

Automation

This initial project was intended to explore the viability of producing various interpretations of recorded documentation from a teaching session. As such, the researcher did not develop automated workflows or pipelines to convert media from one format to another, nor to automatically record positional and motion data during the teaching session.

Future projects could explore ways to automate the media conversion process, possible with real-time processing and layered data visualizations, as well as motion capture for the robot performance. These projects could also explore ways to implement cues in the processed recordings, based on data and guidelines provided by interdisciplinary fields such as dramaturgy, kinesiology, speech, in addition to educational theory and practice.

Immersive review

With automated recording of spatial and motion data, more robust volumetric capture of a session could be explored – both the textures and three dimensional meshes representing the scene and interactions can be recorded for playback. Once volumetric capture is implemented, immersive or virtual reality playback and review can be explored.

Structured reflection

Exploring how these technologies can be used as scaffolding or frameworks to guide teachers through meaningful reflection by connecting practice to theory. Can structured

performance data and reflection prompts be integrated in the recorded or processed media? How would this change the experience of a teacher or observer who reviews a recorded session for the purpose of teaching development?

REFERENCES

Adobe Software. (2017). Adobe Premier Pro CC 2017. https://www.adobe.com/ca/products/premiere.html

Apple Inc. (2017a). QuickTime. https://support.apple.com/downloads/quicktime

Apple Inc. (2017b). macOS Sierra. https://www.apple.com/macos/sierra

Canon U.S.A, Inc. (2017). Canon Vixia HF G30 Camcorder. https://www.usa.canon.com/internet/portal/us/home/products/details/camcorders/consumer/vixia/vixia-hf-g30

RØDE Microphones, Inc. (2017). RØDE VideoMic Pro Microphone. http://www.rode.com/microphones/videomicpro

SoftBank Robotics. (2017a). Nao v5 Humanoid Robot. https://www.ald.softbankrobotics.com/en/cool-robots/nao

SoftBank Robotics. (2017b). Choregraphe software download. https://community.aldebaran.com/en/resources/software (login required)

Zeglin, G., Walsman, A., Herlant, L., Zheng, Z., Guo, Y., Koval, M. C., et al. (2014). HERB's Sure Thing: A rapid drama system for rehearsing and performing live robot theater (pp. 129–136). Presented at the 2014 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO), IEEE. http://doi.org/10.1109/ARSO.2014.7020993