All the protestors fit to count: using geospatial affordances to estimate protest event size

Austin Choi-Fitzpatrick; Tautvydas Juskauskas; Md. Boby Sabur

Abstract

Protest events are a hallmark of social movement tactics. Large crowds in public spaces send a clear message to those in authority. Consequently, estimating crowd size is important for clarifying how much support a particular movement has been able to garner. This is significant for policymakers and constructing public opinion alike. Efforts to accurately estimate crowd size are plagued with issues: the cost of renting aircraft (if done by air), the challenge of visibility and securing building access (if done by rooftops), and issues related to perspective and scale (if done on the ground). Airborne camera platforms like drones, balloons, and kites are geospatial affordances that open new opportunities to better estimate crowd size. In this article we adapt traditional aerial imaging techniques for deployment on an “unmanned aerial vehicle” (UAV, popularly drone) and apply the method to small (1,000) and large (30,000+) events. Ethical guidelines related to drone safety are advanced, questions related to privacy are raised, and we conclude with a discussion of what standards should guide new technologies if they are to be used for the public good.

Keywords: Protest, methods, crowd estimation, privacy, surveillance, drones, unmanned aircraft systems

Protest size matters

Size matters for social movements (DeNardo 1985; Lohmann 1994; Oberschall 1994; McCarthy, McPhail, and Smith 1996; Chenoweth and Stephan 2011; Popovic and Miller 2015; Biggs 2016; Wouters and Van Camp 2017). Whether it be the number of names on abolitionist-era petitions or the number of people present at a “million-man” march, the ability to mobilize people (especially as citizens and consumers) and engage in coherent claims-making is a hallmark of collective action. Visible and sizable mobilization matters for both the movement’s target as well as the general public that so often mediates a movement’s effects (Agnone 2007; Burstein 2003).

Visibility matters because the ability to clog a major thoroughfare or fill a notable landmark demonstrates strength in numbers. This observation, like so many others, is strikingly similar to something Charles Tilly (1999) has already said: public collective action efforts demonstrate WUNC—worthiness, unity, numbers and commitment. This is not to say that the only path to movement
success is mass mobilization—legislative strategies, violent struggle, elite brokerage, court decisions, and opt-out tactics like boycotts—have all proven their value in helping challengers secure gains from incumbents. While mass mobilization is not the only path to success, it is an enduringly important part of the movement repertoire for the past two centuries (Klandermans 2008; Caren, Ghoshal, and Ribas 2011).

Not all success requires mass mobilization, and not all large-scale protests are successful. Heaney and Rojas (2015), for example, demonstrate the extent to which broader changes in the political landscape—e.g., the election of a Democrat to the U.S. presidency—eliminated the impetus to protest American wars in Iraq and Afghanistan. The disconnect between the turnout in climate change gatherings and American environmental policy is particularly striking. Erica Chenoweth and Jeremy Pressman have initiated a Crowd Counting Consortium to open-source the estimation of turnout to public political events.¹ By their estimation, between 5.9 and 9 million people protested in 2017, with the vast majority (89%) estimated to be protesting Donald Trump. These protests have not dampened Republican support for the regime and its policies. It is quite likely that they had the opposite effect, pushing moderate Republicans to demonstrate a sort of counter-protest support for the President. Not only are some large-scale social protests unsuccessful, future empirical analysis may prove them to be counterproductive, as suggested anecdotally in the case of Trump.

If large-scale protests have bifurcated outcomes (leading to success or counter-mobilization and even repression), there is no disputing their symbolic impact. The notion that 2-3% of Americans are protesting a sitting President over a year’s time is an important barometer for public attitudes. The cumulative, crowd sourced approach adopted by the Crowd Counting Consortium is premised on the idea that turnout matters, whether it is large or small.

This is different than historic approaches to turnout. To date, much of the conversation about protest size has focused on newspaper data. A number of problems have dogged this usage, however. It turns out that the New York Times and Washington Post covered fewer than half of all disorders that occurred between 1968 and 1969, for example (Myers and Caniglia 2004). In that period, coverage was determined by event intensity, distance from the paper, event density, the city’s population size, the type of actors involved and the day of the week. Newspaper coverage matters (or mattered in the 20th century) for media cycles, public opinion, and the concomitant sense of urgency policymakers feel regarding the issues that have brought people onto the streets. All news is not created equal. The punch line here is that violent riots in big cities got covered but the kind of events that comprise much of the Consortium’s data were ignored.

¹Crowd Counting Consortium: https://sites.google.com/view/crowdcountingconsortium/home
See also the work of Count Love: https://countlove.org/
Recent work by Michael Biggs (2016) suggests that it is the size of an event that matters more than the total number of events. His analysis undermines an entire vein of movement scholarship that has drawn on event-count data on protests. Biggs argues that it is protest size that explains the newspaper coverage that gets indexed in the first place. Relevant here is Biggs’ observation that protestors do their best to maximize their size at single events, not to spread themselves over many smaller events. Why else do they gather in capital cities and in front of Parliaments? His observation reinforces the findings of Myers and colleagues (Myers and Caniglia 2004; Ortiz, Myers, Walls and Diaz 2005).

However unintentionally, this critique provides a backhanded compliment to newspaper data: journalists and editors do a remarkable job of noting large and significant events. Large movements also create opportunities for attracting new supporters, whether on the street or as conscience constituents who support from home. They also have the effect of creating hospitable environments for counter mobilization by other civil society actors (Meyers and Staggenborg 1996). Large numbers of people on the street also represent symbolic challenges to authorities and practical challenges for administrators and bureaucrats. The temptation, then, may be to engage in repressive or co-opting responses in the event movements’ target entrenched interests. This is true whether the target is a university, hospital, church or government (Walker, Martin, McCarthy 2009). Size matters for targets, for the general public, for newspaper editors, and for social movements themselves.

Two arguments can be identified thus far: the Crowd Counting Consortium’s implicitly cumulative argument and Bigg’s explicit emphasis on large-scale protests. This essay does not set out to resolve this tension, but to provide a method for obtaining better data on the turnout for all protests, whether large or small. The reason for this is that both approaches rely on accurate data on event size. For social movements, perceptions matter. Political opportunities, it is widely noted, are only as real as they are perceived (Goodwin and Jasper 1999). If a movement perceives an opportunity where there is none, it is possible they may respond with enthusiasm and a redoubling of their efforts (Rasler 1996). In this way a closed opportunity opens. Perception might not be everything (after all, if a movement lacks the resources necessary to complement their enthusiasm, all may come to naught), but it cannot be ignored altogether. The same can be said of the threat experienced by institutional targets facing a challenge from a newly formed bloc of voters in a Parliamentary plaza or group of students mobilized on the campus quad. The salient point here is that perceived protest size matters. This is why so much effort has gone into contesting exactly how large an event is—demonstration turnout is a crucial political resource for social movements (Wouters and Van Camp 2017: 450).

In sum, a “Million Man” March has nothing but alliteration going for it if it turns out the number is inflated by one million. Protests are inherently political and politicized events. Thus, the actual number of protestors matters to at least
one of these four parties (i.e., movement, target, media, general public). The Million Man March itself is often cited as a prime example of the inadequacies of crowd size reporting (McPhail, McCarthy 1996; Watson, Yip 2011). Organizers of the event placed attendance numbers between 1.5 and 2 million. The United States Park Service estimated the crowd to be around 400,000 people. The discrepancies between the two numbers resulted in the legal action taken against the National Park Service by March organizers.

Nobody doubted that a tremendous number of people took a stand with Louis Farrakhan against the economic and social conditions of African Americans. But once again, it is not just the number itself that matters, but its relationship to perception. Was the march a success or a failure? Whose interests were served by the varying answers to that question? In some ways the answer is mediated by the gap between perception and reality of the event’s size, factors themselves directly connected to the movement’s perceived worthiness. Of the many important factors at play in studies of protest turnout, this study focuses in on a key methodological puzzle: how best might the size of an event be estimated? What is important overall, and the subject of this article, is the process involved in getting the numbers right.

**Estimating protest size methods**

A broad survey of crowd estimation techniques suggests there is significant methodological fragmentation across media, authorities, academics and social movement actors. Lay approaches range from naïve guestimates to politicized declarations of “actual size.” Official approaches are often plagued by political factors (Kielbowicz and Scherrer 1986). Gitlin (1980) cites instances in which the *New York Times* simply passed along police estimates of Vietnam War protest sizes. Mann (1974) found newspaper estimates of crowd size often matched the publisher’s political leanings (as measured by their editorial board). Edelman (1986) found higher police estimates for established political candidates and lower for more radical groups from the left and the right, when compared to his use of the industry-standard Jacobs Crowd Formula (JCF) (which we used in this study, as discussed below.). Several of these examples are emphasized by Michael Biggs (2016), who explicates these complications in great detail.

In what follows we will leave aside these politicized and haphazard approaches and focus our attention instead on the development of estimation methods within the scholarly literature on protests. Here it seems there is little debate, since the crowd size estimation method is fairly well established, despite a relative lack of attention to the issue. Estimation techniques among movement scholars appear to have remained virtually unchanged since the 1960s. Those readers eager for a significant reimagining of the status quo will be disappointed. What we propose here is rather a transposition of the existing methodological approach to a new platform. We suggest an extension and improvements rather than a radical revision.
The industry standard method of estimating the size of static crowds has been relatively stable for the past five decades (in this study we leave to the side moving crowds, a matter for another day). Herbert Jacobs, a journalism professor at UC-Berkeley, pioneered the approach from an elevated angle as he observed the Free Speech Movement’s birth outside his office window.

He noticed the concrete pattern in Sproul Plaza provided the perfect grid format for consistent estimation size. The refined version of this approach appeared in the *Columbia Journalism Review* in 1967. The central assumption is that loose crowds were comprised of one person per 10 square feet (0.93 square meter) of space, while the same person occupies only 4.5 (0.42 square meter) square feet in a dense crowd and a mere 2.5 square feet (0.23 square meter) in the front of an event, assuming of course that there is a “front of the event.”

The task, then, was to accurately estimate the (1) square footage of the site, (2) the percentage of the site occupied by participants, and (3) the density of the crowd. Considered together, these factors underline the principal of the Jacobs’ Crowd Formula (JCF) and would allow any individual an accurate estimation to any crowd size. In table 1 we apply general assumptions to several recent sites of protest.

**Table 1: **Public gathering places and carrying capacities at different density levels

<table>
<thead>
<tr>
<th>Area in square meters (in sq. feet)</th>
<th>0.23 m² (2.5 ft²)</th>
<th>0.42 m² (4.5 ft²)</th>
<th>0.93 m² (10 ft²)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Int'l football field</strong></td>
<td>10,800 (116,250)</td>
<td>46,956</td>
<td>25,714</td>
</tr>
<tr>
<td><strong>US football field</strong></td>
<td>5,364 (57,733)</td>
<td>23,321</td>
<td>12,771</td>
</tr>
<tr>
<td><strong>National Mall (US) (total area between the Ulysses S. Grant Memorial and the Lincoln Memorial)</strong></td>
<td>1,200,000 (12,916,692)</td>
<td>5,217,391</td>
<td>2,857,142</td>
</tr>
<tr>
<td><strong>Trafalgar Square (UK)</strong></td>
<td>21,000 (226,042)</td>
<td>91,304</td>
<td>50,000</td>
</tr>
<tr>
<td><strong>Tiananmen Square (China)</strong></td>
<td>380,000 (4,090,286)</td>
<td>1,652,173</td>
<td>904,761</td>
</tr>
</tbody>
</table>
NOTE: Area calculations were done on Google Earth Pro (Trial version), though similar results can be obtained using ArcGIS, GIS Atterbury, Daftlogic, etc. While calculating we also included surrounding areas that also have crowd carrying potential. Those surrounding areas might include green areas, parks, wide streets, crossroads, etc. Our estimates occasionally differ from those found elsewhere (e.g., http://en.wikipedia.org/wiki/List_of_city_squares_by_size).

Jacob’s principal has been redefined and adapted a number of times (Seidler, Meyer, Gillivray 1976; Swank 1999; McPhail, Clark and McCarthy 2004). In the 1970s, the United States Park Police developed a formula of their own (McPhail and McCarthy 2004). Others incorporated aerial photography from helicopters and official site measurements from city square footage plans. Taken together these factors allow for a more accurate assessment than what Jacob’s formula in general would account for. These improvements to accuracy were made at the margins however, and the importance of the original three factors—site dimension, percentage occupancy, density—remained intact.

The JCF reached its current industry standard formulation through the work of Clark McPhail, who has consulted extensively on the issue. McPhail and McCarthy (2004) add one component (comparative data) to suggest four rules for the most credible estimation of crowd size:

1. Carrying capacity of site;
2. Density of the crowd;
3. Observations from multiple vantage points, some of which must be elevated;
This approach is notable for its integration of both the direct estimation recommended by Jacobs as well as complementing that data with assessments of other measurements, such as the number of busses used to bring people into an event from far away (a practice as common in New Delhi as in Washington D.C.).

We have established that real and perceived crowd size is an important signaling mechanism (whether it is followed with political action is another matter altogether—see Heaney and Rojas 2015), that accurate assessments of crowd size are important, and that there is in fact a relatively stable approach for measuring crowd size. The shortcoming in this method, we argue, is that it is difficult to secure multiple vantage points from which to watch or photograph a crowd. Movement actors do not usually have access to the roofs of the buildings surrounding the protest space. Significant crowds may form in places other than those anticipated by authorities, journalists, or even the movement itself. Multiple crowds may converge in different locations simultaneously.

In these, and countless other conditions, observation from multiple elevated vantage points is simply impossible. Of course these obstacles can be overcome by having an airplane or fixed-wing aircraft secured for the day of the event and deployable to consecutive locations on a moment’s notice. This solution, however, has two significant weaknesses: (1) it is expensive, usually well beyond what any movement actor is able to afford; and (2) it assumes open airspace, something that cannot be counted on in many of the political contexts where authorities feel threatened (e.g. the US Federal Aviation Administration closed the airspace over Ferguson, Missouri at the height of the 2014 protests over state repression there, perhaps in response to the deployment of drones by journalists).

In what follows we argue that geospatial affordances—new ways of doing things from the air, here including drones, balloons and kites—provide the benefits of a helicopter or fixed wing aircraft (multiple vantage points at altitude) without the associated challenges (cost and airspace access). In providing an extension of the JCF to a new geospatial affordance (the drone) we provide civil society actors with a means for securing affordable, easily deployable, high quality, aerial footage of protest events and a method for easily analyzing this visual data.

An aerial-based crowd estimation method

We use a consumer-grade unmanned aerial vehicle (UAV, or drone) to implement the Jacobs Crowd Formula (JCF, hereafter). It is important to note that the same technique works regardless of how the image was made, so long as (1) the camera is at sufficient altitude, and (2) the imaging sensor is

---

2We like the term “remotely piloted aircraft system,” as it reflects the wide range of payloads and the reality of a pilot (of any gender), but fear it is not long for this world as algorithms make more in-flight decisions, rendering useless the phrase “remotely piloted”.

303
perpendicular to the ground (i.e., the camera is pointing straight down). While we suggest several modifications (listed below) they are simple extensions of the JCF. Thus, the main advantage of the proposed method is its ease of use. While the technical details of the method are spelled out in greater detail elsewhere (Choi-Fitzpatrick and Juskauskas 2015), a brief overview of the approach bears mentioning.

**Step 1: Drone platform** – All tests in this study were conducted with a commercially available DJI Phantom 2 Vision+. We chose this device for five reasons: it was the industry standard at the time of testing; no additional equipment is required for flight; its GPS capabilities allow it to be flown quickly and safely by pilots with a range of experience; it has a “return home” function that ensures a safe landing if the operator is detained or the link is broken; and it is a “prosumer” product, meaning it combines some professional features with a consumer price point.3

**Step 2: Digital image** – We made one important change to our device: We modified the UAV to ensure the camera was angled perpendicular to the ground, effectively eliminating issues related to estimating at an angle—an issue that plagues Jacobs estimates from rooftops. We used commercially available software to eliminate the round lens flare known as the “fish-eye effect”.

**Step 3: Area measurements** – The process for securing an area measurement are described in greater detail in Choi-Fitzpatrick and Juskauskas (2015). We began by laying a 10-meter marker onto the ground and used that as our reference point. Once the exact length of the reference point or line had been determined, we used publically available software (GIMP) to translate it into pixels as this is the unit of analysis for digital imagery. Table 2 shows a few dimension-sizes at three standard altitudes.

---

3When purchased, small consumer drones ranged in price from approximately US $300 to around US$3000. This device was purchased for US $1000.
**Table 2. Area Measurements and Crowd Estimation**

<table>
<thead>
<tr>
<th>Altitude in meters (feet)</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Photo dimensions in pixels after fish-eye correction</td>
<td>Fish-eye correction in GIMP software (main, edge)</td>
<td>Reference on ground in M (ft)</td>
</tr>
<tr>
<td>50 (164)</td>
<td>4384x2466</td>
<td>-20, -20</td>
<td>10 (32)</td>
</tr>
<tr>
<td>100 (328)</td>
<td>4384x2466</td>
<td>-20, -20</td>
<td>10</td>
</tr>
<tr>
<td>150 (492)</td>
<td>4384x2466</td>
<td>-20, -20</td>
<td>10</td>
</tr>
</tbody>
</table>

Source: Choi-Fitzpatrick and Juskauskas (2015)

**Step 4: Grid digitally applied to image** – Placing a digital grid over the digital image allows for the rapid estimation of individual unit density and counting of total units. After determining the number of pixels that correspond to the 10 m. reference line, a simple grid can be applied to the picture. A grid application is accomplished in two basic steps using open sourced software and described in Choi-Fitzpatrick and Juskauskas (2015).

**Step 5: Estimating the density levels of each grid** – With the grid then applied, and with each grid measuring 10 meters between each gridlines, it is now possible to estimate the number of individuals within each grid. Using (Western) density levels established in the literature, we are able to base estimates on five density levels, effectively, where there are no people, where the crowd is very loose, relatively loose, relatively dense, and very dense. Specifically, the five possible density levels are as follows:

- **Empty** (Density Level 0) – A rooftop, or any other empty space, counted at zero.
- **Very loose** (Density Level 1) – A very loose crowd with a very low density level. You could ride your bike through this crowd easily. It is counted manually.
- **Loose** (Density Level 2) – A somewhat loose crowd with a pretty low density level. This is a crowd you could walk through easily without bumping into too many people (imagine about 1 person per square meter).

---

Recent work by Sorokowska et al (2017) suggest that personal space varies significantly by culture, meaning that a loose crowd would be looser in Romania (where people prefer to stand about 120cm from one another) than in Bulgaria (preferring only 90cm apart).
On average, at this density level there are usually about 109 people in the grid. [one person in 10 ft² or 0.93 m²]

Dense (Density Level 3) – This is a dense crowd. You would have a hard time moving through this crowd, but it would be possible (imagine more than 2 people per square meter). On average, at this density level there are usually about 238 people in the grid. [one person in 4.5 ft² or 0.41 m²]

Very dense (Density Level 3) – This is an extremely dense crowd. It would be nearly impossible to move your arms in this crowd (imagine more than 4 people per square meter). This is the same as the very front of a concert, just in front of the stage. On average, at this density level there are about 435 people in the grid [one person in 2.5 ft² or 0.23 m²]. NOTE: this density level rarely occurs.

Step 6: Compile estimate of crowd size – The sixth step is counting how many squares of different density levels the grid has. The actual number of the crowd is summed up.

Step 7: Determine intercoder reliability – Some users may choose to incorporate Cohen’s Kappa—a statistic that measures agreement between different estimators—as an optional seventh step in this estimation methodology.

Implementing the drone-based crowd estimation model

We applied this method in two public gatherings in Budapest, Hungary. The first was a concert and the other was a protest event. General detail about each event (date, time, weather and GPS coordinates) and specific information regarding estimation parameters (i.e., inter-coder reliability) can be found in Choi-Fitzpatrick and Juskauskas (2015) and briefly in Appendix I.

First field test: concert

The image was made at 160 meters in altitude. Clearly larger crowds will require “zooming out,” an action accomplished by increasing the UAV’s altitude so that a greater surface area is covered by the image. Prior to photographing the crowd we made the estimation necessary to insert the grid in GIMP.
To do this we identified a line that was clearly visible from this altitude. With knowledge of the line’s actual length on the ground (15.6m), we used GIMP to measure the pixel length of this referent. The 15.6 meter line on the ground is equal to 237 pixels within the digital photo file. As we need a grid of regular 100 m² squares, we need to convert 10 meters into pixels. The formula for determining this ratio is described in Section “Step 3: Area measurements”. In this case, 10 meter reference equals 152 pixels in the picture. A 10m x 10m square on the ground is therefore a 152px x 152px grid in the image (Image 1).

We recruited research assistants from a cohort of graduate students. Volunteers received a modest voucher and brief verbal introduction to the process and were given approximately 90 minutes to accomplish this task. We found that 80 minutes was the average amount of time required to accomplish this task, and that the instructions led to very few misunderstandings about the task, or any particular step in the task. As detailed in Choi-Fitzpatrick and Juskauskas (2015) coders were instructed to determine the density level within each grid (X, O, 1, 2, 3), to manually count any persons within density level 0, and to then determine what percentage of each grid was filled at the indicated density level (25%, 50%, 75%, 100%). These tasks were accomplished with an 8x10-sized printout of the photograph and a white marker. Coding decisions were made directly onto the image itself.

This data was then entered into a spreadsheet by the article’s second authors and a Cohen’s Kappa, an industry standard inter-coder reliability estimate, was applied to the data. Our final iteration of the test resulted in an inter-coder reliability estimate of .7 and a crowd estimate of between 37,112 and 37,695.

While we are pleased with this level of agreement, we would have preferred to offer a benchmark for comparison. Three are desirable but in this case were not possible. First, ticket sales or turnstile counts; unfortunately for our purposes
(but fortunately for concertgoers) this was a free event and neither data point existed. Second, other media sources: several bloggers after the event claimed the event was attended by several tens thousands of people. Third, “indirect passenger volume estimation,” such as busses: this event took place close to a public transportation hub, making comparative data hard to obtain.

**Second field test: protest**

The second field-test of the method was implemented at a demonstration held by a civil society organization. The event was held during a national holiday and targeted social injustices and lack of democracy in Hungary. In our coding of this data we determined there were 2,609 people present at the event. Using the process described above, external coders (who were unaware of our own estimate) determined that between 2,589 and 3,750 individuals were present, with a Cohen’s Kappa of .85.

While we could have cropped the image to make counting easier, we have left it untouched in order to emphasize one additional question unaddressed by this method: Who is part of the event? Who counts? Are the people in squares E3, E4 and E5 part of the event? We can ask the same question of almost everyone in columns 1, 2, 3, 5, 9, and 10. Presumably the answer to this question varies based on the size of the event—for a large event the cells bounding the central mass of individuals may be counter-protestors, police, reporters, or bystanders. They may also be comprised of individuals debating whether to join the event. Clearly birds-eye data must be complimented by on-the-ground data.
UAV data are agnostic about turnout — e.g., we would never know if lots of counter-protestors infiltrated an event in order to disrupt it. It is important to augment the single method of measurement introduced here with observations on the ground, and with comparative benchmarks, where possible. In this regard both quantitative and qualitative data matters. At this event, media estimates of the turnout ranged from 700 to 3,000. The most frequent estimate was in the 1,500-2,000 range. Here we face a question deserving additional exploration: should the significance of mass mobilization events be measured by supporter turnout or total turnout?

If an event attracts 10 protestors but 100 counter-protestors, this ratio is salient. But if an event attracts 100,000 protestors and 5,000 onlookers, should the onlookers be included in the estimation of the event size? Presumably a large number of onlookers indicates that the event is important not only to the protestors, but to other publics as well. We leave this puzzle to others.5

5One concerned reviewer suggested that our approach shifts debates from the public domain to that of specialists, as it sets up technical experts to debate particular definitions, parameters, personal distance rates, crowd composition, boundary puzzles, and so forth. We would be quite disappointed if this turns out to be the case. Our objective is to make the estimation of crowds easy to perform and audit. Our goal is explicitly anti-specialist, as it were. Subsequent adoption and debate will suggest whether our optimism is warranted.
Discussion

Unmanned aerial vehicles are the subject of increasing attention in public, policy and commercial arenas (Choi-Fitzpatrick, et al 2016). Yet the bulk of this attention has remained focused on two debates: the first on how the state should regulate UAVs used for commercial purposes, and the second on what should be done about the use of UAVs for military purposes.

This essay is meant to provoke discussion in a third area of inquiry related to the use of drones by a broader array of actors. This contribution is timely, as protestors flew drones over Maidan in Kiev during the upheaval that led to the ousting of then-President Viktor Yanukovych and used them to document police abuse of water protectors at Standing Rock. Russia Today documented the protests that followed a police shooting in Ferguson, Missouri, researchers documented anti-regime protests in Budapest, Hungary, and a South African arms manufacturer has begun shipping “anti-riot” drones equipped with non-lethal armaments, including rubber bullets and tear gas.

These developments, and our intervention, raise three critical questions regarding the relationship between technology and surveillance. The first question is whether movement communities should broaden their focus from the state’s use of drones for surveillance and targeted killing to the use of drones by the state (for other purposes), corporations, and civil society actors. Clearly we believe the answer is unequivocally in the affirmative – if new technology is encroaching on (and perhaps expanding) the public sphere, then this matters for both scholars and practitioners of protest. While we have answered in the affirmative, it appears activists have not had a broad and vibrant debate over drone use by civil society actors.

The second question is whether drones should be used by state, corporate, and civil society actors. If the answer is a simple no, then a significant amount of hard-nosed pragmatic work must be done to undo a decade’s worth of technological innovation in terms of robotics and artificial intelligence. It is more likely that the answer is more complicated, and will require some sort of disaggregation of actors, intent, space, etc. At present it appears that corporations have taken the lead in developing this technology, states have taken the lead in weaponizing and deploying this technology, and that change-oriented actors within civil society have been regulated to a reactive stance.

The third question is what sort of general principles should guide the use of this new category of digital devices. Legislative frameworks are being hastily constructed at the international, national, and sub-state level, but these frameworks are technical prescriptions, and elide broader ethical puzzles. To this end we follow earlier work in advancing a six-fold set of guiding principles and puzzles for the use of UAVs by civil society actors (Choi-Fitzpatrick 2014).

1. Subsidiary – Should drones only be used in those situations where other actions or technology already yield the desired result? Can new technology be original without being useful? If so, how might we know the difference?
2. Physical and Material Security – Appropriate measures (training, flight-planning, etc) must be taken to ensure the security of people and things in the area where a UAV is used. As drone use increases, who will coordinate these efforts? How will anti-establishment actors (e.g., protestors) fit into this space?

3. Do No Harm – This concept, pioneered by Patrick Meier and colleagues, emphasizes the importance of the public good: benefits must outweigh costs and risks. Yet the nature of the public good is a matter of great debate; is documenting an embarrassingly small turnout on a key social issue harming the movement’s cause (assuming the key issue is for the public good)?

4. Newsworthiness – This concept is borrowed from journalism’s focus on the greater good and emphasizes the importance of a free press (in both corporate media and citizen journal models) in holding the powerful to account. Must pro-social and advocacy footage only be made “for the greater good” or is aerial data collection important in its own right? Is ubiquitous drone surveillance a simple step up from Google Earth in terms of frequency of coverage or is it a scale shift that represents a fundamental threat to privacy?

5. Privacy – While debates about privacy and technology are ongoing, and users of digital media appear less worried about the issue than advocates, what is the proper balance between the privacy of private citizens and newsworthiness and the public good? Privacy is treated differently across national contexts, and no blanket legislation is possible, meaning the increased use of drones is likely to lead to very different policy approaches.

6. Data Protection – Data protection is critical. Social movements who use camera equipped drones to monitor police action at a political protest, for example, must take great care to ensure that the privacy of protestors is protected and that the digital data is kept secure.

Subsidiarity – Is it possible to estimate the size of medium to large crowds using existing approaches. At present there is no auditable method for estimating the size of a crowd in an unbounded space. By sealing off a space and adding a turnstile, one can easily measure ingress and egress—but this violates the unbounded space condition that applies in most public events. Use estimators to count off through the crowd in a rigorous way and you can generate an estimate—but this violates the auditable and affordable conditions that makes this approach apolitical. We determine that we have met the subsidiarity threshold.
**Physical and Material Security** – We did our best to launch, fly, and land our UAV beyond the edge of the crowds depicted in this study. New parachute technologies have emerged in the time that has transpired between our data gathering and this publication. These would add a further level of safety to our flight. Furthermore, the first two authors have tested camera-equipped balloons which remain tethered to the ground, thereby eliminating a host of safety concerns. We determine that we have nearly met the security threshold in our past efforts, but future efforts will certainly meet them fully.

**Do No Harm** – The gathering and publicizing of data about public events is inherently in the public interest and the provision of this data is for the public good. Our activities could have caused harm had our camera captured individually identifiable faces. Critics of our work have pointed out that the current analysis overlooks another significant area of potential harm: our approach could highlight the extent to which important events suffer from low turnout rates, thereby amplifying criticisms from opponents. We thus leave open the question of whether we succeeded in doing no harm.

**Newsworthiness** – Gathering and publicizing data about events that social actors desire to make public are inherently worthy of public attention. As a result, our documentation of a protest event (Image 2) is decidedly newsworthy. Whether our documentation of private citizens at a public concert (Image 1) is newsworthy is less clear-cut, although we feel that such events are regularly covered by newspapers in the arts and entertainment section. We leave open the question of whether we met the newsworthiness threshold in one of our two cases.

**Data Protection** – Data captured during public events should be secured. How it is secured, and at what level of protection, is a matter of ongoing debate. All of the raw footage for this project is stored in the first author’s Dropbox account, which synchs over password-protected WiFi connections to the hard-drive of his password-protected MacBook Air. Is this a secure arrangement? This approach is sufficient for apolitical data, but would be easily hacked by a sovereign, or state-sponsored agents intent on disrupting protest activity. Our data protection is sufficient at one level and insufficient at another.

**Privacy** – By engaging the camera function on our UAV only at a high altitude, we elided the complicated issue of privacy. Activating the camera at a lower altitude, however, was technically feasible and would have certainly captured discernable faces. Here we face a puzzle: should activists document public events in such a way that capture individually-identifiable features? To date citizen journalists have argued in the affirmative, and a wave of scholarship on new digital technologies (i.e., smartphones) has suggested that these new tools level the playing field when it comes to capturing and telling stories (Milan 2013). The first author’s sense, however, is that individuals who express enthusiasm for smartphones are often more sanguine when it comes to UAVs equipped with cameras (or other sensors). Why might this be? A sustained conversation about the deployment of drones by protesters, police, and the media is long overdue, and will raise far more questions than this essay will...
answer. Returning to this study, we have respected the privacy of individual actors by capturing and presenting data that obscures individual identities. We will let the reader determine whether or not we have met these thresholds. More broadly we hope our guidelines are subject to debate, as they represent an initial effort to establish broadly applicable ethical norms. Our thinking is that these could guide individuals and institutions in establishing specific guidelines around questions like 1) who gets to fly these devices, 2) where, 3) with what training, and 4) under what conditions.

It is critical for movement communities to debate these issues. While free spaces are critical in fostering the kind of solidarity and commitment necessary to sustain radical politics (Cross and Snow 2011), these are consistently subject to encroachment by the state. This is one way the state kills movements (Davenport 2014). The advent of new digital tools means efforts to create solidarity (online for example) are subject to a host of new threats. Cress and Snow (2011: 119) argue that a “security culture” must be developed within activist circles if free spaces are to remain “free”. Movement engagement with these tools and the development of new practices that balance solidarity and security should always be kept within view.6

Returning to the methodological intervention that lies at the heart of this article, the combination of a camera-equipped UAV with a simple but accurate methodology improves on the status quo established by Jacobs and extended by others. This improvement is six-fold.

Firstly, with regard to scalability, the method can be used to estimate a crowd of 100 or 100,000. The linking of altitude to square meters of ground cover, and of ground coverage to image pixels, represents a fresh approach to crowd estimation. As a result, crowds of all sizes can be measured using this method.

Secondly, with regard to cost, the results produced in this study were performed using equipment costing one thousand US dollars at the time of purchase and half that at the time of publication (doubtless a comment on both the youth of the technology and age of this essay!). The same results could be obtained by balloon for a fraction of this amount. These expenses pale in comparison to the cost of renting a fixed-wing aircraft or helicopter to perform an estimation of similar accuracy.

Third, portability: while it may be too obvious to deserve mentioning, this solution can be deployed from a backpack or carryon-sized luggage. Even more easily deployed technology is available and new devices are quickly entering the market.

The fourth benefit, ease of use, relates to the fact that off-the-shelf units such as the one used in this test, and indeed any others utilizing GPS capabilities, can be deployed comparatively quickly.

---

6Doing so is not always easy, as creators of technology, users of technology, and critics of technology rarely come from the same milieu (c.f., Hoople and Choi-Fitzpatrick 2017).
The fifth benefit, replicability, refers to the fact that the method we introduce produces comparable data regardless of location, crowd-size, camera dimensions, UAV-type etc. The first author has captured aerial imagery using a helium balloon as well as a kite. Both platforms provide the exact same level of coverage as a UAV, but without attendant concerns about safety and novelty.

The sixth improvement we bring is in regard to the incorporation of an inter-coder reliability estimate and a relative standard error term. Together, these benefits combine to recommend this solution to anyone interested in quickly deploying inexpensive equipment to accurately estimate the number of people present in crowds of all sizes.

Listing these benefits should not obscure the complexity involved in using this technique. The entire enterprise raises a host of issues, especially related to privacy and security. As suggested earlier, it is not at all clear how to best balance privacy and transparency, especially when social movements set out to challenge those in positions of authority. This study is an example of innovative use of a new technology in the absence of a policy framework. Regulations devised for an earlier age are unwieldy and ill-matched to new technologies and uses.

Taken a step further, UAVs push a broader question regarding whether privacy is a core collective good, as some have recently suggested (Livingston and Walter-Drop 2014). Any attempt to answer this question will surface deep philosophical divisions between the United States and the United Kingdom and much of continental Europe. Recent recognition of the “right to be forgotten” in Spanish courts has hardly elicited a shrug from Americans actively uploading all manner of content to the cloud, despite the thin guarantees provided by click-through user agreements. While a majority of Americans are pessimistic about commercial and personal drone use, this discomfort may decrease with familiarity, although this depends entirely on developments in both regulatory and commercial spaces. Whether the technology is emerging or settled, the best approach is an ethical approach.

In brief, we believe we have managed to blend old methods with new technology in such a way that respects provisional guidelines for its ethical use. Of course, caveats abound. To begin with, it is important to note that while we have used a quadcopter, this approach should work with both fixed wing UAVs as well as satellites.7 Also, the method is guided by several main assumptions: the first is that the crowd is static—not going anywhere—which is mostly the case in protests and demonstrations that gather and remain at a particular public place. More sophisticated methods are required to address the flow of crowds found in marches.

---

7My colleagues at the University of Nottingham are, for example, applying machine learning to large datasets of satellite imagery in an effort to establish a baseline of brick kilns in the “kiln belt” in India, Nepal, and Pakistan, an area disproportionately plagued by bonded labor and human trafficking (Boyd et al 2018).
Secondly, our methodology assumes individuals are standing on level ground. It is not clear to what extent our calculations would have changed were the ground uneven. Shifting the drone off-center for safety purposes, for example, would increase security but make subsequent imagery harder to inspect visually (“ocular inspection” as some say). Thirdly, we made these images during the day in order to ensure we could capture imagery of discrete individuals. Modifications would be necessary to extend this method to count crowds in low light conditions.⁸

Working in the Global North there were fewer security issues related to theft of the device. Security may present an issue in more densely populated countries where there might not be as many places suitable for the safe launch and landing of the craft. It may also be that crowds are more dense or loose in other parts of the world. A final consideration when working with this method outside the Global North, but present worldwide at the moment: anonymity is hard when the novelty of UAVs attracts the attention of passersby.

Of course, nothing about the technology prohibits a drone operator from securing footage during ascent and descent, or from navigating the drone through a crowd in an effort to, for example, capture footage of police brutality. The framework introduced here only begins to address the ethical considerations related to the use of this setup for citizen journalism.

This method is platform independent, as it can be applied to images made at altitude by an airplane, helicopter, satellite, drone, balloon, or kite. In choosing to test the method with a UAV platform, however, we hope to initiate a broader conversation about the role new technologies play in the protest repertoire. At a time when artificial intelligence and machine learning are being coupled with autonomous devices (especially drones and robots) in order to gather data that is subject to pattern analysis and facial recognition, scholars and advocates have an opportunity to decide whether or not they want to experiment with these technologies, call for their abolition, or ignore them altogether.

We anticipate these preliminary tests can easily be augmented with more sophisticated methods and techniques. For example, from the very beginning the biggest puzzle for us was area measurements. If area measurements are automated or expressed in an algorithm, it would make things easier. We are confident overhead imagery can be combined with current innovation in the field of computer vision to begin automating the estimation of crowd size (Ryan 2013). Ongoing research has also produced more sophisticated methods for estimating density levels. Both issues might be addressed by the development of a mobile application or purpose-built software that could automatize the whole estimation process. Others are also working on the issue of automating the assessment of visual data (e.g., Marana et al 1999; Zhan et al 2008; Ryan et al 2009; Ryan 2013; Kong, Gray and Tao 2005 and 2006), though not from the same platform as ourselves. There is plenty of room for growth in this area.

---

⁸Presumably, future work could incorporate infrared cameras rather than traditional cameras to capture images that are amenable to the same methodological treatment.
Conclusion

But what does any of this tell us about social movements? We hope our method will prove useful to those with an interest in the actual size of protests, riots, marches and other politicized mass gatherings. In referring broadly to “those with an interest” we mean to describe police, policy-makers and protestors alike. McCarthy, McPhail, Smith (1996) have established the close link between protest size and media coverage. To date the gap between estimated and actual protest size have fluctuated based on the location (it’s easier to estimate events in popular locations where prior estimates have been established) and the media’s decision to report police or protesters’ estimates (the latter almost always being higher than the former).

More accurate estimates are not necessarily good news for social movements, who sometimes take advantage of the perception of large events to advance claims. This issue aside, our method frees movements to make their own estimates independent of the state, which is often more likely to possess the resources necessary to produce credible estimates. Additionally, thanks to social media, this information can be easily and instantly uploaded and disseminated. Social movements have the technology, capability and ethical framework to use UAVs in order to ensure accurate and verifiable crowd estimates. Whether they do so is another matter altogether.
References


APPENDIX I: General Details of Flights

<table>
<thead>
<tr>
<th></th>
<th>Test 1</th>
<th>Test 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>16th June 2014</td>
<td>23rd October 2014</td>
</tr>
<tr>
<td>Time</td>
<td>20:25 (GMT +2)</td>
<td>16:00 (GMT +2)</td>
</tr>
<tr>
<td>Weather</td>
<td>+24, clear</td>
<td>+10, rainy</td>
</tr>
<tr>
<td>Wind</td>
<td>5 km/h</td>
<td>4 km/h</td>
</tr>
<tr>
<td>GPS</td>
<td>9 satellites</td>
<td>10 satellites</td>
</tr>
<tr>
<td>Altitude</td>
<td>160 m.</td>
<td>80-90 m.</td>
</tr>
<tr>
<td>Take-off</td>
<td>Heroes square, Dozsa Gyorgy Way, south-east side</td>
<td>Blaha Lujza Square, Budapest</td>
</tr>
<tr>
<td>Reference (px)</td>
<td>10 m (152 px)</td>
<td>10 m (308 px)</td>
</tr>
<tr>
<td>Grid square (px)</td>
<td>100 m² (152x152 px)</td>
<td>100 m² (308x308 px)</td>
</tr>
<tr>
<td>Total number of people (est)</td>
<td>36,000</td>
<td>2,609</td>
</tr>
<tr>
<td>Cohen’s Alpha Intercoder reliability</td>
<td>.73</td>
<td>.85</td>
</tr>
</tbody>
</table>

About the Authors

**Austin Choi-Fitzpatrick** is assistant professor of political sociology at the University of San Diego’s Kroc School of Peace Studies (US) as well as associate professor of social movements and human rights at the University of Nottingham’s School of Sociology and Social Policy (UK). He is co-founder of the Good Drone Lab and author of a book tentatively titled *ProtestTech*. He can be contacted at achoifitz [at] gmail.com.

**Tautvydas Juskauskas** is European public business development manager at DJI. Tautvydas is in charge of business development and key account management in public domains, such as NGO sectors, Research and Education, as well as Public Safety. During his master studies at the Central European University, he co-founded the Good Drone Lab, an award-winning collaborative space for the development of drone uses for the greater good.

**Md Boby Sabur** is working as Monitoring, Evaluation, Learning and Analytics Specialist with United States Agency for International Development (USAID) Bangladesh Mission.