

WE SHOULD BE MAKING GREAT VIDEO TALKS

Ian Mertz

Abstract

Video talks, a staple of modern conferences and a great accessibility tool, are nonetheless maligned for being inferior versions of traditional live talks. We push back on this from two perspectives: first, that with clearer goals in mind video talks can be made to perform different functions than live talks; and second, that with basic editing video talks can be every bit as good as live talks, and can even reach creative heights unachievable in non-recorded settings. We make the case that creating such purposeful, engaging video talks is a benefit to the field, and should become more normalized therein.

1 The Current State of Video Talks

Since the onset of the COVID pandemic in early 2020, prerecorded talks have become a regular feature of TCS conferences everywhere. In a time where a large fraction of the community does not travel, whether for health reasons, home obligations, budget restrictions, environmental concerns, visa issues, or the shifting political climate of the hour, there is much to appreciate about this dual component to our live events. Even without these considerations, having talks which are rewatchable on demand, which can be slowed down or sped up at will, and which can reach anyone in the world (including those not on the conference circuit) is a great boon.

It is also widely agreed that these talks, frankly, suck. Trying to conjure up the energy of a live talk while going through slides in a silent, empty room is nearly impossible, and the effect is made worse by a disparate audience often watching in silent, empty rooms themselves. Every fumble hangs in the air, technical issues and poor mic placement quickly become grating, and the impersonality of it all can be more tangible than the content itself. If a traditional slide talk is meant to be polished, as a recording it often feels more mechanical than anything else.

This situation reinforces, and in turn is reinforced by, the unspoken yet assumed role of video talks, made especially clear since the overall return to in-person conferences: video talks are a useful yet inferior substitute for the live presentation, designed chiefly for those who cannot attend or for those scouting out which sessions they will attend. Only a handful of conferences, such as ITCS,

have given distinct roles to the in-person and video talks, where one serves explicitly as the condensation of, and advertisement for, the other, and even in these cases there is little emphasis put on the recorded portion on the conference website, during the submission process, etc. In short, recorded talks are treated with a utilitarian purpose, whether for accessibility at the conference or a quick memory refresh down the line, rather than something to be produced for, or considered on, its own merits.

In this paper we seek to push back on all of these assumptions and practices. We will argue two closely entwined points. First, video talk requirements should be designed with clearer purposes in mind, including ones orthogonal to, but no less important than, those of live talks. And second, in this capacity video talks can be not just good but spectacular, and can take advantage of the unique advantages of prerecording to even surpass live talks in many ways. We believe that making better and more purposeful video talks is a great potential asset for our field, and that, for reasons of equity, creativity, and outreach, we should make such content more respected, encouraged, and, ultimately, commonplace.

2 Why Make Video Talks?

We begin with the question of purpose. To start, we should ask: what is the purpose of a talk in general? While there are immediate ideas that come to mind—that they give the motivation and ideas of a paper at a higher level, for instance—the real answer is that it entirely depends on the format and context.

Let us take an example. Imagine a well-designed conference talk, timed and practiced to 30 minutes on the dot: 10 minutes of introducing and motivating the model, 5 minutes of prior work, 2 minutes of explaining the current results, 10 minutes of high-level proof ideas, 2 minutes of extensions and other implications, and one last minute for open questions. This talk is occurring at a top-level conference for the subfield, meaning much of the core language is shared but most people will come from outside the specialization, and thus over half of the talk is setup and the proof is restrained to the big ideas, yet leaving enough to add in a few hooks for the experts in the audience.

Now imagine the same speaker is asked to present the same work for a flash talk of only 5 minutes; is it correct to linearly scale everything down, with 3 minutes of setup, 2.5 minutes of proof, and thirty seconds of wrap-up? What about an invitation to give a two-hour technical seminar; is it correct to leave any mention of the proof until after the break, filling the first hour with more expansive details about the model and prior work?

Specifying the length of a talk is not just a matter of time constraints, but rather a signal as to the purpose of the talk itself, such as whether it is meant to be an

advertisement or a fully technical deep dive. The best talks use this knowledge, just as much as the target audience and the contours of the result and subfield, in a fundamental way.

Timing is just one technical component of a presentation, and this question directly applies, and is perhaps more widely understood, when considering different mediums. Just as adapting books to film requires understanding the differences between written and visual medium, a paper and a talk are fundamentally trying to fulfill different goals even when given comparable constraints on the amount of material to be covered. This further applies between variants within a given medium; within the realm of talks, a slide talk is a more polished and prepared exposition, while a board talk is typically more conversational and interactive; meanwhile plenty of written forms exist beyond papers, such as lecture notes and blog posts, both of which are more high-level but with the former typically discussing proof/technical ideas and the latter more context-driven.

Going back, then, to prerecorded talks, the question becomes what purpose we want to accomplish with a conference video requirement, taking into account what video itself is best suited for, and how to set the specifications around these goals. The most obvious factor of a video talk is its medium, namely an audio-visual presentation probably focused around slides and designed for a technical audience, but frequently we stop there and lump them together with slide talks without further ado. This also leads to time constraints being set in line with the corresponding live presentation, but unlike the conference version there are no obvious reasons to put such limits on a video (one byproduct of this is that people frequently overrun these time constraints, as they undoubtedly would, if they do not already, given the option in a live setting). Thus rather than deciding the length and type of talk as a result of the time constraints inherent to running a conference, recording and uploading videos automatically gives us the freedom to work backwards and set these restrictions based on the purpose we want these talks to fulfill.

Even more salient than the lack of hard time constraints is the relationship of presenter to audience. On one hand, the creator has unprecedented level of control over a video; on the other, it is the ultimate model of one-way information, akin to papers in lacking even basic feedback. We can consider a spectrum of styles, where casual, relaxed interactivity constitutes one pole while a tightly curated information channel forms the other, on which we can consider any given medium. For written works this may consist of forum/social media discussion on one end and journal publications on the other, with blog posts and lecture notes sitting in between; for presentations, discussion groups and videos sit at opposite extremes and with board and slide talks falling in the middle towards the discussion and video ends respectively. This does not prescribe videos to being the most purely informational form of talks, but rather encourages us to look to practices in good

paper-writing—specifically those practices which sit in opposition to more casual forms—to inform how we might use this spectrum to our advantage.

With these ideas in mind, just as we imagine conferences whose purposes are specialized and distinct from simply soliciting top theory papers writ large—again ITCS comes as an example—we can imagine a myriad of alternatives to recording 25 minute slide talks in an empty room. A conference can solicit 5-minute advertisements for the main talk, in which case we can look to actual advertising practices and hone in on a blitz of big takeaways with broad appeal, respecting the listener’s time by giving them everything they need to decide whether or not to commit to the longer session. We can flip the script and have the live talk be the advertisement for a longer technical presentation on recording, which can serve as a lasting guide to deeply reading the paper. In service to the community, we can even ask for videos which focus only on preliminaries and surveying past results, becoming a go-to reference for understanding a subfield long after the conference is over.

Besides a call to conference organizers to revisit the question of why and how to solicit video talks, as presenters we can think about how best to use video not just as an reflection of the live presentation but as a complementary product serving purposes the live presentation cannot meet. Beyond the clear accessibility benefits we all accept from video talks, there is a freedom to explore them as a completely different mode of communication, one which utilizes both time and technical constraints to tightly curate the viewing experience.

3 Can Video Talks Be Good?

We now move to the second question: can prerecorded talks ever be good enough to make these freedoms and benefits worthwhile? We argue that, alongside leaning into the aforementioned differences between prerecording and traditional talks, the ability to edit can make videos truly shine. This holds along a large spectrum of approaches: basic editing, i.e. cutting out dead space and bad takes, is doable with low additional effort and would immediately make many videos more enjoyable to watch and rewatch; while for the truly motivated and inspired, making a fully formed piece of video content can be both personally rewarding and of incalculable benefit to the field.

3.1 Basic Editing: Low Cost, Major Benefits

Let us start from the classic image of a recorded talk, a speaker presenting a slide show to an empty zoom call. A practiced talk generally has few flubs, but when speaking to the computer it seems that many people will inevitably misspeak a few

extra times, especially since such a mistake, hanging awkwardly in the air, tends to beget another. The more devoted may stop and restart after a take with too many errors—such is one of the more obvious advantages of recording—but even for such speakers, or in fact especially in cases where a speaker is willing to spend extra time and effort to produce a polished product, the speaker’s engagement can visibly begin to wane, important lines or explanation get skipped without notice, and their voice, bereft of earlier sharpness, starting to trail off as the technical section goes on. And in this situation, one overly long pause, possibly coming minutes after the stated time limit has been passed without comment, is space enough for the audience, sitting at home with a playlist of videos to watch the week before the conference, to zone out or simply move on.

The absolute basic tenant of editing is the ability to cut a recording into arrangeable and malleable clips, from which the final video is assembled. From this alone we can immediately tackle many of the issues plaguing our previous speaker. A flub or pause is no reason to restart, as one can simply stop and restart the sentence, cutting out the bad take in post-production. If they feel that they are making too many errors, or that the video is starting to drag, the speaker can stop, go take a break, and come back some time later to continue shooting. In some important sections it may even be worth stopping line by line, in order to make sure that each one comes across clearly; this has the added benefit that the cuts create a slight discontinuity which can snag the viewer’s attention before it starts to drift. Missed lines can be recorded and added even days later, pacing can be controlled through massaging the space between and during lines, and fitting timing constraints need not require timing the talk so precisely ahead of time, as less crucial lines can be removed at the end of the process.

From here it is a small step to add additional minor elements that break up the monotony of a longer talk. For those who bolstered their Zoom setup during the pandemic by buying a separate camera and microphone, patching audio and video together is just a few clicks on any editing software. Adding external sounds, such as a snippet of music in place of speaking between talk sections or as an opening grab, is equally trivial. And playing with the prominence of video elements can bring an entirely new dimension beyond slides, from looping small animations for key subroutines, mimicking a “board talk” through recording a live derivation on a tablet, running code live in the background to understand performance guarantees in situ, or simply having the presenter front and center with slides and other technical parts only appearing as necessary.

There is, of course, a cost to editing, namely time and effort. While these basic edits are routine for those experienced with editing, there is still a learning curve and the first few videos do undoubtedly take longer, but resources are available to get over the hump. Sophie Huiberts has designed a suite of tutorials in tandem with this piece such that those interested can quickly progress and start knocking

out edits as a matter of course. Universities often have their own resources, both in terms of equipment and software and with regards to tutorials, in-house video producers, or even funding to hire editors, a minuscule fraction of the total cost of attending the conference itself. And when calculating the overall effort, it is worth remembering that even early unpracticed editing may be faster than constantly restarting, and with the latter coming across much cleaner in the final release.

3.2 Video Production: Creativity Unbound

If basic editing is the floor when it comes to videos, then the ceiling is very high indeed. Far from being bound to a slide deck, modern video editing software and accompanying tools allow virtually limitless creative freedom, a freedom which, in turn, can be used to convey information in ways more tailored to the audience and content at hand.

We can look to the top three most viewed STOC video talks on the SIGACT YouTube channel, all of which utilize completely distinct styles borrowed from popular science videos on the same platform, for examples of this creative potential. On the topic of shortest even cycles we have no human face but a tightly edited audio set over beautiful visualization software, in this case the Manim package developed by the hugely popular 3Blue1Brown channel (in fact this software is already in use for many live popular talks as well). The smoothed complexity of the simplex method video, by contrast, uses a face-plus-overlay approach to counteract the “impersonality” of recordings by putting the presenter front and center and using an easygoing conversational presentation style; this appears in popular videos like those of Dr. Angela Collier, who was the stated inspiration for the video. And a variety of experimental essay-style videos, particularly those by political creators, have become one of the hallmarks of YouTube as a whole in the modern era, with a crossover attempt¹ on the topic of space complexity appearing for the STOC audience as well.

Experimentation, in this case with form and presentation, is itself the essence of good science, and therein carries on a tradition—perhaps now less common with papers but still alive and well with blogs, threads, and lectures—of finding new ways to engage with audiences and teach mathematics. This is in many ways a necessary experimentation, as each field, result, and presenter is unique and should be handled as such. This meta-form of scientific discovery, moreover, is egalitarian in a way that the cutting-edge science it teaches inherently struggles with, namely that it can come from novices with a creative vision or even by

¹This video, made by the present author, is shot over four locations and features costume changes, music, a sped up tablet section, color correction, audio balancing, and more; judgments on the success of this endeavor, however, are left to the reader.

chance rather than demanding years and years of study with near surety. This holds doubly so for video, a medium where, if anything, younger researchers have a massive edge with regards to intuition and technical know-how. Not every video will succeed in the creative goals it sets out to achieve, but there is little lost in an unsuccessful experiment besides effort on the part of the presenter, and this inevitable cycle of failing and learning is as integral a piece of the research process to teach to junior researchers as the science itself.

4 Why Put Effort Into Video Talks?

Finally, even accepting that recorded talks can be made both purposeful and polished, it seems nevertheless reasonable to ask whether making purposeful, polished recorded talks is worth the corresponding time and effort. Such videos are an unambiguous good for the viewers themselves, akin to the clarity that comes from a well-written paper, but even from the presenter's perspective this work demonstrably pays important dividends.

To frame this discussion, we must acknowledge first and foremost that these videos do, in fact, perform well. Revisiting the three STOC videos discussed above, we note that, excluding live recordings of plenary talks, these are the three most watched videos on the SIGACT page, with 2000-2500 views each to date. By contrast, the first non-edited talk, again excluding plenary and live recordings, clocks in at 1200, and even the most popular recording of any best paper talk over the past four years sits at just over 1000. Besides having double the view count, 2000 viewers is *four times* the average live attendance at STOC itself over this time; in fact we could confirm that for at least two of these three edited videos, their view counts had surpassed their respective live attendances before those conferences had even started.

Thus from the perspective of the speaker, this positive viewer experience directly translates to more exposure of their ideas to the scientific public. Many first-day conference discussions revolve around which videos people have watched and are interested in attending as a result—such is the goal of prerecorded talks as live talk advertisements, after all—and so more eyes can mean greater attendance for the “real” presentation as well. When a talk, whether live or online, stands out from the crowd, it also leaves a greater impression on viewers that increases the chances of them revisiting the result later; in this, a strong recorded talk can perform double duty, as it will also be the version any viewer who wants to revisit a presentation invariably returns to.

Beyond this, however, it is interesting to consider the further reach of those ideas to the public writ large. As mentioned above, popular computer science and mathematics thrive on platforms like YouTube, with well-produced videos at all

levels of rigor and detail garnering massive view counts; to take one example, a video by the generalist producer Vsauce on the Banach-Tarski paradox has the second-highest view count on the channel, having been watched over *46 million times*. These videos are often further disseminated through a variety of hubs, such as video competitions and festivals, amateur forums for boosting and discussing such topics, reactions and other “follow-on” content made by viewers, and other casual channels generally not touched by the more secluded work at the forefront of the field. And they are monolithic in neither style nor purpose; these can run the gamut from popular science articles which give viewers a taste of high-level research to more thorough investigations of a particular question or field.

In choosing a length and level of technical depth, in choosing what are the necessary preliminaries and which gritty details to sweep under the rug, above all else these presenters are choosing a target audience from among the internet public, an unrepresentative yet extremely broad swath of the broader population, and building a talk to fit both the purpose and aesthetic tastes of that audience. While STOC or FOCS attendees, for example, should be expected to have more background and mathematical maturity than an average viewer on YouTube, this approach still fits squarely with the question of speaking to a general theory crowd, and so there is no reason a well-constructed video talk hosted on a public platform cannot act as both within the field as a paper overview and externally as a more-technical-than-average advertisement for the field. At the very least, the success of such “amateur” videos should, perhaps, prompt us to look at the power of making recorded talks with purpose and quality in mind.

5 The Future of Video Talks

The current essay began from a thought, namely that it seems strange that many of us have collectively decided that a talk which we can edit—a talk where we can cut out dead space, redo bad takes, or restructure after reviewing a full draft—is *inherently worse* than a traditional talk. It is one thing to assert that awkward pauses, vocal tics, and other interruptions are more halting in a video than live, but it is quite another to assert that a video without these interruptions is still worse than a live talk laden with them, let alone when the video allows itself the audiovisual freedoms of the medium. And while it may be intimidating or even paralyzing to enter into the world of editing, the costs of learning this new skill, once amortized over a handful of submissions, are more negligible than they may feel from the outside.

The prerecorded talks component of conferences are not, and for many reasons should not be, leaving anytime soon. While the field has generally made peace with this fact, we may have missed the opportunity to truly embrace video talks

as a novel frontier for presenting computer science and mathematics at the highest levels. They are not without disadvantages; even for the most engaging styles the lack of human connection stands in sharp contrast to the live talks they so often walk in the shadow of. However, in light of their advantages, such as clarity and comprehensive level of control, it seems more useful to think of them as an orthogonal, rather than inferior, medium. And with in-person conferences still featuring live talks front and center, we have no need to pit these orthogonal modes against one another: we can really have it all.

Insofar as we accept and embrace video as a mode of presentation, it seems only natural to put in the effort to make it work. As with both papers and live talks of all stripes, videos with a baseline level of production and editing are more enjoyable and should be lauded, while those that go above and beyond and succeed therein should be viewed as perennial classics. While there may be a chicken-and-egg problem with winning such acceptance, the existing forays into this realm, both inside and outside our community, are already viewed in this way; the question, then, is simply whether we choose to make them the exception or the norm.

This is not meant to be prescriptive of how such videos look and feel. There is no single correct way to make a video just as there is no correct way to write a paper or give a talk; it is the diversity of presentation styles, along with experimentation therein, which drive the field forward. However, if we have resigned ourselves to feeling that every way to make a video is wrong, a view that is both untenable and plain incorrect, then it is our contention that we all benefit from working towards a new paradigm where videos are more meaningful, impactful, and, above all else, enjoyable.

Afterword: A Call for Videos for EATCS

This piece accompanies the beginning of the new EATCS YouTube channel for pre-recorded talks, an initiative launched by Stefan Schmid and hosted by Sophie Huiberts and myself. We invite speakers to pitch potential videos at any technical level with an emphasis on clarity of presentation and quality of production rather than topic; the submission and editing process will go through the channel rather than the bulletin and will be released throughout the year. We also invite all contributors to the Bulletin to submit video accompaniments to their submitted article, or even video-only contributions, to be released alongside the written bulletin, which will go via the ordinary bulletin soliciting/editing procedure. Our primary goal is to help lead the way in videos occupying a prominent place in TCS by becoming a hub for the best video talks our field has to offer.

Acknowledgments

We thank Sophie Huiberts for years of conversations (and proofs of concept) on the topic of producing video talks and Stefan Schmid for the opportunity to publish this work in tandem to the channel launch. We thank Sophie Huiberts, Fran McManus, and Kristie Petillo for feedback on earlier drafts of this work.