NORDICOM-INFORMATION 39 (2017) 2: 15-21

Paul McIlvenny & Jacob Davidsen

A Big Video Manifesto

Re-sensing video and audio

For the last few years, we have witnessed a hype about the potential results and insights that quantitative big data can bring to the social sciences. The wonder of big data has moved into education, traffic planning, and disease control with a promise of making things better with big numbers and beautiful visualisations. However, we also need to ask what the tools of big data can do both for the Humanities and for more interpretative approaches and methods. Thus, we prefer to explore how the power of computation, new sensor technologies and massive storage can also help with video-based qualitative inquiry, such as video ethnography, ethnovideo, performance documentation, anthropology and multimodal interaction analysis. That is why we put forward, half-jokingly at first, a Big Video manifesto to spur innovation in the Digital Humanities.

The Big Video manifesto critically draws upon our experiences with new capture and recording technologies, as well as with the workflow from data collection through to the publication of findings. We call this a re-sensing of video and audio, which leads to reflections on the nature of spatiality, mobility, agency, presence, intercorporeality, relationality, materiality, observation and documentation. Here are some of the key points of our manifesto:

1) No recording technology is neutral

Moving image and sound recording technologies are evidently not able to produce transparently faithful documents of an "original" event. Not only are there biases built into the technology, the underlying algorithms and its (default) uses, there are also issues with the function of the technology and the recordings (registrations) as "witness" and "testimony" for observers.

As a result, it is essential to appreciate both the affordances *and* the theoretical and practical limitations of making recordings as a methodological stipulation, as well as to question or re-establish the validity and reliability of legacy recordings.[1]

2) Proliferate the cameras and microphones

No setting or scene of social action and practice is too complex to record and analyse. Although using multiple recording devices might seem like overkill, we prefer to overlap the targets of the recording than risk losing a phenomenon we are, or will become, interested in.

There is, of course, a danger that increasing the number of visible recording devices may unduly influence natural human behaviour in the setting. This is also why researchers need to practice and reflect on their recording techniques and procedures even more in Big Video as every setting will require a tailored recording setup.

Paul McIlvenny, Professor, Department of Culture and Global Studies, Aalborg University.

Jacob Davidsen, Assistant Professor, Department of Communication & Psychology, Aalborg University.

3) Rummage through public video archives

Rather than insist on doing data collection ourselves, we also encourage collecting video from public video archives online, what we call "found video".[2]

With the everyday proliferation of video recording devices, one can often find swarms of video records of the same event from multiple perspectives, eg. using smartphones, that would otherwise not be accessible, and in such diversity, to the analyst. Be wary of fake videos!

4) Draw on and adapt professional solutions

In professional fields such as cinema, video, TV, game and theatre production, there is much to learn and adapt, though costs for high-end solutions may be beyond reach and some solutions are simply not applicable to our concerns.

For example, we have had to move to RAID disk array solutions usually offered to commercial video production for storing and editing multiple 4K video streams in which speed and capacity are crucial.

Also, we are using gaming engines such as UNITY[3] to enable the collaborative creation and editing of inhabited, immersive data.

5) Scavenge computation for qualitative enquiry

Computation can have its uses for enhancing qualitative methods. We are assembling video data sets which span many terabytes and hundreds of hours of footage, and as a result they are becoming unwieldly.

For example, there are low-level tasks in the qualitative workflow which burden us that could be handed on to a high-end workstation, high performance computing centre or cloud processing.

For example, the synchronisation of multiple data streams, composite video rendering, multicam stitching, low-level tagging, motion/feature tracking, anonymization efx, and gross collection building are all ripe for (parallel) computation.

6) Think scenography and virtuality, not just cinematography and televisuality

For too long we have been confined when using video cameras by the cinematic frame (4:3 or 16:9) and the flat screen. With expanding computational power, micro-sensors and massive storage, an important trend is the virtualisation of functionality so that there is increasing choice *after* the recorded event to change the framing, POV, focus, orientation, etc. in both video and audio. For example, using a 360-degree camera and an ambisonic microphone permit a greater degree of control over the observer's point-of-view and place-of-audition after the event.

On the horizon is volumetric video which combines 360-degree lightfield camera arrays with depth maps to do videogrammetry, which is to mathematically reconstruct a virtual video camera from multiple lenses in different locations. This will move us towards six degrees of freedom – being able to look around in every direction and move about in a limited space to view the same scene from different angles and positions – leading to a more active engagement with the video recording.

7) Capture and map space, depth and volume

A corollary of 6) is that up until recently the role of embodiment, space, materiality and performance has been missed, but this is the domain of expanded scenography and choreography.[4] We can learn a lot from such artistic approaches to encounter and exchange founded on spatial and material relations between bodies, objects and environments!

We need lightweight solutions to capture embodiments and materialities, and to track motion, movements and micro-mobilities. And we need to use the vertical dimension, not just the eye-level horizontal plane.

8) Inhabit the data

It is imperative that we creatively visualise our video and sensing data and develop tailored representations. Rather than 2D visual frames and stereo sound, we must think in terms of

three-dimensional fields that we perceive, such as lightfields, soundfields and sensefields. With VR developing rapidly, we are experimenting with building a system that supports researchers in different locations who will be able to simultaneously and collaboratively inhabit their video data streams in a 3D virtual space. They will be able to view and edit a segment of stitched 360 video in VR/AR, and annotate and comment together using gestures, controllers and 3D tools.

This will invert the traditional solution of 2D video clips clumsily tacked on to the primary, online written articles; instead, we can have immersive environments that "stage" video data in new visualisations, e.g. a 3D interactive world modelled on the setting in which the data was collected.

9) Down with planocentrism!

When it comes to presentation and publication, we quickly hit a barrier. We should lobby publishers and universities to provide web infrastructure, such as webVR, and flexible physical infrastructure spaces, so we can escape from the flat plane of one 2D screen, which is what Schröter[5] calls planocentrism.

Instead, they should support more immersive venues and transplanar viewers for publishing our work.

Paradigm shift	Feature
From analogue to digital	*Computationally intensive*
From singular to plural	*Multiple recording devices, such as cameras and microphones*
From sound as secondary to sound as covalent	*In-built microphones versus spatial audio*
From frame to field of vision	*16:9 versus 360-degrees*
From flat to depth	*2D versus stereoscopic 3D*
From cinematography to scenography	*Screen versus volume and body*
From spectator to POV	*Roving, mobile, virtual cameras*

Diverse paradigm shifts that influence the emergence of Big Video.

What paradigms and technologies are relevant?

From our trials and reflections, we contend that today there are a set of paradigm shifts that define the rise of what we call Big Video (see table above).

Following these paradigm shifts, one of our principal aims has been to collect richer moving image and sound recordings in a variety of settings. And this means developing a sense of good camerawork and micwork (with both existing and new technologies) in order to collect analytically adequate and malleable recordings.

We have used swarm "found" video gathered from online video archives; sports cameras; 4K video cameras; network cameras; 360-degree cameras; stereoscopic S3D cameras; wireless, binaural and ambisonic microphones; multi-track video and audio streams; GPS and heart rate tracking; and multi-track subtitling and video annotation. Also, we are beginning to work with local positioning systems (LPS) and beacons, as well as biosensing data and eye tracking, to see what is applicable to our qualitative methodological concerns.

Lately, we have also experimented with team filming, in which a pool of researchers and students specialise and refine their expertise with a particular recording device in order to collect a richer data set together.

One key area of development is finding ways to integrate the experimental tradition that supports more quantitative behaviour studies – e.g. eye-tracking, biosensing, motion capture – with more qualitative "in the wild" studies.[6]

How did we get here?

The history of moving images, and video in particular, is marked by milestones and landmark advances, such as the first attempt at a battery-operated portable film camera (Etienne-Jules Marey, 1899), the synchronisation of sound and film in a portable camera (NWDR Pilot-tone, 1954), the first portable video camera and recorder (Sony Portapak, 1967), and most recently the development of commercial omni-directional cameras and VR head-mounted displays. As studies have shown, there has always been a close, yet critical relationship between the advances in moving image technology and scientific inquiry, also in the humanities and social sciences.[7]

For example, anthropologists were using film as early as 1898 (A.C. Haddon) and experimenting with sound recording in the 1930s (Franz Boas). Later, in the 1960s and 70s, portable video equipment was taken up by some scholars to study human behaviour both in and out of the laboratory.

We argue that, in the 2010s, developments in the miniaturisation of digital sensors, computation and storage have precipitated a revolution in academic scholarship, leading to new ways to "see-with-a-camera", "be-with-a-camera", "sense-with-a-camera" and "hear-with-a-microphone". Scholars increasingly dabble with video as a tool or method in their field, though they tend to rely on the default settings of consumer kit – what we describe as the "algorithmic normativity" of audio-visual recording technologies. We also need to learn from professional solutions in cinema, video, sound, theatre and game production to advance our ways of understanding human sociality.

Nevertheless, it is important to appreciate that there is no one-size-fits-all solution. To further our methodological inquiry, we wish to use new technologies to open fresh perspectives on the lived work and local order of social practices, especially in regard to video ethnography and qualitative video analysis.[8]

Applications and uses

Given the potential for new understandings of human conduct and social practice, there are many uses and applications of Big Video.

First, we can capture and document the salient phenomena in an event or scene for later replay, staging and analysis. Second, we can use the data collected to better understand the praxeology of more complex social practices involving many human participants and non-human agencies.

Third, we may develop alternatives to standard solutions available to established cinematic genres such as documentary. Below we illustrate with a few images from data collected by us on a guided nature tour, a PokemonGo hunt, and a human-robot choreographic performance experiment.

Within the limitations of the planocentric 2D print publication format, each image illustrates one single moment at the event in question taken from multiple camera perspectives in a composite layout. Sometimes, there is an indication of who the participants were, though they are anonymised, and a list of the audio channels available that were recorded at the scene. These projects are ongoing and publications are forthcoming.

More information about Big Video and some of the applications can be found here: http://www.bigvideo.aau.dk.

Big Video Project in Aalborg

Over the last two years, with the support of our departments in the Faculty of Humanities at Aalborg University and the national Digital Humanities Infrastructure (DigHumLab) project, we have been quietly developing an enhanced approach to video ethnography and qualitative video analysis that we call Big Video. For us, Big Video focuses on complex human data capture scenarios, involving (1) the recording, storage, archiving and access of digital video; (2) the visualisation, transformation and presentation of data collected; (3) collaboration and sharing; and (4) qualitative tools to support analysis. A link to the project site: https://dighumlab.org/.

Examples of Big Video Projects

Nature tours

Pirkko Raudaskoski and Paul McIlvenny are exploring how citizens encounter nature(s) on guided nature tours conducted by experts. Rather than speculating about how we think of "nature", they are using Big Video in the wild to document how citizens and scientists encounter and debate Nordic nature in place.

Multiple cameras and mics recording a guided nature tour

PokemonGo in action

Jacob Davidsen is investigating social gaming in urban spaces using Big Video methods. The problem with many studies of locative AR gaming is that they do not have an appreciation of the rich texture of practices that support and enable the game in situ amongst a social group of participants. Big data may be able to shed light on gross, hidden patterns across a user population, but it misses completely the practices that made such "data" collectable and computable in the first place.

Multiple cameras and smartphones on a PokemonGo hunt

Archaeological site

Paul McIlvenny and Pirkko Raudaskoski are studying how particular types of reasoning about the past are conducted across disciplines such as archaeology, psychoanalysis, cultural theory and philosophy. In this project, they are using Big Video to track an interdisciplinary collaboration with cultural psychologists on the practical work of abductive reasoning at a Viking age archaeological site, both in the museum and at the site itself.

Human-robot performance

In a recent project, Paul McIlvenny and Pirkko Raudaskoski have used some of the latest Big Video technology to find answers to the question: How does one become a cyborg in practice? What role does a multimodal scenography and a choreographic preparatory phase have on exoskeleton rehabilitation? This is a collaboration between the university and two international artists renowned for their challenging works on misbehaving robots.

Multiple cameras and mics recording a human-robot performance experiment

Danish TV journalists at work

Tobias Boelt Back, a doctoral student in the Interdisciplinary Discourse Studies doctoral programme at AAU, is collecting Big Video data on the trajectories of affect in television production by following journalists building a story for national television over time and across sites and modalities. He has used a range of mobile cameras and microphones to track and capture the delicate teamwork that goes into assembling an interview segment from first lead to final studio production.

Notes

1. Mengis et al. (2016).
2. McIlvenny (2017).
3. UNITY is a software suite for designing and prototyping computer games. It is increasingly used for developing virtual reality and augmented reality.
4. McKinney and Palmer (2017).
5. Schröter (2014).
6. Kendrick (2017).
7. Evers (2011), Tosi (2005).
8. Davidsen and McIlvenny (2016).

References

Davidsen, Jacob & McIlvenny, Paul (2016). *Guest blog post: Jacob Davidsen and Paul McIlvenny on Experiments with Big Video*. ROLSI journal guest blog, October 17, 2016. Available: <https://rolsi.net/guest-blogs/guest-blog-jacob-davidsen-and-paul-mcilvenny-on-experiments-with-big-video/>. [Accessed August 28, 2017].

Evers, Jeanine C. (2011). From the Past into the Future. How Technological Developments Change Our Ways of Data Collection, Transcription and Analysis. *Forum: Qualitative Social Research* 12(1). [Online]. Available: <http://www.qualitative-research.net/index.php/fqs/article/view/1636/>. [Accessed August 28, 2017].

Kendrick, Kobin H. (2017). Using Conversation Analysis in the Lab. *Research on Language and Social Interaction* 50(1): 1-11.

McIlvenny, Paul (2017). Mobilising the Micro-Political Voice: Doing the 'Human Microphone' and the 'Mic-Check'. *Journal of Language and Politics* 16(1): 110-136.

McKinney, Joslin & Palmer, Scott (Eds.) (2017). *Scenography Expanded: An Introduction to Contemporary Performance Design*. London: Bloomsbury.

Mengis, Jeanne, Nicolini, Davide & Gorli, Mara (2016). The Video Production of Space: How Different Recording Practices Matter. *Organizational Research Methods* Epub ahead of print. [Online]. Available: <http://journals.sagepub.com/doi/abs/10.1177/1094428116669819>. [Accessed August 28, 2017].

Schröter, Jens (2014). *3D: History, Theory and Aesthetics of the Transplane Image*. London: Bloomsbury.

Tosi, Virgilio (2005). *Cinema Before Cinema: The Origins of Scientific Cinematography*. London: British Universities Film & Video Council.