Journal of Human Rights Practice, 2018, 1–20
doi: 10.1093/jhuman/huy005
Policy and Practice Note

Policy and Practice Note

Reconstructing Human Rights Violations Using Large Eyewitness Video Collections: The Case of Euromaidan Protester Deaths

Jay D. Aronson, McKenna Cole, Alex Hauptmann, Dan Miller, and Bradley Samuels*

Abstract

The widespread availability of mobile phones with high quality cameras means that events around the world can be live streamed or captured on video and rapidly shared via social media. Because this video is multi-perspectival, it can tell the story of an event from many different vantage points, providing a synthetic and composite form of documentation that has the potential to enrich our understanding of events of interest. While video has the potential to provide valuable information, variability in recording platforms and metadata can make a large video archive complex and very difficult to analyse. This paper describes a platform developed by a multidisciplinary team to organize and analyse a large collection of event-based video. It also explains how the system is being deployed to aid in the investigation of allegations of abuses by security forces during the 2013–2014 Euromaidan Protests in Kiev, Ukraine. This platform includes a video archiving system, semi-automated tools for video synchronization and geolocation, and visual interfaces for exploring video data. This system will be useful for the investigation and analysis of protests and demonstrations, mass government repression, police brutality, conflict events, and disasters. The paper concludes by noting that video—even in high volume—does not tell the entire story of an event. As with all other forms of evidence, it must be combined with other available data and relevant knowledge in order to provide a nuanced understanding of what has taken place.

* Jay D. Aronson (aronson@andrew.cmu.edu) is the founder and director of the Center for Human Rights Science, and associate professor of science, technology, and society, at Carnegie Mellon University. Alex Hauptmann is a principal systems scientist in Carnegie Mellon University's Language Technologies Institute. Bradley Samuels is a founding partner at SITU Research, an interdisciplinary practice working in design, visualization, and spatial analysis in the service of human rights reporting and fact finding, where McKenna Cole and Dan Miller carried out research, design, and software development on this project.

Keywords: collaboration; computer science; data visualization; human rights documentation; social media; video analysis

Project background

In the era of social media, widespread mobile phone coverage, and the availability of Internet access in the majority of the world, user-generated video has become an important dimension of conflict monitoring and the documentation of war crimes and human rights abuse (Padania et al. 2011). Faced with a proliferation of content, practitioners are increasingly finding themselves overwhelmed by vast quantities of visual evidence. To date, most information extraction from video related to conflict and human rights has been accomplished manually (Koettl 2016; Silverman 2015b). An investigator will review each relevant video individually, noting whatever particular attributes are of interest. When original videos are obtained, technical metadata such as location and time stamps can be extracted from the file. When videos are obtained through social media, this information is stripped and must be gleaned from clues in the footage.

Once extracted, data can be expressed in a prose summary or as entries in a database. Such analysis is vital, but time-consuming and very expensive if people have to be paid to do the work. It is also emotionally challenging to watch numerous videos and extract information from them if the data being gathered deal with issues such as torture, rape, mass death, or extrajudicial killings (Dubberley et al. 2015). Additionally, language skills or the need for local knowledge can limit the number of researchers who are capable of carrying out such work.

The same characteristics that make video difficult to work with—its volume, its lack of structure, and its diverse sources—also make it a potentially rich source of evidence for human rights practitioners. Because this video is multi-perspectival, it can tell the story of an event from many vantage points rather than a single authoritative view. This enables investigators and ordinary people to understand a situation from multiple points of view—in some cases lending credibility to one particular account of what happened, in others demonstrating that a variety of accounts are plausible, in some cases suggesting that no extant account gets the story right, and in others generating a new narrative. The full potential of multiple video accounts of the same event is that it makes possible both a spatial (3D) and temporal (4D) reconstruction of an event.

Visual evidence has become an increasingly important aspect of advocacy, journalism, and law in recent years (Feigenson and Speisel 2009; New Tactics in Human Rights 2014; Wardle et al. 2014). Video, for instance, has changed the conversation about police brutality in the United States (Cave and Oliver 2016; Uberti 2015) and has been used to investigate human rights abuses and war crimes around the world (Aronson 2018). Claims of self-defence, active resistance, or defying orders that are often used to justify lethal shootings by police and military personnel are now being called into question as bystanders and victims are filming these encounters and sharing the video with media outlets and the public. These video accounts sometimes complement, and sometimes show a different perspective, compared to the footage captured by the police body worn or dashboard cameras. Video evidence is also being used in a variety of other domains, including environmental justice, labour law, and monitoring gender-based violence.

Visual evidence brings with it numerous challenges and potential pitfalls (Feigenson and Speisel 2009). Putting aside the obvious concern that video can be faked or manipulated, even authenticated video is an incomplete, mediated view of the world.[1] It shows only a very narrow slice of what is happening in a location, generally over a short period of time, and may give a false sense of omniscience. Large volumes of multi-perspectival video can help overcome this challenge, but this still does not guarantee complete, unbiased coverage. Further, even if there is a reasonably complete video archive of an event, the actions depicted are open to multiple interpretations and must still be placed in context to be fully understood. In the case of police brutality, for instance, it is important to know about previous interactions between the police unit involved and the person or people in question. For instance, was the interaction being recorded a continuation of a set of events in which the previous events were not captured on camera? Had either party issued threats that were not recorded? Is the police unit in question normally restrained, or was the violent event captured typical of their behaviour? The answers to all of these questions depend in part on the visual record, but also on one's assumptions and interpretive frameworks. The same video or videos can be used to substantiate vastly different claims, as Bill Nichols shows in his analysis of how different social groups interpreted the circumstances and video recording of the 1991 Los Angeles Police Department beating of Rodney King (Nichols 1994).

There is also a long history of concern about the prejudicial effect that images and video might have on legal fact-finders and decision makers (Feigenson and Speisel 2009). In most cases, some sort of expert judgment is necessary to understand what is taking place. Projectiles such as bullets might appear to come from a particular direction based on visual cues, such as how a person falls after being hit, but ballistics experts are crucial to understanding the path of the projectile and determining where it may have come from. Police or protesters may appear to be doing something in a video, but only individuals with local knowledge or first-hand experience of the conflict can explain what is taking place and why it is happening. In other words, video evidence does not replace expert knowledge and eyewitness narratives in legal contexts (Whiting 2015). Rather, the proper use of such evidence requires expertise and testimony for it to be appropriately contextualized, and it will often be subjected to adversarial or inquisitorial questioning.

Video evidence is undoubtedly persuasive and compelling in a world where 'seeing is believing', but it cannot tell the entire story of an event. Video evidence is not an unmediated window into the 'facts' of a case (Landman and Carvalho 2009) or a virtual witness to a crime that can be viewed by a naïve fact-finder without analysis (Feigenson and Speisel 2009; Shapin 1984). As with all other forms of evidence, it must be viewed with healthy scepticism and combined with other available data, expert knowledge, and eyewitness testimony in order to provide a nuanced understanding of what has taken place. It is also crucial to remember that not all human rights violations will be caught on video and in many cases, an event will only be filmed by a single witness, limiting the value of the tools and methods discussed below.

Overview

This paper describes a platform developed by a team of computer scientists, designers, and humanists to organize, analyse, and present large volumes of event-based video. This

1 On faking or manipulating video, see e.g. King et al. (2015); Silverman (2015a).

platform includes a video archiving system, algorithms for video synchronization and geolocation that do not require metadata, and video-based spatial analysis in a straightforward and useful manner. After providing a general overview of the platform and workflows, we will present the case study that generated much of the work on this tool: a request by human rights lawyers in Ukraine to help make sense of several hundred videos depicting clashes between protesters and riot police during the Euromaidan protests in Kiev on the morning of 20 February 2014 that may involve illegal use of force by law enforcement officials.

Event reconstruction

The goal of video-based event reconstruction is to be able to construct a visual and spatial narrative of an event that is as comprehensive as possible. Ideally, investigators should be able to gain detailed information about what happened at the time and place that an event occurs, but also what happened before and after the event in question, as well as what was taking place in the geographic vicinity of the event (Weizman 2017). At the most basic level, the reconstruction involves building an archive of videos that are geolocated and time stamped based on a universal clock. A more complex reconstruction allows the investigator to play multiple video accounts of the same event side by side in a synchronized fashion. This work builds on a long tradition of forensic crime scene reconstruction, but relies on actual video, not animation or simulation, to render an event over time.

In an effort to move beyond these limitations, we are exploring various ways of extracting information from multiple videos and presenting it in a more digestible fashion. One method that we describe below is to depict people, or groups of people, on a map in two dimensions and dynamically model their interactions with other people or groups. For instance, in a clash between protesters and police, mapping makes it possible to understand how these two groups engaged with one another over time and space. This map can be linked to the video collection in such a way that the viewer can see an overview of the action on a map, and then view videos that reflect what is happening on the ground at a particular time or place.

Another form of event reconstruction we use, building on previous work done by Forensic Architecture and SITU Research, is to select key frames from one or more videos that enable modelling of particular features of interest, such as the trajectory of a projectile or the physical behaviour of an individual hit by a projectile (Forensic Architecture and SITU Research 2010). It may also be possible to extract details of forensic and legal interest from these frames. Some examples include the presence or absence, and spatial arrangements, of particular objects (for example, weapons or tear gas canisters) or people; the directionality of blood spatter of a person hit by a projectile or blunt object; or the angle of entry of a bullet into a victim. Event reconstructions can also confirm or refute details of an autopsy or crime scene report of a person whose death or attack was caught on video. We also explore the possibility of using various computer vision and 3D modelling systems to develop digital reconstructions of particular events. In some cases we are doing so in static fashion in order to understand the positions of various actors at the time of death, and in others we animate these reconstructions in order to help fact-finders reach conclusions about what happened at a specific place over time. The type of analysis we describe here is best understood as a composite approach—no single video or piece of data unlocks a case, but when examined together new understandings can emerge.

Euromaidan

In this section, in order to provide context for our technical work, we provide a brief overview of the Euromaidan protests, which began in central Kiev on the night of 21 November 2013 and climaxed in a wave of violent clashes between government forces and protesters in mid-February 2014. The initial protests in Independence Square (Maidan Nezalezhnosti, or just 'Maidan') were a response to the government's decision to suspend preparations to sign a formal economic and political partnership with the European Union (EU) at the upcoming Eastern Partnership Summit taking place in Vilnius, Lithuania. Talks on the EU association agreement had been going on since 2007. As the time to sign got closer, President Viktor Yanukovych and his supporters became concerned that such an agreement, while enhancing ties with the EU, would alienate Russia. In the years since the collapse of the Soviet Union, Ukraine had become increasingly dependent upon Russia for cheap natural gas, loans, and other financial necessities due to its poorly planned, corrupt, and inefficient economy (Menon and Rumer 2015). Protesters were equally angered over what they perceived to be rampant corruption within Ukrainian politics, decreasing quality of life, and increasing authoritarianism within society—especially in attacks on the independent judiciary and civil society (Matviychuk and Pavlichenko 2015).

The demonstrations began peacefully on the nights of 21 and 22 November with no more than a few thousand people (many of them students) participating. By 24 November, though, tens of thousands of people had joined the protests. While most protesters were out to express their political views about European integration that night, a small contingent attacked the Government Building, which houses the Ukrainian Cabinet of Ministers, on the square. This transgression led police to strike back, and the situation devolved into violence, with police and protesters both using tear gas in the fighting. According to news sources, the protesters also deployed firecrackers to repel police (Matviychuk and Pavlichenko 2015). These engagements continued sporadically over the next few days. On the morning of 30 November, a day after the Ukrainian government ultimately failed to sign the EU Association Agreement, the government ordered an elite police unit, the federally controlled Berkut ('Golden Eagle') special forces,[2] to disperse protesters from Maidan. Berkut personnel chased, beat, and kicked protesters who posed no threat to them, and they also attacked civilian bystanders. The stated justification for their action was to enable the decoration of the square for Christmas and New Year (ibid.). Later in the day, the Minister of Internal Affairs admitted that the Berkut overstepped their bounds, and promised an investigation. All told, nearly 80 protesters and seven police officers were injured in the clashes. That night, opposition groups organized another rally near Maidan, with thousands of people participating. Protesters and activists also began forming self-defence units called 'hundreds' to support the protest activities and also set up medical centres to treat wounded participants. These actions continued around Maidan over the next two days (ibid.).

2 The Berkut unit was initially set up in 1992 as a special forces unit to fight organized crime but was reoriented towards riot and protest control. Berkut commanders reported directly to federal Interior Ministry officials, and officers were better trained and paid than regional police personnel. According to the BBC, there were between 4,000–5,000 Berkut in Ukraine at the time they were disbanded in 2014 (BBC News 2014).

Beginning on 2 December and continuing through 6 December, protesters settled into Maidan for the long haul even though the government had declared a ban on all protests in central Kiev. They set up tents, a stage, and video monitors, and fortified their positions with barricades. Observers described the atmosphere as generally calm and hopeful, even festive. Opposition leaders gave speeches, musicians entertained protesters, and protesters gathered around campfires to keep warm. Organizers and participants believed that peaceful political change was possible (Matviychuk and Pavlichenko 2015). Over the next few days, Parliament continued to debate various no confidence motions brought against the government, and protesters added to their fortifications at Maidan. The government told protesters to clear blockades of government buildings, and complained that the media was giving them too much positive coverage.

From 8 to 11 December, protesters and police clashed sporadically in the wake of economic aid talks between Yanukovych and Russian leader Vladimir Putin, which culminated on 17 December with a signed agreement that included financial support from Russia, concessions on gas prices, and the restoration of customs regulations regarding the import of Ukrainian goods in Russia that had been removed in the run-up of Ukraine's negotiations with the EU. The anti-government movement gained significant support during this period. Some 200,000 protesters celebrated New Year's Eve at Maidan and throughout the first two weeks of January, tens of thousands of protesters (the numbers vacillated during the period depending on the news and planned activities) maintained control of the square, clashing with police from time to time.

On 15 January, courts declared the protests illegal, and the next day Parliament passed a series of anti-protest laws that essentially criminalized all activities taking place at Maidan, making it impossible for opposition groups to organize or even to openly criticize the government. These laws included a ban on concealing one's face with a mask or helmet at a public gathering or protest (Polityuk 2014). These political manoeuvres suggested to opposition leaders that the government was planning to use the Berkut and other police units to forcibly clear the site once and for all. Clashes between protesters and government officials became increasingly violent over the next two weeks and President Yanukovych said that he would use 'all legitimate means' to restore order if a political compromise could not be reached. 'Titushky', violent pro-government civilians who had Yanukovych's support, also took to the streets to confront and intimidate protesters (Matviychuk and Pavlichenko 2015; Goncharenko 2014; US Department of State 2015).[3]

On 18 February 2014, the opposition's anger turned into action. Armed with a variety of weapons and explosives, protesters marched on Parliament demanding constitutional reform and new elections. For the first time, police fought back with guns (using both live ammunition and rubber bullets), tear gas, flash grenades, and blunt force weapons. They also attacked opposition encampments at Maidan. More than two dozen people were killed that day, including ten Ministry of Internal Affairs soldiers from the 'VV unit'. Hundreds of people were injured in the fighting. The situation would only get worse over the next two days, as the government redoubled its efforts to gain control of Maidan and explicitly

3 Activist groups call the titushky/titushki 'criminals', 'mercenaries', or 'thugs', but in this paper, we prefer the neutral term civilian to denote that they were not directly affiliated with the government. According to news reports and the US State Department, they have been involved in beatings, kidnappings, carjackings, and were reportedly paid by the Yanukovych government to intimidate its political opponents.

authorized the use of assault weapons with live ammunition on protesters. On the other side, opposition leaders felt that they were on the verge of changing the course of history. The stage was set for a massive and deadly standoff on the morning of 20 February. It is this event that we were asked by Ukrainian human rights lawyers to analyse in detail.

Events of 20 February 2014

Between 5 a.m. and 9 a.m., three VV unit soldiers were killed and 20 VV unit soldiers sustained injuries in fighting over control of the Academy of Music, where protesters had set up a medical station. At approximately 9 a.m., a police officer from the Berkut unit was killed as protesters began to push the police units away from Maidan. Perpetrators have not been identified, but the direction of where the shots were coming from suggests that they were made from the side of protesters. Between 9 a.m. and 12:30 p.m. at Instytutska Street, 47 civilians sustained fatal gunshot wounds at the hands of Berkut and other Interior Ministry Force personnel, and that morning approximately 200 people sustained injuries (Open Dialog Foundation 2014). An additional person was fatally shot later in the afternoon, bringing the total of deaths to 48 (Matviychuk and Pavlichenko 2015: 80).

Liberal Ukrainian political activists, human rights groups such as No Borders Ukraine, and the lawyers of families of victims who requested our assistance believe that the government's use of violence at Maidan was unjustified and are demanding accountability. They note that protesters were generally armed with little more than large metal or wooden shields and helmets and did not have any deadly weapons. Thus, in their view, those protesters who were shot or killed did not pose an immediate deadly threat to police officers and special forces personnel. According to rights groups, government attacks were 'widespread and systematic', and 'directed at a group of [nonviolent] citizens according to their political views characteristics, in this case—disagreement with actions of authoritarian regime' (Matviychuk and Pavlichenko 2015: 15 and 22). Moreover, the violence was 'a deliberate policy of suppression of the protest movement by any means available' (ibid: 21). Such tactics, in their view, violated OSCE Guidelines on Freedom of Peaceful Assembly as well as numerous other legal frameworks that support the right to assembly, including Article 20 of the Universal Declaration of Human Rights and Article 11 of the European Convention on Human Rights (ibid: 25).

The police and security force personnel, on the other hand, argue that they were being threatened and fired upon by protesters and snipers in the Hotel Ukrayina and other locations, and were firing to protect themselves and to regain order. Ultimately, the case focuses on addressing the question of whether excessive force was used and whether protesters killed by police and security force personnel were an immediate and direct threat. In addition to the specific legal questions addressed by the attorneys in the legal context, the work also provides a portal to bring into focus specific actions and events of an incredibly complex, chaotic, and dynamic day. The ambition of this work is to present a tool that can be used to pursue truth and accountability in the face of the many competing narratives that surface after conflict and human rights violations.

Mandate

The Ukrainian legal team that contacted us is representing families of protesters killed by a particular group of 25 Berkut officers active during the morning of 20 February 2014. Four

of the officers and a low-level commander are currently being held in a Ukrainian jail while on trial, and the others have fled the country and are living in exile in Russia, where several have been granted Russian citizenship. All members of the Berkut unit and their commander were initially charged with the murder of 47 civilians and the attempted murder of 80 others. Later, the charge of terrorism, defined in Ukrainian criminal law as killing with a weapon with the goal of intimidating the public, was added to the list. The private attorneys who contacted us are acting alongside, and in partnership with, Ukrainian prosecutors who are representing the state (that is, the Prosecutor General of Ukraine). The case is being heard in the Kievo-Svyatoshinskiy District Court of the Kyiv region—case no. 59/3498/15к (1кп/759/43/16)—by a hybrid decision-making body composed of two professional judges and three lay jurors. The trial began in March 2016, but it is unclear when it will end due to the scope and complexity of the forensic studies and the number of witnesses that need to be questioned. More than 400 such witnesses have been identified to date, ranging from eyewitnesses and survivors to forensic specialists to high-level government officials including former president Yanukovych, the former Minister of Internal Affairs, and the former commander of the internal security troops at the time of the killings.[4]

The legal team's case rests on determining whether or not the protesters posed a direct threat to security force personnel in a way that would legitimize a lethal response. They asked us to develop a system that would allow them to track specific groups through all extant videos, including various categories of civilians, persons in uniform, and members of the press. Further, they asked us to carry out an event reconstruction for specific moments from the protests on 20 February and glean as much information as possible about when killings or significant injury to civilians took place; the area/location where the individual sustained an injury; the location from which the shot that injured or killed the person was most likely fired, and, if possible, who fired the shot; and who was in the area at the time the killing/injury took place. Additionally, they asked us to analyse the movements of various civilian and uniformed personnel over the course of 20 February in order to better understand the context in which uniformed security personnel used deadly force against protesters.

For videos showing the moment of injury of a person of interest, the legal team asked us to develop a reconstruction that operates at a range of scales—from the urban scale that helps illuminate the relative positions of the actors all the way down to the corporeal scale of the body that will be useful in representing information provided by the autopsy reports (entry and exit wounds, for example). In cases where video captures munitions being discharged, the legal team asked us to determine whether their trajectory can be visualized and whether videos could aid ballistics experts in making such a determination. When a video does not show an injury but has the moment of firing, the legal team asked us to identify the area at minimum, and the specific gun if possible, from which the shot could have been fired, as well as the trajectory and potential targets of the bullet. This information can then be used to augment, verify, or rebut information gleaned from other sources such as witness testimony, medical reports, media accounts, or GPS location information from mobile phones of suspects (if available).

Further, the legal team asked us to develop a visualization platform that would enable them to identify and isolate the movements of particular groups or individuals over time and space. A key component of presenting this information was the ability to link source content that served as a basis for reconstruction (that is, the name of the original video,

4 Trial information supplied by attorneys Pavel Dykan and Alexandra Yatsenko on 1 June 2017.

timing for the event in the video, and potentially also other text, including links to medical reports, and so on). Although our services were requested by lawyers representing families of victims, we are submitting our findings directly through the court. We will make as much of our report as possible available to the public (while taking steps to mitigate any risks to privacy and safety of those depicted in the videos), and will be clear about the limitations of our work—most importantly that our conclusions are limited by the videos that we have access to and were able to synchronize and geolocate.[5]

Video evidence base

The attorneys shared with us approximately 65 hours of video of the events that took place around Maidan on the morning of 20 February. In order to reconstruct key events of interest such as troop movements, clashes, or shootings, all available video had to be geolocated and synchronized to a global time clock. In this case, the vast majority of the video in the attorneys' possession had been retrieved from social media and other Internet sources, and therefore lacked location and time stamps. The attorneys initially requested the assistance of an analyst who was intimately familiar with the events of Euromaidan and the geography of Kyiv. Working over the course of eight months, this individual was able to stitch together a small percentage of the total video using visual and audio cues in the recordings. This work was incredibly demanding because videographers tended to be relatively far away from one another and focused on different points within a larger scene. Smoke, combined with a drab, relatively monochromatic environment and many people wearing very similar clothing, made the task even more difficult. Often, there are no obvious visual cues that two videos are synchronous. The analyst put this work together in a 9-channel video (that is, videos of the same moment in time displayed in a three-by-three grid) that encompasses 4 hours 37 minutes and 19 seconds, with large gaps throughout (See Fig. 1).

While this work produced an impressive synchronization of a portion of the video content, it became clear that it was prohibitively time consuming to analyse all of the available footage in this manner. Recognizing that this manual synchronization approach was helpful but not comprehensive, the attorneys reached out to their partners in the international human rights community for assistance, who introduced them to us. We quickly realized that developing a metadata-independent approach to making sense of large volumes of video would be useful not only in the context of the Euromaidan investigation, but also for many other public events that were in the news at the time (such as the Arab Spring, Ferguson, Gezi Park, Baltimore, Paris, and so on). As such, we decided to use this case as a test-bed to understand what sorts of techniques would be needed to be able to organize large volumes of video over time and space, and then present it in a visually and cognitively compelling way.

Workflow

During the first nine months of the project, SITU Research and Carnegie Mellon University's Center for Human Rights Science (CHRS) worked on parallel tracks to create

5 We did not verify the authenticity of each video we analysed in this case, although we would certainly do that if we were conducting an independent investigation or if we collected the material ourselves.

Figure 1. Still excerpt of the 9-channel video grid showing manually synchronized video from various perspectives and locations in and around Maidan during the morning of 20 February 2014

the systems necessary to fulfil our mandate. Building on institutional strengths, SITU focused on the platform design and event reconstruction portion of the project, while the CHRS team focused on the challenge of developing semi-autonomous methods for synchronizing video over time and geolocating the camera. While both groups had distinct roles aligned with their core competencies, a robust feedback loop was established to share progress at regular intervals that informed the development of the project overall—in this sense it was an interdisciplinary collaboration that produced something greater than either institution could have accomplished alone.

Semi-autonomous synchronization and geolocation

The systems created by CHRS for synchronization and geolocation of video are semi-autonomous, not fully autonomous. This means that the algorithms developed do not make the final determination about where videos are filmed, and which were filmed at the same time and place. Rather, they assign probabilities that two or more videos were filmed at the same time, and generate a list of the most likely known locations where a video was filmed. Human analysts then review these probabilistic determinations through two separate custom-built user interfaces to confirm which ones are likely to be correct and which ones are erroneous. They can also review those videos that the algorithms were unable to synchronize due to sound quality or lack of distinct sound signatures, or were unable to geolocate due to poor quality or lack of distinct visual reference points. Thus, our systems do not lessen the need for human judgment—indeed many videos simply do not contain enough useful information to allow the system to generate reasonable locations or synchronization. They do, however, allow human analysts to be significantly more efficient in synchronizing and geolocating videos because they only have to sort through a few potential matches, not dozens or hundreds, for many videos.

Synchronization process[6]

We received 520 videos from the 20 February Euromaidan protests, totalling approximately 65 hours. Because many of the videos were edited combinations of many scenes, the videos were segmented into unique, unedited segments using a computer vision algorithm that looks for scene changes. For computational reasons, all clips under seven seconds were excluded, leaving 4,537 clips which totalled 52 hours and 24 minutes. Much of this footage was shot on the ground by protesters, bystanders, and photojournalists, while some was filmed from buildings or other aerial positions. Later in the project, the legal team also sent CCTV footage from a bank near Maidan, although this evidence had time stamps and we knew exactly where the camera was mounted.

As noted in Fig. 2, visual cues could not be used to synchronize this vast quantity of metadata-free video because most videos shot at the same time were shot from very different perspectives. The videos in the collection were rarely shot side by side at the exact same time. Rather, videographers tended to be dispersed over space and each one was focusing on something slightly different. This is not surprising given the size and scope of the protests. Had the event been very focused and the location compact, visual synchronization may have been possible.

As such, the CHRS team created a unique sound print for each clip using an algorithm that recognized a standardized vocabulary of 'features' (such as wind, screaming, gunshots, airplane noise, amplified speeches, music, and explosions). The algorithm then compared the sequence of these features in each clip to all others and looked for reasonable matches (see Fig. 3).

In comparing all of the videos in the collection against one another, we had to include a certain amount of tolerance for difference because each recording device captures sound in slightly different ways, and the actual characteristics of the sound will differ based on the distance from the source and the immediate environment of the camera. When a high-probability match was made, we confirmed the synchronization manually to ensure that it was reasonable. Because sound travels more slowly than light, there was often a slight offset when two videos were shot at different distances from the sound. As such, CHRS was able to manually encode an offset into videos shot at the same time so that they could be visually aligned when shown together. All told, CHRS was able to synchronize 4 hours 16 minutes and 13 seconds of video, some of which overlapped with the manual synchronization described above. Through these efforts, the combined total of synchronized video was increased approximately 50 per cent, to 6 hours 52 minutes 10 seconds. We also determined that 10 hours 40 minutes 2 seconds of the video we analysed were duplicates of other video in the collection.

Geolocation process

The CHRS team also developed a semi-automated method of geolocating videos. After doing some experimentation, we decided that the most efficient approach would be to mine the Internet for accurately geotagged images of major landmarks taken from known locations around Maidan to create the basis for rapidly comparing scenes from the video collection. CHRS created an algorithm that could compare these images—in this case we used Google Street View and Flickr because of the high probability of accurate geotags—with key frames taken from the Euromaidan protest video. This method also has the advantage

6 For a technical overview of the process described in this section, see Liang et al. (2016).

Figure 2. Close-up view of the same moment in time and place taken from two different perspectives, demonstrating the impossibility of synchronization using visual cues

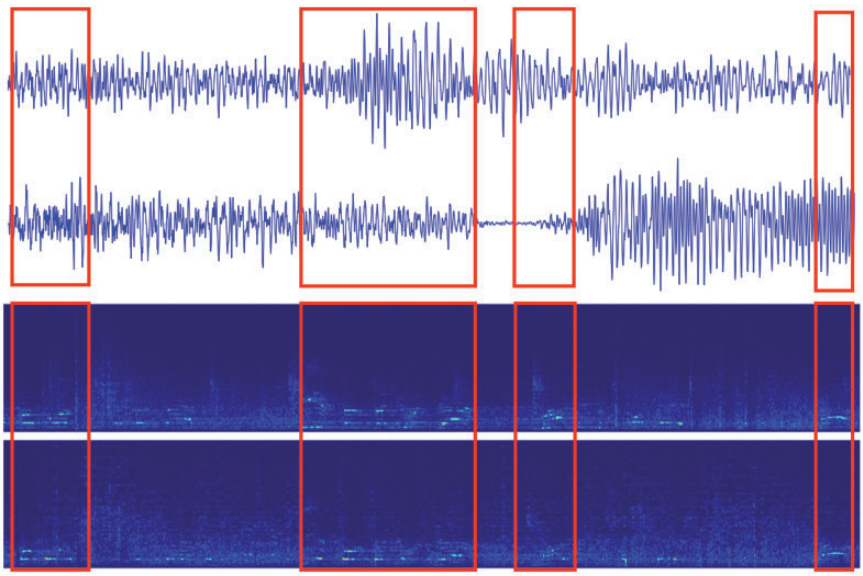

Figure 3. Sound print from the two videos in Fig. 2, demonstrating the utility of sound cues in synchronization. Each boxed section represents an instance of a predefined audio feature, such as wind, explosions, gunshots, loudspeaker, music, etc.

of being reproducible anywhere one can find geotagged images of the space in which an event of interest takes place.

In order to allow human analysts to quickly confirm or reject a potential match, CHRS developed a system that enabled side-by-side comparison of the most likely locations of a given video clip and a key frame from that clip (see Fig. 4). The human analyst either confirms or rejects the match. If rejected, then the next most likely location image shows up and the human analyst can accept or reject that match.

It is important to note that the system is not always able to come up with a reasonable locational match. It was able to geolocate approximately 12 per cent of the videos in question, determine that 10 per cent were shot indoors, and that 11 per cent contained no information that could be used for geolocation. The first version of this system failed to

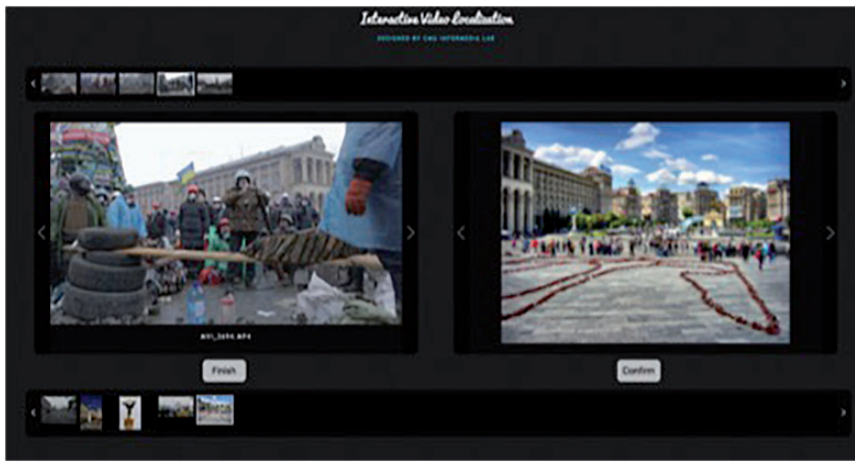

Figure 4. Screenshot of the geolocation tool developed for this project

determine the location of approximately 66 per cent of the videos, which can likely be improved by further refinement of algorithms.[7]

Platform design and development

Once the CHRS team determined the location and time of the videos delivered by the Ukrainian legal team, this information was sent to SITU Research in a database to be processed and converted into a functional video archive. The primary use for this archive would be gathering information about the killings being investigated by the Ukrainian legal team. The SITU team was motivated not just by this particular case, but structural questions of how to best create a tool to leverage large amounts of citizen documentation gathered during a mass event in an urban context. The sheer volume of videos from 20 February 2014 necessitated a critical reflection on workflows—SITU needed to develop an approach to automating, archiving, hosting, and analysing the relevant data.

Driven by these questions, SITU designed an interactive platform to host the data in a format that could ultimately be used by legal fact-finders to better understand the interactions of the killed protesters and Berkut personnel: a curated diagrammatic timeline of the day serving as a useful tool for supporting a more detailed, spatial analysis of specific events. The resulting platform is a digital, interactive application that presents specific evidence and analysis of localized events in a format that allows the user to also interact intuitively with a vast, structured media archive.[8] The platform is structured in four sections: Introduction, Timeline, Analysis, and Archive. Each section plays a distinct and complementary role in the reconstruction and presentation of the events of 20 February.

7 These results are based on currently unpublished data that is available upon request from the authors.

8 The platform will be made public after it has been formally introduced into evidence at the trial in late 2018.

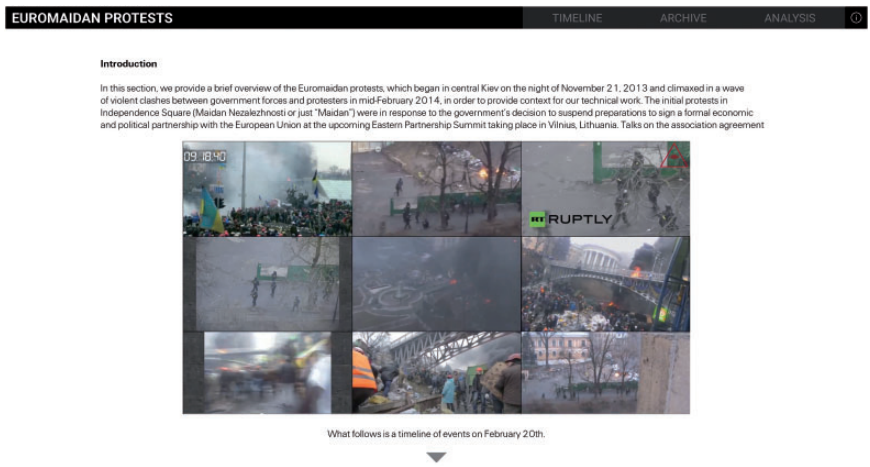

Figure 5. Screenshot of the Introduction section of the platform

Platform components and functionality

The platform is a tool that provides a concise narrative of what is otherwise an overwhelmingly complex sequence of events. It functions as an archive that hosts all of the video assets that document the period of interest—an intuitive interactive tool for accessing the hundreds of videos that captured the events of that day through either spatial or temporal queries. It also presents analysis of specific moments that have been studied in greater detail. The nature of both collective action and conflict in urban environments is often characterized by complexity and seeming disorder. Reconstruction of these events is challenging and can hinder analytical efforts. The ultimate ambition of this work is to make relevant documentation and the analysis it enables more readily accessible in similar contexts where human rights violations may have been perpetrated.

The first section of the platform provides an overview of the case, the legal brief, and the evidence/data used (Fig. 5). In this section, narrative text gives a context for the protests at Maidan and goes into greater detail for the sequence of events as they unfolded on 20 February 2014. Following this overview, the brief is outlined providing the specific objectives for the case as defined by the Ukrainian legal team.

The second section introduces a linked timeline and map interface (Fig. 6). The map provides a view of Maidan Square and its immediate surrounds. A timeline below the portal begins at 8 a.m. and extends to 12:30 p.m.—as it was during this time that all but one of the protester deaths occurred on 20 February. Using arrows to advance the timeline, the viewer is taken through the key developments for the time period in question. From troop and protester movements to locations and durations of temporary barricades that were key to understanding the nature and positions of the fighting that occurred, this timeline is a distilled chronology that provides the user with an understanding of key events in space and time. Javascript tools were utilized to draw features with accessible attributes and animate them in sync with the timeline below. As one advances the narrative, the timeline moves to the next event and an animated sequence is initiated on the map above. A short narrative text also appears with each event.

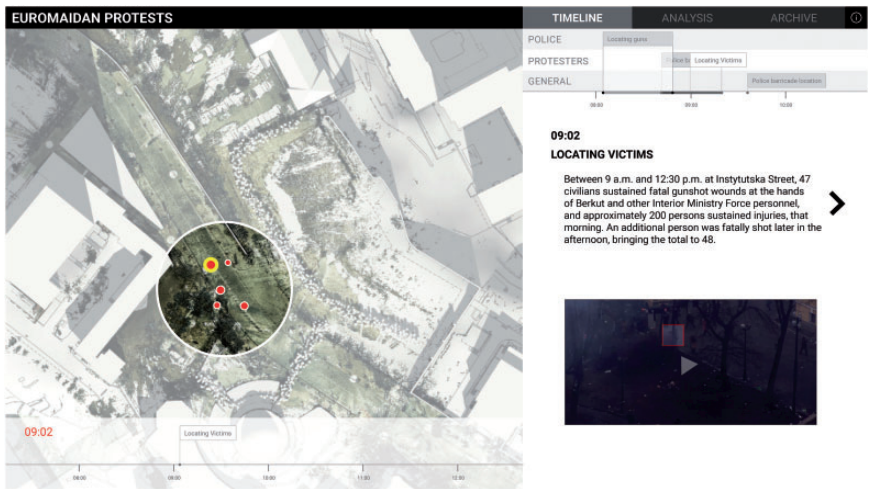

Figure 6. Screenshot of the Timeline section of the platform

The Analysis section of the platform synthesizes components of the case into evidentiary narratives for protester deaths that were prioritized by the legal team. Medical reports, expert testimony, relevant videos from the archive, and event reconstructions in the form of analytical animations are presented alongside a map of the events. The page is designed to be a useful tool for courtroom presentation, making accessible the underlying data contributing to the event reconstructions. Because each event may have different evidentiary assets associated with it, users can click on the different events on the timeline to pull up the analysis specific to each. The structure of this section and its focus on a specific moment in time and space reflects a strategy of identifying and isolating key moments within larger complex events for increased scrutiny and analysis.

The next section, titled 'Archive', retains the basic map interface and timeline structure (Fig. 8). Here however, the platform moves from the curated content of the previous section to a format that acts as an archive for the totality of the video assets available for the time period in question. The geolocation and global time stamping data produced by the CHRS team's semi-automated system is fed into the platform's underlying database and visualized within the user interface to provide an accessible and comprehensive catalogue of all video assets. The timeline is expanded into a field that represents unique video assets as horizontal lines—the starting point of each aligned with the time on a global clock, the length of each corresponding to the duration of the video. A magnified view of the highlighted section of the timeline appears, allowing users to view videos of interest or automatically sync a series of videos that correspond through a specific duration of time. In addition to appearing on the timeline, the positions from which videos were taken are located the map.

With all video assets appearing on both the map and the timeline, users can query any content based on location and/or time of event. This functionality also allows the user the ability to quickly cross-reference video content capturing the same moment in time, making multi-perspectival event reconstruction possible. In addition to the ability to query all video assets through multiple access points, key events in time and space are layered on top of the

Figure 7. Screenshot of the Analysis section of the platform

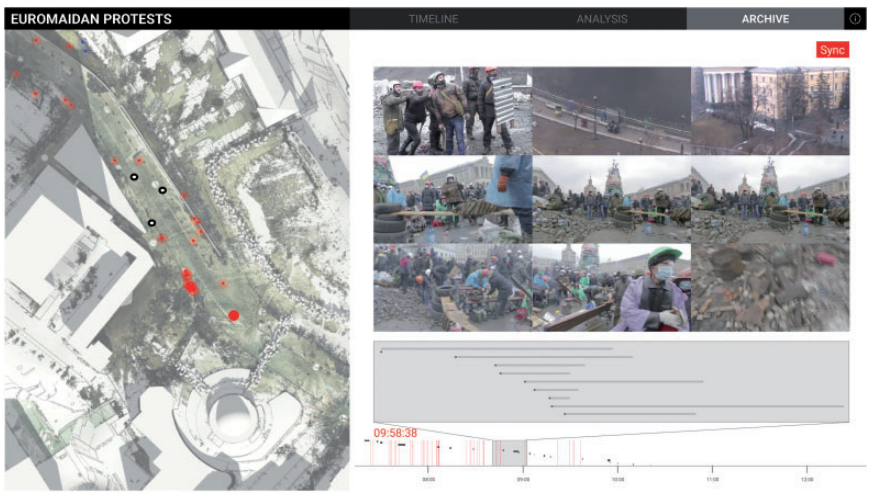

Figure 8. Screenshot of the Archive section of the platform

data. All times and locations of known civilian deaths are placed on both the global timeline and map providing nodes around which users may want to focus inquiry or attention. Any of these nodes may be clicked to access more attributes (for example, name and photo of the individual, as well as relevant primary documents such as autopsy reports). While the Timeline section was linear and narrative in nature, the Archive section offers an entirely interactive tool, allowing the user to query across time and space non-hierarchically.

Conclusion

The platform and associated methodology described here makes it possible for investigators and advocates interested in justice and accountability to harness the informational potential of large event-based video collections in a way that is not possible using strictly manual analysis methods. The human investigator is not taken out of the loop, but rather is able to work in a more efficient and productive manner, and can more effectively communicate their findings to decision makers and fact-finders.

In the case described in this paper, for instance, we are able to reconstruct interactions between protesters and police officers, determining relationships among various groups over time and space, as well as the most likely places from which fatal shots were fired. When combined with other available information, eyewitness accounts, and expert testimony, these reconstructions can provide evidence that attorneys and fact-finders can use to determine where lethal shots were fired from and whether or not particular killings were legally justified. Video reconstructions can also help corroborate or refute other evidence or be used to provide context that might help in the sentencing phase of a trial (for example, whether perpetrators had faced threats before killing protesters, or whether perpetrators had sought to intimidate other protesters in the period leading up to the killings). While we have described the use of this capacity in a legal context, it can also be used in other kinds of advocacy and accountability work. Some examples include advocating for reform of law enforcement practices when instances of police brutality or suppression of the right to assemble are filmed, or corporate policies when labour violations or environmental degradation are caught on camera.

Ultimately, the greatest strength of the platform we describe here is the ability to analyse large volumes of video footage and integrate it with other forms of evidence in a package that can be viewed and understood by non-specialists. Our platform enables visualization of evidence from the level of the built environment (geographic space and the interaction of groups) to the level of the individual (for example, the moment of death and the autopsy report that describes its characteristics). This capacity was only made possible by the collaboration of computer scientists, designers, legal professionals, social scientists, and most importantly, activists with detailed knowledge about the events that occurred and the space in which they took place, to provide spatial and temporal guideposts upon which the reconstruction could take shape.

It remains to be seen how this platform will be utilized in court, how persuasive it will be for fact-finders, and how the defence may challenge its validity (Accatino and Collins 2016). We can say with some certainty, however, that the increasing availability of video evidence will require that the human rights community integrate video analysis into its fact-finding protocols and advocacy strategies. For the bulk of its existence, human rights documentation has relied heavily on eyewitness testimony for data, and the written report as its core mode of information transmission (Alston and Knuckey 2016). Visual elements, like quantitative analyses, have typically served an auxiliary role—they have not been the primary means for communicating the existence of human rights violations. This trend has begun to shift, but there is still a long way to go (Amnesty International 2015a, 2015b; Rall et al. 2016). We hope that the platform we have developed will enhance the ability of human rights practitioners to integrate video evidence into their advocacy and accountability efforts.

We also recognize that the human rights community will have to pay particular attention to the risks that come with the diffusion of technology into human rights practice. There is a possibility that the new tools and methods we describe here will shift so rapidly from being at the cutting edge to being necessary to secure prosecutions or advocate for change that only the biggest and most well financed groups will be able to keep up. In order to address this concern, we plan to make all of the components available for public use and will open source as much of the code as possible. This is, indeed, an expectation of the foundations that have supported our work. Some of the architectural rendering software we rely on is commercially sold so we cannot make that open to the public, but we are currently working on new methods of developing three-dimensional models of space directly from videos to bypass the need for proprietary products. The reality, though, is that for the forseeable future most human rights groups will not be able to use the tools we create out of the box, so we plan to continue to provide assistance to interested parties for as long as possible. Because the tools and methods described here require large volumes of multi-perspectival video in order to be effective, the number of human rights organizations and practitioners who need help will likely be manageable to us, for the next few years at least.

Aside from resource issues, there is also a risk that human rights violations will only be taken seriously by courts and public opinion in the future if victim or eyewitness testimony can be corroborated with video evidence. Finally, and most concerning, the same technology and evidence used to secure a conviction or support an advocacy campaign can be used by perpetrators and their allies to seek retaliation against individuals who film human rights violations or the individuals depicted in such videos. There are already tools available to obscure faces and make individual identification from a video difficult, but our experience suggests that nefarious actors will eventually be able to defeat these safeguards.

We would like to conclude by noting that we have endeavoured to be as transparent as possible in the methodological choices we have made (not all of which have been detailed in this paper because of the need to be accessible to a wide audience) and recognize that visual evidence must be met with appropriate scrutiny in legal and advocacy contexts. We plan to evaluate the impacts of our work moving forward, and also seek to develop the platform for broader use in other contexts.

Funding

This work was supported by grants from MacArthur Foundation [#15-107795-000-INP to Carnegie Mellon Center for Human Rights Science and 15-106572-000-INP to SITU Research], Oak Foundation [OUSA-15-025 to Carnegie Mellon Center for Human Rights Science and OCAY-15-150 to SITU Research], and Humanity United [BP10696/14PROJ to Carnegie Mellon Center for Human Rights Science].

Acknowledgements

The authors would like to thank: Olena Bondarenko, Katherin Machalek, Evelyna Nefertari, and Nick Turchak for their assistance at various points throughout this project; Griffin Frazen, Lindsey Krug, and Gideon Finck for help with visualization and animation; and Desai Fan, Po-Yao Huang, and Junwei Liang for developing the synchronization and geolocation systems described in this paper and providing technical support throughout the project. Thanks especially to Pavel Dykan and Alexandra Yatsenko for presenting us with the challenge of integrating video evidence into human rights practice and for their feedback on this paper.

References

Accatino, D., and C. Collins. 2016. Truth, Evidence, Truth: The Deployment of Testimony, Archives and Technical Data in Domestic Human Rights Trials. *Journal of Human Rights Practice* 8(1): 81–100.

Alston, P., and S. Knuckey. 2016. The Transformation of Human Rights Fact-Finding: Challenges and Opportunities. In *The Transformation of Human Rights Fact-Finding*, pp. 3–24. Oxford University Press.

Amnesty International. 2015a. 'Our Job is to Shoot, Slaughter and Kill': Boko Haram's Reign of Terror in North-East Nigeria.

———. 2015b. Stars on their Shoulders, Blood on their Hands: War Crimes Committed by the Nigerian Military.

Aronson, J. D. 2018. The Utility of User-Generated Content in Human Rights Investigations. In M. K. Land and J. D. Aronson, *New Technologies for Human Rights Law and Practice*, pp. 129–48. New York: Cambridge University Press.

BBC News. 2014. Ukraine's Berkut Police: What Makes them Special? 26 February. http://www.bbc.com/news/world-europe-25895716 (referenced 22 February 2018).

Cave, D., and R. Oliver. 2016. The Raw Videos that have Sparked Outrage over Police Treatment of Blacks. *New York Times* http://www.nytimes.com/interactive/2015/07/30/us/police-videos-race.html (referenced 25 October 2016).

Dubberley, S., E. Griffin, and H. M. Bal. 2015. Making Secondary Trauma a Primary Issue: A Study of Eyewitness Media and Vicarious Trauma on the Digital Frontline. http://eyewitnessmediahub.com/research/vicarious-trauma (referenced 25 October 2016).

Feigenson, N., and C. Speisel. 2009. *Law on Display: The Digital Transformation of Legal Persuasion and Judgment*. New York University Press.

Forensic Architecture and SITU Research. 2010. Report: Summary of Findings on the April 17, 2009 Death of Bassem Ibrahim Abu Rahma, Bil'in. http://www.situresearch.com/works/bilin-report (referenced 5 October 2016).

Goncharenko, R. 2014. Titushki—The Ukrainian President's Hired Strongmen. Deutsche Welle. http://dw.com/p/1BBjy (accessed 5 October 2016).

King, G., J. Pan, and M. E. Roberts. 2015. How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, not Engaged Argument. Working Paper. http://j.mp/1Txxiz1 (referenced 1 June 2016).

Koettl, C. 2016. Citizen Media Research and Verification: An Analytical Framework for Human Rights Practitioners. Centre of Governance and Human Rights, University of Cambridge. https://www.repository.cam.ac.uk/bitstream/handle/1810/253508/Koettl_Citizen%20Media%20Research%20and%20Verifcation_FINAL%20(1).pdf?sequence=1 (referenced 5 October 2016).

Landman, T., and E. Carvalho. 2009. *Measuring Human Rights*. Routledge.

Liang, J., S. Burger, A. Hauptmann, and J. D. Aronson. 2016. Video Synchronization and Sound Search for Human Rights Documentation and Conflict Monitoring. http://www.cmu.edu/chrs/documents/Video-Synchronization-Technical-Report.pdf (referenced 26 May 2017).

Matviychuk, O., and O. Pavlichenko (eds). 2015. The Price of Freedom: Public Report by Human Rights Organizations on Crimes against Humanity Committed during Euromaidan (November 2013–February 2014). https://issuu.com/irf_ua/docs/hr-2015-2engl (referenced 5 October 2016).

Menon, R., and E. B. Rumer. 2015. *Conflict in Ukraine: The Unwinding of the Post-Cold War Order*. Cambridge, MA: MIT Press.

New Tactics in Human Rights. 2014. Video as Evidence: To be Evidence, What Does Video Need? https://www.newtactics.org/comment/7427#comment-7427 (referenced 1 June 2016).

Nichols, B. 1994. *Blurred Boundaries: Questions of Meaning in Contemporary Culture*. University of Indiana Press.

Open Dialog Foundation. 2014. Ukraine: More than 100 People have been Killed since the Beginning of the Conflict; Approximately 300 are Considered Missing. http://en.odfoundation.eu/d/3120,ukraine-more-than-100-people-have-been-killed-since-the-beginning-of-the-conflict-approximately-300-are-considered-missing (referenced 25 October 2016).

Padania, S., S. Gregory, Y. Alberdingk-Thijm, and B. Nunez. 2011. Cameras Everywhere: Current Challenges and Opportunities at the Intersection of Human Rights, Video, and Technology. New York: Witness.

Polityuk, P. 2014. Ukraine Passes Anti-Protest Law. *The Independent* (London). http://www.independent.co.uk/news/world/europe/ukraine-passes-anti-protest-law-9066054.html (referenced 5 October 2016).

Rall, K., M. L. Satterthwaite, A. V. Pandey. et al. 2016. Data Visualization for Human Rights Advocacy. *Journal of Human Rights Practice* 8(2): 171–97.

Shapin, S. 1984. Pump and Circumstance: Robert Boyle's Literary Technology. *Social Studies of Science* 14(4): 481–520.

Silverman, C. 2015a. Lies, Damn Lies, and Viral Content: How News Websites Spread (and Debunk) Online Rumors, Unverified Claims, And Misinformation. New York: Tow Center for Digital Journalism/Columbia Journalism School.

———. (ed.). 2015b. Verification Handbook for Investigative Reporting. Maastricht: European Journalism Centre.

Uberti, D. 2015. How Smartphone Video Changes Coverage of Police Abuse. *Columbia Journalism Review* (online). http://www.cjr.org/analysis/smartphone_video_changes_coverage.php (referenced 19 October 2016).

US Department of State. 2015. 2014 Human Rights Reports: Ukraine. http://www.state.gov/j/drl/rls/hrrpt/2014/eur/236588.htm (referenced 5 October 2016).

Wardle, C., S. Dubberley, and P. Brown. 2014. Amateur Footage: A Global Study of User-Generated Content in TV and Online News Output. New York: Tow Center of the Columbia Journalism School.

Weizman, E. 2017. *Forensic Architecture: Violence at the Threshold of Detectability*. New York: Zone Books.

Whiting, A. 2015. The ICC Prosecutor's New Draft Strategic Plan. https://www.justsecurity.org/24808/icc-prosecutors-draft-strategic-plan (referenced 1 June 2016).