Documatic: Participatory, Mobile Shooting Assistant, Pre-Editor, and Groundwork for Semi-Automatic Filmmaking

Andrew Quitmeyer
Georgia Tech – Digital Media
TSRB 85 Fifth St.
Atlanta, GA, USA
andrew.quitmeyer@gmail.com

Michael Nitsche
Georgia Tech – Digital Media
TSRB 85 Fifth St.
Atlanta, GA, USA
michael.nitsche@gatech.edu

ABSTRACT

While digital filmmaking has broadened the accessibility of documentary film and television creation, legacy production practices persist of first gathering information (shooting) and later structuring the audio-video narrative (editing). The open-source app Documatic combines shooting and editing processes via synchronized smartphone–camera annotation systems that automatically generate non-destructive video "rough-cuts" for *Adobe Premiere*. Harnessing the correct digital affordances can enable small documentary productions to enjoy some of the production efficiency traditionally belonging only to large scale filmmakers, while maintaining the artisanal quality control of a small team. The paper outlines the design, implementation, and example usage of this production and editing assistant, which is aimed at supporting small independent documentary filmmaking teams.

Categories and Subject Descriptors

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Design, Experimentation, Human Factors.

Keywords

Television, filmmaking, documentary, editing assistant, pre-editor, smartphone, collaborative, categorical documentary, procedural model, automation, Android, open-source

1. INTRODUCTION

Digital filmmaking has significantly impacted documentary television and film by decreasing the costs of production, editing, and distribution. Few digital affordances, however, have been applied to improve the actual filmmaking process. Currently, most documentary productions continue to abide by the legacy practices that were formed by the conditions of traditional, celluloid film practices. First, documentarians gather massive amounts of subject information from archival footage, recorded interviews, and text. Next, the documentarians are forced to sort through the collected data and derive a structure for the eventual

audio-video narrative [1]. While this structural synthesis period distinguishes documentary from other film formats, as a stand-alone process it can be quite arduous.

Some video logging systems attempt to address this problem of sorting through droves of audio and video. These systems, however, are typically only used in large commercial or theatrical filmmaking as they rely on pre-established concrete master structures (such as shot lists).

Following a different approach altogether, database film projects automate the structuring of video into procedurally arranged segments or present spatialized, interactive clips. With these systems, the construction of any sort of distinct narrative that could function in linear TV still requires the sorting and editing process of traditional filmmaking.

Documatic[1] is a free, open-source Android app created as a proof-of-concept for a dynamic and efficient digital video creation process. It simplifies the arduous structural synthesis process by combining it with the more exploratory, spontaneous information gathering period. Via a synchronized Android app and an arbitrary digital camcorder, annotations can be added in real-time to recorded footage. Users structure these tags and – guided by *Documatic*'s open design – generate a malleable structure for their film during production itself. Finally, as the amassed data is downloaded to a computer, *Documatic* utilizes this structure of tags with its custom-built Project Generator to create "pre-edited," rough-cut video sequences for *Adobe Premiere Pro* [2]. In that way, it utilizes digital affordances during production to allow for a new form of content assembly. Since annotations utilize universal timestamps, the system can be paired with any digital camcorder, and audio-visual fidelity need not be sacrificed or limited to the capabilities of current phone cameras.

Documatic's structural basis evolved from theories and research of analogue documentaries in order to preserve the cinematic grammars culturally developed over the past century. The end product is not meant to be an interactive movie but will be more or less indistinguishable from a traditional, linear documentary film. The audience's ability to "read" the final artifact will be preserved, while the new formative process exemplifies simpler, more efficient ways for digital documentary production.

2. PROBLEM SPACE

Even though both, small independent films and large studio productions can be shot, edited, and distributed completely digitally, a major production company is still able to create cinematic content more efficiently due to its ability to sort through

---

[1] http://www.semiautomaticfilmmaking.org

and organize massive amounts of collected content. Where a small documentary team may take years to plan, shoot, and edit a 120 minute film, the sheer manpower of a large studio can produce a similar amount of content in a matter of weeks.

Concerning his thesis on the "Evolving Documentary," Michael Murtaugh describes the "traditional" method of making a documentary film: "filmmakers collect a large amount of raw material -- original film footage, archive photographs, text articles. These raw materials are organized in progressively larger chunks: shots, scenes, and sequences. Finally, sequences are edited together to form the final "cut" of the film. [...] In this way the filmmaking process may be seen as a kind of funnel [where] a large collection of content [...] is gradually refined and reduced" [3]. This is a work- and time-intense approach.

Another hindrance arises in that documentaries, in particular, tend to be produced in a more dynamic fashion than large theatrical studio films. A fiction film can follow *a priori* storyboards and scripts as blueprints for shooting and editing. Documentaries, on the other hand, must allow for unforeseen events and interviews with outcomes far different than could have been planned.
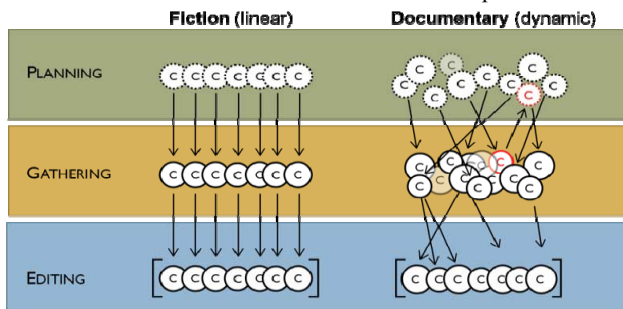

Figure 1. Differing content management in fiction and documentary video production

Whereas fiction film production has pre-assembled content (e.g. in the form of a screenplay and shot list), documentarians often face unforeseen changes in their content units (marked "c" in fig. 1). This applies even more to the practices of *YouTube* directed filmmakers catching unpredicted events on video.

Rigid filmmaking structures grant process efficiency to large, assembly-line-style teams but impede the work of small-team video creators. If we could however, digitally automate and/or parallelize some of the steps of the production process, the individual can begin to enjoy some of the filmmaking efficiency afforded to the big filmmakers, while maintaining the quality control and flexibility of a small team.

3. CURRENT FIELD
3.1 Video Tagging/Metadata Systems
Video logging systems, such as Adobe's *OnLocation* software [4], permit filmmakers to add supplementary information to movie files such as shot number, scene, description, camera information, timestamp, and ownership. Most of these logging systems are intended for large commercial or theatrical video productions, however, and they function as little more than a modern "clapboard." This linear style of metadata (shot, scene, and camera information) must be used with a previously established master structure since the tags themselves bear little semantic meaning. Thus, these current systems do little to aid documentary filmmakers whose movies' structures are dynamic and malleable.

3.2 Database Filmmaking
Meanwhile, other systems, like Michael Murtaugh's "evolving documentary", *ConTour*, Manovich's *Soft Cinema*, or the work of Murray's Experimental TV Lab, were developed to automatically generate cinematic experiences from information databases that reference collections of video clips [5]. These systems permit dynamic exploration of large content bases but creating these complex databases can be just as hard as editing the footage directly. Furthermore, the output of these database systems is usually a dynamic and interactive explorative moving image piece. It more resembles a web-based browsing approach and does not fit into the traditional format of linear TV and film.

While both approaches offer unique affordances, they also highlight the lack of an assistive, dynamic system for small-scale productions that could benefit from a combination of the two.

4. DESIGN
Analysis of the historical and structural difficulties in television, cinema, and documentary filmmaking, form the here-proposed design goals for a "semi-automatic filmmaking" activity. Overall, *Documatic* strives to enhance filmmaking; it does not replace the established techniques. It focuses on five core design goals: A) *Efficiency*: data needs to be intelligently collected with the automated structuring. B) *Agency*: creators should have full agency to direct the outcome of the project, and the digital component should be entirely non-destructive; that is, even if the system breaks, or performs oddly, the documentarian is still left with all the footage, as if they have shot a traditional documentary. C) *Readability*: The system's output should be indistinguishable from a traditional documentary and follow conventional film grammar. D) *Independence*: a filmmaker should not have to rely on external data networks, e.g. cell signals to capture the shots needed. This is necessary as many documentary productions operate in difficult territory such as nature reserves or foreign places where instant connectivity is not given. E) *Adaptability*: Finally, as Bernard notes [6], a system should react to the dynamics of the documentary creation process. A balance is needed to allow pre-planning of certain situations, while also allowing fast reactions to unforeseen developments.

Technically, the functionality of this product is derived from the pairing of a pool of raw, unorganized data (folders of unlabeled video content) with a collection of semantic data (xml annotations of specific time-periods) via global timestamps in order to automatically generate an editable sequence in an editor like *Adobe Premiere*. Building on the design goal of technological independence, the raw data can be collected from one-to-many digital video recording devices, while separately annotation data is captured by the *Documatic* app running on an Android device. By synchronizing the internal clocks of the desired camcorder(s) and the Android device with the current internet time (such as nist.time.gov), no further communication is necessary to allow the devices to work together.

4.1 Procedural Model: The Categorical Documentary
The first step in any automation process is the creation of rule sets. These rules are necessary to transform any continuous real-world process, into a discrete series of manageable steps that can be completed programmatically [7].

In their book, *Film Art*, Bordwell and Thompson identify a primary rule-set governing many documentary films, the "categorical documentary" [8]. This format follows a simple,

consistent pattern regardless of the topic. First, the subject is introduced, and the viewer is then presented with series of interviews or narrations grouped into topics associated with the overall subject. The order in which these sections are arranged forms the overarching narrative experienced by the viewer. The narrative flow and continuity of these films is provided by characters and themes recurring throughout the distinct categories.

Because this approach is clearly defined, yet broad enough to allow for a high level of flexibility, the categorical model provides a firm basis for organizing footage. This is why *Documatic*'s tagging system uses a basic categorical documentary approach to structure the referenced video segments. It automatically provides "Introduction" and "Ending" segments and allows users to create their own sections, which they tag while filming to group semantically similar pieces of footage across different interviews.

To aid filmmakers in editing the footage once it has been organized into categories, a supplementary, syntactical rule-set is included. This secondary rule-set is based on a study of categorical style documentaries such as Errol Morris's *Fast Cheap and Out of Control* [9]. It breaks down individual sections of a documentary into three fundamental footage elements: *Interview*, *Exhibit*, and *Narration*. *Documatic* uses these three sub-elements to automatically intelligently layer footage in the editing stages.

Interview footage serves as the bulk of the content for most categorical documentary films. It consists of video from a camera pointed at a person (typically a close-up head shot), who is making a statement or answering questions. A single interview is often cut into smaller clips and takes from different interviewees can be grouped according to topics regarding what was being said in each clip. These groupings by topic form the individual sections of the overall categorical documentary.

Exhibit footage consists of superimposed shots that illustrate what is being discussed by an interviewee or narrator.

Narration adds the filmmaker's voice as a transition between categories or explanation a particular piece of exhibit footage.

By using these simple rules to guide the production process many steps of the film material assembly process can be automated. They can support the filmmaker in terms of organization and structuring without limiting the artistic approach or content. Using a logical, yet semantic, system has the additional benefit in that individual projects can be handed off to a third-party with a clear and pre-implanted structure.

4.2 System Overview and Walkthrough

What follows is a sample walkthrough to illustrate the functionality of the *Documatic* system, based upon a set of actual use cases/user tests. "Mary", one of the first users of *Documatic*, is interested in creating a documentary about people and their dogs in the park. Mary asks her friend "Adam" for help filming the documentary, but wants main directorial control of the project.

4.2.1 Pre-Production

As a very first step, Mary launches the *Documatic* app from her Android-powered smartphone (Google's Nexus One [2.3.6]). She creates a new project called, "Long Dogs." A new project is automatically generated for her with initial "Introduction" and "Ending" sections. The team discusses possible topics and interview questions. For each topic they add another section to the overall project via the "New Section" button. Eventually, they compile a list of topics which represent both the questions they want to ask, and the sections that will comprise the final film (see

fig. 2 left). This dynamic list can be re-arranged at any time, but provides an initial structure to the evolving documentary.
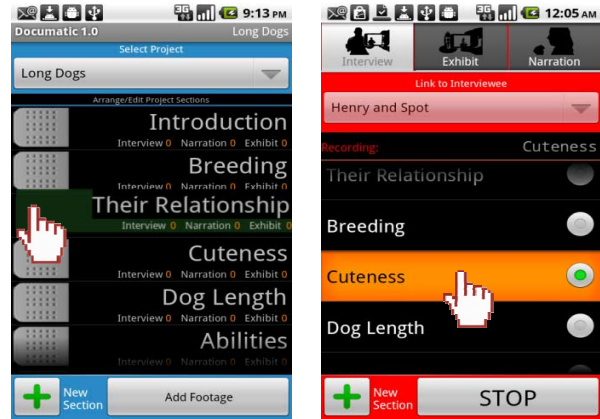

Figure 2. Example screen shots for setting up the initial structure (left); tagging during shooting (right)

4.2.2 Production: Recorder and Annotator

When the duo arrives at the park, they synchronize the clocks on Mary's phone, and Adam's camera. After they find an interviewee, Mary, acting as the "annotator," inputs the subject's preliminary information into the Android app. From now on, the new subject's information – e.g. "Henry and Spot" – can be automatically applied to shots. Adam acts as the "recorder," sets up the shot, and starts recording as Mary begins the interview.

At this point, the recorder's role is exactly the same as in a larger film production. He is responsible for the quality of the shot itself, keeping the person in the frame, in focus, and monitoring audio levels. He does not have to engage with the responses and is thus able to get the highest quality footage possible.

Similarly, Mary's job of conducting the interview is made easier through this division of labor. The initial layout of topics serves as a supportive list of talking points for her to cover, and it also provides the necessary real-time tagging annotation of the footage recorded by Adam's camera. Mary taps the sectional tag in the interface's list corresponding to the topic being discussed by the interviewee and creates additional tags as the interview evolves. If the subject, for instance, begins by discussing the breed of the pet, but then immediately starts talking about what factors affect its cuteness, Mary simply taps the "Cuteness" menu item, and any video being recording during this time is automatically categorized into the "Cuteness" section and linked to the subjects, "Henry and Spot." While footage is being annotated, the theme of the user interface flashes bright red to indicate that virtual clips are being recorded. New sections can be added on the fly, even in the middle of an interview.

After the interview, one can use the system to collect *Exhibit* footage that illustrates what the interviewee was discussing, and these clips will be automatically linked to that specific person. The documentarians can also optionally record audio snippets of *Narration* to introduce specific sections of the film for later use.

4.2.3 Post-Production: Pre-Editing / Refinement

Whether the team is prepared to produce finalized, distributable video, or just wants to get a quick feel for how the video is coming together, the post-production process is made simpler with the *Documatic* system. One first copies the project's folder from the Android device to the editing machine. Next, they copy the "raw" video files from the camera's card to this folder. Then, they can use a Project Generator program. *Documatic*

automatically synthesizes a new, "pre-edited" XML sequence based on the information from the tags and the corresponding time code on the raw footage. Dragging the new sequence file into a Non-Linear editor like *Adobe Premiere CS5*, reveals this labeled, annotated, and pre-structured video sequence (fig. 3).

Technically, all the arranged clips are tagged subclips of the full video files. If anything important was cut off a particular clip by the automatic generator, the human editor can adjust the in- and out-points of any subclip at any time.

Figure 3. Outcome: automatically, intelligently segmented video, with pre-generated name-tags for interviewees

Unlike rigidly structured video systems, the documentarians can go back and forth throughout these Pre-Production, Production, and Post-Production phases at will. At any time they can generate a full sequence, plan new topics, or collect more interviews and add them to the editing project.

4.2.4 Collaboration

The *Documatic* system is not a unidirectional process. Instead, it harnesses Murtaugh's "Evolving Documentary" idea to let documentarians iteratively augment and share their works. Multiple project folders from distinct film crews can be copied together, and the Project Generator synthesizes them together intelligently. Since the Documatic production process is based on structuring the collected footage in a meaningful way (as opposed to the pure syntax of shot/scene numbers), it can be much easier to share or pass along a documentary project to other filmmakers without prior knowledge of an overall theme or outline.

5. DESIGN EVALUATION / EVOLUTION

One of the most important discoveries made during the design process and the testing phase of *Documatic* was the value in exploiting the unique timestamps. In the initial phases of the design, there was much focus on determining the optimal way to directly communicate and pair the semantic data of the annotations with the raw data of the video files in real-time. Several different systems were designed to use either Bluetooth or network communication to share these two sets of data between phones, but none of these methods met the design criteria for agency and independence.

To better situate the annotator's attention more on the interview than the device, the UI's design went through several iterations. Haptic feedback and categorical clustering led to a less obtrusive interface.

6. Extensions

Since *Documatic*'s underlying framework is based on universal timestamps it can be adapted to a wider context. The interactor's tags can be used not only as annotations of a specific video, but rather of a unique *time* and *place* in history. This opens up several ways to extend the system to encompass a wide variety of additional features and uses.

For large events, like a political protest, one could have an indiscriminate amount of people, filming and annotating throughout the day. Then they could separately upload these video files, and time-stamped XML annotations to a central server. Based on combinations of this data interesting views of the day could be automatically generated representing individual or merged experiences from the group as a whole. Films like *Burma VJ* [10] or the Beastie Boys' documentary [11] would be ripe for this approach to digital participation.

A documentary is just one possibility of this fully digital video production process. By modeling other cinematic genres, one could begin producing digital, semi-automated sitcoms, thrillers, or dramas. In fact, one could create a procedural model creation system, where users generate their own rule-sets of custom complexity to replace the current categorical model – leading to an auto-generation of custom-build content tailored to their own preferences. A single audience member might query a large database with her list of interests, which are checked against the existing tags, and a pre-cut movie could be generated for the viewer. This supports the database approaches mentioned above.

7. CONCLUSIONS

The most important aspect of *Documatic* is that it provides a complete workflow for producing and assembling video data procedurally. By pairing annotation with video, and forming intelligent rule sets, it simplifies the labor intensive video-editing process. However, it remains an assistive system that supports but not replaces a filmmaker's creativity. It thus avoids any limitation to the filmmaker's ideas while easing the workload.

8. ACKNOWLEDGMENTS

9. REFERENCES

[1] Murtaugh, M. 1996. *The Automatist Storytelling System: Master's Thesis*. Cambridge: MIT.

[2] Adobe. *Premiere Pro CS5*. http://www.adobe.com/products/premiere.html

[3] Murtaugh, M. 1996. *The Automatist Storytelling System: Master's Thesis*. Cambridge: MIT.

[4] Adobe. *On Location* http://www.adobe.com/products/onlocation.html

[5] Manovich, L. (n.d.). *Form*. Retrieved Sept. 2010, from Lev Manovich: Soft Cinema: http://softcinema.net/form.htm

[6] Bernard, S. C. (2004). *Documentary Storytelling for Film and Videomakers*. Oxford: Focal Press.

[7] Manovich, L. (2000). *Language of New Media.* 27

[8] Bordwell, D., & Thompson, K. (2004). *Film Art.* 132

[9] Morris, E. *Fast Cheap and Out of Control.* 1997.

[10] Ostergaard*,* Anders. *Burma VJ.* 2010.

[11] Yauch, Adam. *Awesome I Fuckin' Shot That!.* 2006