

Motif: Supporting Novice Creativity through Expert Patterns

Joy Kim¹, Mira Dontcheva², Wilmot Li², Michael S. Bernstein¹, Daniela Steinsapir²
Stanford University¹, Adobe Research²
{jojo0808, msb}@cs.stanford.edu, {mirad, wilmotli, steinsap}@adobe.com

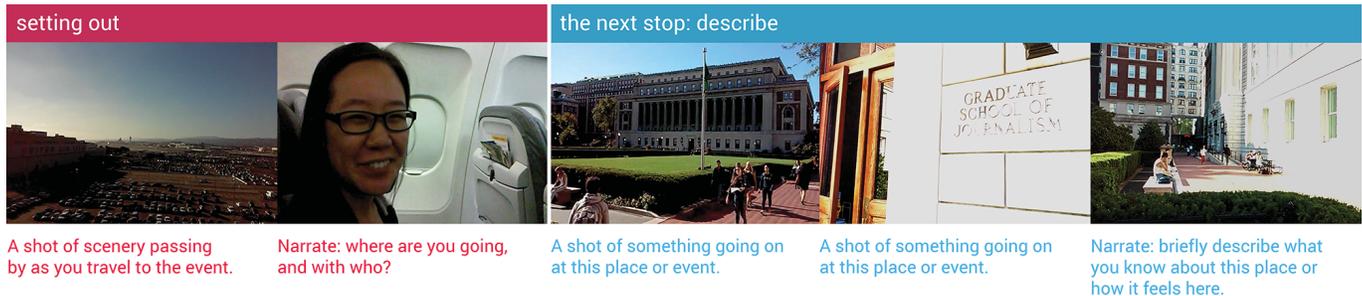


Figure 1. In Motif, a user composes each section of a video story based on patterns extracted from expert work.

ABSTRACT

Creating personal narratives helps people build meaning around their experiences. However, novices lack the knowledge and experience to create stories with strong narrative structure. Current storytelling tools often structure novice work through templates, enforcing a linear creative process that asks novices for materials they may not have. In this paper, we propose scaffolding creative work using storytelling patterns extracted from stories created by experts. Patterns are modular sets of related camera shots that expert videographers commonly use to achieve a specific narrative function. After identifying a set of patterns from high-quality storytelling videos, we created Motif, a mobile video storytelling application that allows users to construct video stories by combining these patterns. By making existing solutions used by experts available to novices, we encourage capturing shots with story structure and narrative goals in mind. In a controlled study where we asked participants to create travel video stories, videos created with patterns conveyed stronger narrative structure and were considered higher quality by expert evaluators than videos created without patterns.

Author Keywords

Novice creativity; video stories; storytelling.

ACM Classification Keywords

H.5.2. User Interfaces: User-centered design

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

CHI'15, April 18–April 23, 2015, Seoul, Republic of Korea.
Copyright is held by the owner/author(s). Publication rights licensed to ACM.
ACM 978-1-4503-3145-6/15/04\$15.00
<http://dx.doi.org/10.1145/2702123.2702507>

INTRODUCTION

Telling stories about personal experiences can help people reflect on their lives and build a shared history with the people around them [12]. For this reason, people often desire to create artifacts representing their personal experiences and histories in order to share them with others [18]. Today, personal stories are often accompanied by digital artifacts, such as photos or videos. However, while the prevalence of full-featured mobile devices makes it convenient to capture such artifacts, assembling them into compelling stories remains a difficult task. Expert storytellers are able to use narrative structure to guide an audience through an event, but novices typically lack consideration for narrative pacing and structure. They make mistakes such as starting and ending a story abruptly or allowing sections of the story to drag on for long periods of time.

A key challenge in creating stories with good narrative structure is capturing the right types of content. While experts typically have a narrative structure in mind and can anticipate what they might need to express that structure (e.g., footage of passing scenery on the way to an event to provide context at the beginning of a video story), novices often do not think about capturing such content in the moment and miss opportunities to do so. Moreover, even if a novice storyteller ends up capturing the right types of content, they may not know how to combine it to achieve specific narrative goals (such as highlighting a key moment by juxtaposing it with a shot of viewers' reactions).

How can novices create better stories? Existing story creation tools for non-experts, such as iMovie Trailers [3] and Animate [1], provide templates that help novices arrange captured content in a heavily structured way. However, these tools only guide users in arranging content they already have and leave users to decide the appropriate types of content for

their stories on their own. Thus, while the output from these tools contains aesthetically pleasing elements, such as stylized transitions or split-screen effects, it typically suffers from the same lack of narrative structure as amateur videos made from scratch. Another drawback of existing template-based tools is that the templates themselves are typically very rigid, forcing users to assemble a fixed number of video clips or images with specific properties in a specific order. While this prevents users from making common mistakes such as including interminable shots with very little action, this also leaves little room for novices to express their own ideas and adjust the story's structure to suit different storytelling needs.

In this paper, we propose an approach to help novices tell stories in a guided way that allows for creative freedom. Our hypothesis is that we can *scaffold novice creative work using patterns extracted from work by experts*. We explore this strategy in the domain of video story production, which is a particularly compelling application area given the prevalence of mobile video capture and sharing, and the difficulty of producing such stories well. In our approach, a *pattern* (Figure 1) is a set of related camera shots that expert videographers commonly use to achieve a specific narrative function. For example, the *Setting Out* pattern is often used at the beginning of video stories to introduce an event and why the storyteller is going to it. It typically includes a shot of scenery passing by as the storyteller travels to the event, along with narration describing where the storyteller is going and with whom. In general, pattern languages describe the knowledge of a typical creative expert in the form of a set of recurring solutions to common problems in a design domain [6]. Unlike templates, *patterns are modular ingredients that can be combined to form a larger creative solution*.

We demonstrate our approach with a video storytelling application for mobile phones called *Motif*¹ (Figure 2). Motif comprises both a story construction tool and a story capture tool. In Motif, the patterns we identified in expert videos are manifested as *story blocks*, which each consist of a set of suggested shots needed to render a pattern. In story construction mode, users can arrange these story blocks to create the outline for a video story, essentially forming a checklist of what a user needs to capture that appears while the user is in capture mode. By strongly linking the activities of capture and construction into the interface, Motif allows users to adapt their story structure as they capture content and vice-versa. In addition, these suggestions may also encourage users to record content they may not have recorded otherwise; novices are supplied with the strategies for what to capture and are only responsible for filling in the pattern with their own content without needing to manually slice video files or generate creative decisions from scratch.

In a controlled study, we asked 13 people from two different locations to create a short travel video story about popular tourist destinations in their home areas. Participants were randomly assigned to use either Motif, where stories were scaffolded by patterns extracted from expert work, or a control version of Motif where stories were not scaffolded. Through

¹<http://motifstory.stanford.edu>

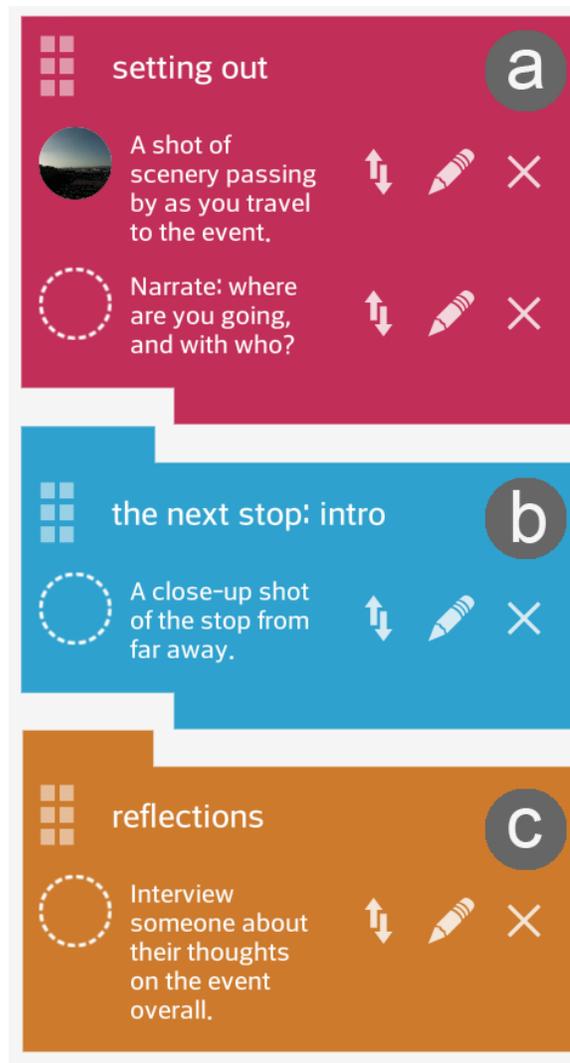


Figure 2. Stories are composed of story blocks in Motif. (a) A Beginning story block. The first shot in this block has already been filled in. (b) A Middle story block. (c) An End story block.

analysis of survey responses, interviews, and story structures present in created videos, we found that videos created using patterns were higher-quality than videos created without patterns, conveying stronger narrative structure and depicting scenes using multiple shots and angles.

RELATED WORK

There are a number of online communities and mobile applications that have made it easy to organize and capture digital artifacts. Applications such as Timehop [5] and Memoir [4] offer ways of retrieving digital content from the past and organizing them by location or time for viewing in the future. Other applications, such as Animoto and Explory [1, 2], have attempted to formalize the slideshow sharing behavior mobile users often employ when sharing photos [7] by creating tools for making narrated or text-supported slideshows.

However, these tools often focus on quick capture and attractive presentation rather than supporting story construction and

planning. How-to books on home video recording or filmmaking often instruct readers to plan their film’s visual content ahead of time with a rough outline so that moments can be anticipated and captured [15]. Veteran scrapbookers often think ahead about how they would want to make scrapbook pages about the event they’re experiencing in order to make sure they capture the right photos and save the right pieces of memorabilia [19], preventing situations where they miss chances to capture something they wanted to remember. This contrasts against the way most people record their lives — without much forethought to what they would want to remember later and how they’ll remember it.

Meaning-making through story construction

The capture and editing phases of a video are often considered separate activities. However, expert videographers view these two phases as intricately linked, where capture incites changes in narrative and where changes in narrative change the plan for capture. The mindful documentary model [8] captures this iterative process of story construction by using commonsense reasoning to suggest storytelling options available for a videographer in real-time. However, enabling real-time suggestions requires videographers to create text annotations for the shots they are capturing. Motif, instead, attempts to make this iterative storytelling process accessible to novices by providing mix-and-matchable patterns for use out-of-the-box.

Past work has also looked at how capturing and handling material for a story helps people construct memories of the past. The activity of building a personal history using a timeline metaphor can provide a vehicle for identifying and reminiscing on key events [17]. Asking families to create time capsules revealed that people did not create an exhaustive record of the past but instead tried to create a detailed representation of a single theme, such as “a typical day” [13]. In addition, many families constructed new artifacts for the purposes of the time capsule rather than logging past content [11]. The suggestion is that memories are not fixed but continuously reconstructed through building and interpreting narrative cues. Motif currently focuses on facilitating this relationship between reflection and story construction for story creators (rather than for an audience) by encouraging users to leave capture mode at any time to change their story’s structure based on new or unexpected events.

Patterns as constraints

A pattern language describes a set of well-worn practices within a field of expertise [6]. Unlike spoken or written languages, pattern languages enable navigation of a design problem rather than communication. Motif employs a simple pattern language to guide novice work: suggestions given to novices can be thought of as a vocabulary, and the constraints Motif places on how these suggestions are put together represent a simple grammar. Use of this simple language places constraints on the creative options available to novices, which may both prevent the novice storyteller from being overwhelmed and allow novices to comfortably experiment within established boundaries [16].

STORYTELLING PATTERNS

Experienced video storytellers have developed tacit and explicit knowledge about their craft. We suggest that, like architecture [6] and web design [10], storytelling patterns can be codified and made more accessible to non-experts. In this section, we discuss the patterns we extracted from expert videos and their constituent elements. Next, we demonstrate how these patterns are instantiated within Motif.

Identifying video storytelling patterns

To develop our set of storytelling patterns, we examined 13 short video stories from the New York Times, two Rick Steves’ Best of Europe travel videos, and two high-quality home videos about a group camping trip and visiting a local food festival (see supplementary materials for links). We explicitly chose a variety of videos to ensure that we generated patterns that were both common and specific across several genres of stories. We included home videos in our corpus because amateurs may have different storytelling goals than professional filmmakers; we anticipated that home videos might reveal different kinds of patterns than professionally produced videos.

Through a grounded analysis of these videos, we created a list of groups of shots that worked together to convey a coherent narrative idea or function. These groupings were detected by looking for scene changes, as they often marked where one narrative chunk ended and another began (e.g., moving from the introduction of the story to an interview with an event attendee). We also focused on repeated narrative elements, since comparing similar pieces of the story could reveal common elements between them (e.g. comparing several different interviews to identify the kinds of shots an interview typically contains). Shot groupings were selected on the basis of storytelling purpose (e.g., a pattern for how interviews are introduced) rather than on visual qualities (e.g., a pattern for showing scenery). We took note of how each shot was framed, the narrative purpose for each shot, and the length of each shot. We then consolidated groupings with similar functionality to create a list of 24 patterns.

Early iterations of the patterns were fairly general. For instance, early patterns included things like *Action/Reaction* (i.e. showing something happening and viewers’ reactions to it) or *Contrast* (i.e. juxtaposing two ideas, situations, or things). We assumed that users would know when these patterns would be best appropriate. However, with early pilots of Motif, we observed that these patterns were too disconnected from the actual situations people wanted to capture. So, we instead focused patterns around specific narrative purposes. This resulted in more targeted patterns which were more easily recognizable, including: *Setting Out* (i.e., leaving for a trip), *Travel Break* (i.e. moving between two locations), and *Serve* (i.e., putting freshly-cooked food in front of people).

For each pattern, we identified the *elements* that were required. For example, a *Travel Break* pattern, which is used while travelers are enroute to a destination, requires 1) a shot of vehicles leaving the current location, and 2) a shot of the surroundings passing by, taken through the window. We noted the time lengths of each element as they appeared in

Pattern	Type	Elements	Example	Timeline
<i>Setting Out</i>	Beginning	<ul style="list-style-type: none"> ■ Shot of scenery passing by as you travel to the event. ■ Narrate: where are you going, and with who? ■ Text: title of event 	GCC-family (0:06–0:32)	
<i>How to Make</i>	Beginning	<ul style="list-style-type: none"> ■ Shot of the final dish. ■ Text: name of dish 	Cuba-Libre (0:03–0:07)	
<i>Set the Location</i>	Beginning	<ul style="list-style-type: none"> ■ Record yourself briefly describing the place this video will be about and why it's so great. ■ Narrate: a short outline of the trip to come. ■ Short shot showing the first part of the trip. ■ Short shot showing the second part of the trip. ■ Short shot showing another part of the trip. 	RS-Italy (0:00–0:12)	
<i>Key Moment</i>	Middle	<ul style="list-style-type: none"> ■ Shot of something going on during this key moment. ■ Shot of something going on during this key moment. (optional) ■ Shot showing people's reactions to the key moment. 	GCC-family (28:57–30:20)	
<i>The Next Step</i>	Middle	<ul style="list-style-type: none"> ■ Shot of the cook grabbing the next ingredient or tool needed for this step. ■ Shot of the cook adding the ingredient or performing this step. ■ Text: name of ingredient or tool 	Cuba-Libre (0:07–0:12)	
<i>Travel Break</i>	Middle	<ul style="list-style-type: none"> ■ Before you head out, record a shot of vehicles heading out to your next location. ■ While you're travelling, take a shot of your surroundings through the window. 	RS-Spain (27:09–27:32)	
<i>Reflections</i>	End	<ul style="list-style-type: none"> ■ Interview someone on their thoughts about the event overall. 	Food-Fest (3:10–5:20)	
<i>Serve</i>	End	<ul style="list-style-type: none"> ■ A shot of the dish being served or placed on a table. 	Cuba-Libre (0:21–0:25)	
<i>Highlight Reel</i>	End	<ul style="list-style-type: none"> ■ Narrate: short outline of what you did on the trip. ■ Short shot showing the first part of the trip. ■ Short shot showing the second part of the trip. ■ Short shot showing another part of the trip. (optional) ■ Record yourself describing the trip and what you got out of it. 	RS-Spain (53:46–54:23)	

Table 1. A sample of story blocks and prompts available in Motif. The Example column specifies a time and expert video where the pattern can be seen (see supplementary materials).

examples and created a timeline illustrating a typical editing pattern for combining those elements into a final video (right column of Table 1).

Patterns

Table 1 presents a subset of the 24 patterns we identified. Some of them apply across a variety of story settings: for example, *Reflections*, *Set the Location* and *Key Moment* all could be used for travelogues, local events, or slice-of-life videos. Other patterns are focused on specific types of events. For example, *How to Make*, as well as *Serve*, are tailored for cooking-oriented videos, and *Travel Break* and *The Next Stop: Tips* best suit travel videos.

We classified each pattern by the narrative role it typically played in example videos. Beginning patterns often opened videos by setting the stage or providing a preview of the final goal; End patterns ensured the final part of the video had a compelling closing; Middle patterns filled in the action in between. Table 1 presents three videos of each type. Strong videos have one Beginning and one End pattern, and as many

Middle patterns as necessary in order to tell the rest of the story.

Patterns not listed in Table 1 include: *How to Make Slideshow*, *This is Perfect for...*, *The Next Step (narrated)*, *Serve (narrated)*, *Describe Travel Plans*, *The Next Stop: Intro*, *The Next Stop: Describe*, *The Next Stop: Tips*, *The Next Stop: Mini Review*, *In the Action*, *Guests/Friends Appearing*, *Guests/Friends Interview*, *Guests/Friends Leaving*, *Group Shot*, and *Time Lapse*. Our supplementary materials include descriptions of all 24 patterns we identified.

MOTIF: PATTERNS IN ACTION

We manifest our use of story patterns in Motif, an Android application that uses patterns to help structure both *construction* and *capture* of video stories. In this section, we describe how Motif transforms patterns into story blocks and guides novices during the video story creation process.

Story Blocks and Prompts

In Motif, we use two user interface elements to help users make use of patterns.

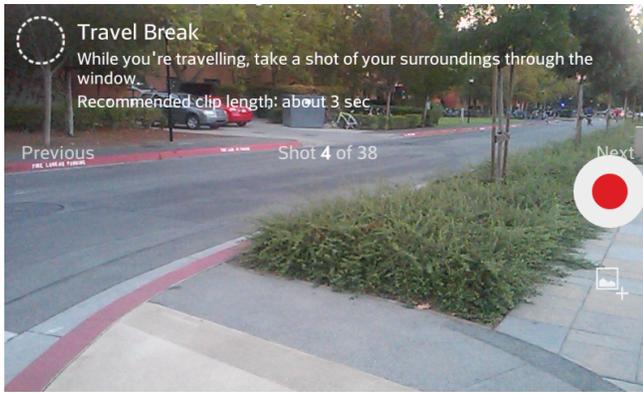


Figure 3. Prompts for story blocks appear at the top of the screen while users are in capture mode.

Patterns appear as *story blocks* (Figure 2), which are used to build an outline of a video story. They can be combined according to a simple grammar: the story must start with a Beginning block, followed by as many Middle blocks as desired, then close with an End block. This grammar enforces narrative structure while leaving the user free to decide the kind of content they want their story to include. Motif’s visual language suggests these constraints through the use of puzzle pieces (e.g., [14]).

The elements required for each pattern — required and optional shots, narration, and title text — are transformed into *prompts* (Figure 2). A user can tap on each prompt to insert corresponding video, audio, or text into their story. Additionally, users can use these prompts as guides during capture mode (Figure 3).

Scenario

Storytelling patterns provide guidance for the content users need to capture to achieve specific narrative goals. Reciprocally, they prioritize recognition over recall — rather than trying to remember what they’ve seen in past examples, novice users can now choose storytelling elements from a provided library. In this way, novices capture content they might not otherwise have captured. Below, we walk through a scenario illustrating how a user constructs a story using Motif.

Story construction through blocks

Jin is heading out for a short vacation to New York City with his friend Clara. He wants to use Motif to create a short video story about his trip so that he can share it with friends and family, as well as have a memento of the trip for the future. Having planned out his travel itinerary, Jin already has ideas for the places he wants to include in his story.

The night before he leaves, he opens Motif on his Android phone and creates a new story with a skeleton he can fill in while travelling. To start out, Jin selects a suitable Beginning block for his story. He adds the *Setting Out* story block to his story, which represents a pattern made up two elements: (a) A shot of scenery passing by as you travel to the event. (b) Narrate: where are you going, and with who?

Jin decides to add a few more story blocks for significant stops during his trip. He wants to make sure that he cap-

tures his visits to Magnolia Bakery, the Empire State Building, and Times Square, so he adds the *The Next Stop: Intro* and *The Next Stop: Describe* story blocks for each of those stops. Additionally, he adds the *The Next Stop: Mini Review* story block for the section in his story about the bakery, anticipating that his friends may want to hear his thoughts about the food and ambience he finds there. Jin continues adding a few more story blocks in this way before heading to bed. In this way, Jin is able to construct his story before any clips have been captured, in contrast to constructing a video story on a frame-by-frame basis after the trip is over.

Prompts encourage users to capture

The following day, Jin heads to the airport with Clara. While waiting at the gate, Jin checks Motif and realizes he forgot to capture the first shot for the *Setting Out* block, which suggested capturing “a shot of scenery passing by as you travel to the event”. Jin decides to set the scene for travel in a different way; rather than capturing scenery passing by, he decides to capture an ambient shot of the airport near the gate.

To do this, he taps the empty circle next to the prompt, which flips the application into camera mode (Figure 3). Jin captures a 3-second shot (as suggested by the prompt), and the empty circle fills in with a thumbnail of the video clip to indicate that Jin has successfully fulfilled the prompt (Figure 2a). Motif acts as a visual checklist for Jin as he fills in his story.

Then, he looks at the next shot in the story block, which suggests “Narrate: where are you going, and with who?”. Again, using the camera, he records himself briefly describing the context of his trip, and even asks his friend Clara for a few words so that she is in the video as well. When Jin is done, he quickly scans over the rest of the list to see if there is anything else he can record. He sees the rest of the story blocks can only be fulfilled at future stops on the itinerary, so he pockets his phone for now.

Jin later approaches the first major stop on his trip — the Empire State Building. He remembers having created some story blocks for this part of the trip and takes a look at Motif before he and Clara arrive there. By the time they approach the building, Jin is ready to record shots introducing and describing it. Along the way, he spots a street musician he finds interesting. Though it was not part of his plan for capture, he is able to click the “Add New Shot” button to record a few seconds of this unexpected experience on the fly. Jin can later move this uncategorized clip (Figure 4) to augment an existing story block in his story, or use the clip to fill in an empty prompt. In other words, story blocks are not rigid but can be modified to suit various situations and ideas.

Motif generates the story

In Motif, Jin is able to preview his video story at any time. When he returns to his hotel room at the end of the day, he presses the Play button to see how his video story is progressing. He notices a few things that he would like to change; he removes a shot he recorded on accident and belatedly records a clip of him picking up his backpack to remedy a jarring transition between when he left the hotel room in the morning and his visit to the Empire State Building. Once he has

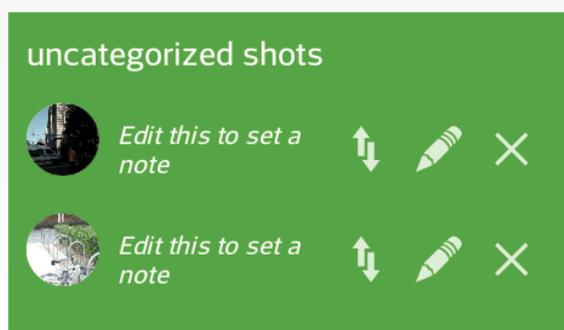


Figure 4. Users can also record shots that are not tied to a story block. These shots can be organized into story blocks later.

made modifications, Jin previews his story once again, this time finding it a much more smooth experience.

EVALUATION

Motif hypothesizes that a storytelling pattern language can support users in telling well-structured stories. In this section, we report on an evaluation exploring whether this overall strategy improves creative outcomes.

Method

We evaluated Motif through a controlled study comparing two versions of the application: the full Motif application (the Motif condition), and a version without story blocks or prompts (the control condition). The control version of the application acted like a standard camera and gallery application, with the additional ability for users to type brief notes for each captured video clip if they wished.

Thirteen participants (eight female, five male) were recruited from TaskRabbit and Craigslist in Seattle and Palo Alto for a task that lasted 1.5 hours. Six participants were recruited at the Seattle location and seven participants were recruited at the Palo Alto location.

Participants were randomly assigned to one of the study conditions and asked to use the application to create a video about a highly popular tourist spot in the area (Pike Place Market for Seattle, Stanford University for Palo Alto). The participants were first given a list of points of interest for the area as well as a map in order to ensure all participants would have some basic knowledge about available subject matter. They were then given 10–15 minutes to create a plan for what to capture using the application. Then, participants walked around the area to collect video footage according to their plans. A researcher followed participants to answer any questions about the interface and to take notes as the participant thought aloud about their video recording process. We then conducted a semi-structured interview with participants about their experience, focusing on what their criteria was for including certain shots, why they made certain design decisions, and the easiest and hardest things about creating their video. Lastly, participants filled out a short survey evaluating their experience and the video story they had created. Participants were paid \$40 for their time.

To evaluate each video for quality, two researchers and one video expert used a rating rubric to independently rate each video, blind to condition, on a 7-point likert scale based on five elements:

- **Structure** ($\kappa = .867, p < .05$). Is there a clear beginning, middle, and end? Are the beginning and endings generic, or are they strongly related to the theme of the story?
- **Shot Coverage** ($\kappa = .626, p < .05$). Is there more than one shot per scene in the story? If so, do the shots cover multiple points of view and angles?
- **Shot Composition** ($\kappa = .745, p < .05$). Is the subject of the shot each clear? The shots should avoid placing subjects in the center of the shot and consider the rule of thirds. Individual shots should be as still as possible.
- **Shot Length** ($\kappa = .775, p < .05$). Are shots long enough to understand (> 2 seconds) but short enough so that it stays interesting (< 10 seconds)?
- **Audio Content** ($\kappa = .593, p < .05$). Does the video contain a smooth audio track that links the whole story together? Is narration (if any) easy to hear and understand? Is narration informative and specific?

For final decisions, disagreements in ratings were resolved through discussion. The rating rubric used was based on guidelines from a popular beginner’s book about creating home video [15].

Lastly, we coded and examined survey and interview responses to look for themes in behavior between the two study conditions.

Results

Seven videos were created using the control application and six videos were created using Motif. Participants were between the ages of 22 and 60, with the average age being 36 ($SD = 12.75$). The videos created by the participants were an average of 4.8 minutes long, though total story length varied widely ($SD = 3.9$ minutes). The median video length was 3.3 minutes.

Most participants stated that they were not “video people”, explaining that when travelling or sightseeing they either focused on the experience at hand, or took photos using their phones. Most participants also stated they had little experience with video editing applications, with five participants stating they had never completed a video editing project before, and six participants stating they had completed fewer than three projects in the past.

Because we gave participants a list of suggested places to visit during the study task, participants tended to visit the same areas while creating their video story. However, ideas for video stories were diverse, ranging from “a food-focused tour of local businesses” to “the story of buying flowers for my partner” to “the ghosts of Pike Place”. Participants (especially those from the Seattle location) took the study as an opportunity to purchase items or pass by spots they personally found interesting about the area but had not yet had a chance to see.

In this sense, participants acted as realistic tourists; they were not experts in the area but had some knowledge of what places were famous as tourist attractions and what places were personally interesting to them. Participants also had to deal with the challenge of navigating an unfamiliar area while carrying items and recording video footage.

Story blocks helped users utilize structure of examples

The median number of story blocks in Motif video stories was five, and the median number of shots in Motif video stories was 10.5. Participants tended to use travel-related story blocks, with the top used blocks being *The Next Stop: Describe* (used in five stories), *The Next Stop: Intro* (used in four stories), *Set the Location* (used in four stories), and *Reflections* (used in three stories). This is unsurprising given the nature of the study task; most videos created by participants followed a structure where the participant moved from one point of interest to the next. Participants used a mix of patterns extracted from both the professionally produced video examples and the home video examples; the Reflections pattern was extracted from one of the home video examples.

Participants in both conditions explained that they used expert videos they had seen in the past as mental examples for the video they created. However, for participants in the Motif condition, story blocks and prompts seemed to help translate these examples into a structure for their video stories. Participants — even those that reported using video quite often to document personal experiences — noticed that this changed the way they recorded their video experience:

I definitely recorded different things than I normally would have. I liked the structure, how it's like a checklist, that really helps. With just the food, for example, now we're walking up to the restaurant, okay now this is the restaurant. I liked those, and normally if it were on my own, I would have done just one video [clip] and that would be it.

Participant 8, Motif condition

Participants in the control condition were less sure about how to apply elements from expert examples from the past in their own story. This was not for lack of knowledge about the area; many participants in the control condition told the researcher small stories relevant to places visited during the task when not recording (for example, describing their experience on fishing boats in Alaska while visiting the various fish stalls at Pike Place Market). Through they were given the same amount of time as Motif participants to plan their video, control participants tended to spend little time doing so, instead jumping straight into shooting whatever caught their eye. As a result, they developed a sense of a potentially good structure for their video story only after the task was finished:

Maybe I would put more historic facts, to tell you what these [statues] were. Or maybe I would start in the center [of the garden] and move out... I didn't have a plan... Starting was awkward and finishing was awkward.

Participant 10, control condition

At the same time, the prompts suggested by Motif did not perfectly support every participant's ideas. However, participants did modify existing story blocks to suit their own storytelling needs or recorded slightly different shots than the ones suggested by blocks:

I kinda had to say like, "Yeah, I guess that fits." It wasn't like, "Yeah, oh, that sounds perfect." It was like, I can kinda fit it in... I had my ideas of what I wanted and I was looking for that and I couldn't find it, so I just found the next best thing.

Participant 8, Motif condition

A Kruskal-Wallis test on experts' ratings revealed that videos created with Motif were significantly more structured according both to raters ($\chi^2(1) = 4.803, p < 0.05$) and to the expert evaluator ($\chi^2(1) = 5.09, p < 0.05$); that is, Motif videos tended not to start or end suddenly, instead having clear beginning and ending sequences. In other words, Motif successfully scaffolded the ideas participants had for their stories, guiding participants in creating plans for capture but allowing modifications to flexibly support a variety of story ideas. Participants found themselves capturing shots necessary for a larger narrative, stating they would not have captured these shots without guidance from Motif.

Story blocks reduced cognitive load during capture

Story blocks allowed users to pay an upfront cost for later benefits. Finding the right story blocks to include in the outline for their story was difficult for some participants; however, developing the story at a high-level prior to capture lessened the cognitive load for participants during the actual capture task:

The hardest thing was probably finding the templates, trying to fit my idea... the easiest was organizing the [video clips]. It was so nice to be like, I know that this is going go here, or I know that this is going to be the intro video. And then I knew I could record it later... I didn't have to everything in order... the organizing is done for me when I'm choosing my templates. Then I'm just filling those in.

Participant 8, Motif condition

In addition, participants did not have to generate the list of what they wanted to capture from scratch; they simply had to reconcile existing patterns with their own ideas.

The suggestions themselves were all well-worded, and they actually helped provide a guide for the goal that I had selected. It was like following a "choose-your-own-adventure" book... you have the free will to do a bit more of where you want to go, but eventually you're going to get to your conclusion, and it's helpful along the way to be pushed along, but gently... it was kind of a reminder, like, "What am I supposed to be doing here? Oh, right."

Participant 5, Motif condition

Participants made use of these reminders for what to capture in slightly different ways. Most participants, as described above, filled in story blocks according to the prompts given

by Motif. However, one participant used story blocks as a higher-level method of keeping track of the purpose of each of their clips. Rather than creating a *The Next Stop: Describe* story block for each stop they visited, they used a single *The Next Stop: Describe* to “store” all clips they recorded where they described a new location in their tour.

Control participants, in contrast, had to mentally juggle their story idea with judgments about what to capture, making it hard to separate how they experienced the task in real-time from the order of events they wanted their story to eventually depict. This resulted in a collection of shots that were often aimless and unrelated. This became clear in one participant’s use of notes in the control condition – he considered writing notes annoying but necessary to help him remember what he was capturing, as there was no structure to ground the purpose of shots:

I wanted to label stuff so... if I ever wanted to go back and look at things then I would, the labels, I guess they’re kind of like tags that help me remember what I took videos of. If I wanted to rearrange them in the future. I guess I did it on the way because I... doing it after the fact takes more time because you have to rewatch the video.

Participant 13, control condition

Similarly, Motif participants had the most success creating structure in their stories when using story blocks that did not require them to think about a story timeline different from the way they experienced the event. As an example, one participant used the *Set the Location* and *Highlight Reel* story blocks at the beginning and end of his story, both of which call for clips showing a preview (or a review) of the spots visited during the video. Therefore, these blocks must be completed in non-linear fashion. Some of the participant expressed confusion when encountering these prompts in their story, and ignored them. Other participants explicitly removed these particular prompts.

Prompts helped with shot coverage, but not composition

A Kruskal-Wallis test observed a marginal increase in shot coverage (that is, whether participants recorded multiple shots with different points of view for a scene) for participants in the Motif condition according to raters ($\chi^2(1) = 3.6708, p = 0.056$) and the expert evaluator ($\chi^2(1) = 4.33, p < 0.05$), indicating that patterns may help novices at least think about breaking down scenes into smaller, more consumable pieces.

However, there was no difference in how well shots were composed and framed (raters: $\chi^2(1) = 0.915, n.s.$; expert: $\chi^2(1) = 1.6878, n.s.$). Videos from both conditions generally lacked still shots, with most participants capturing scenes using panoramic shots or while walking.

While neither version of the study application limited how long participants could record for each clip, the Motif application suggested along with prompts how long each clip should ideally be based on expert patterns. We expected Motif videos to comprise of many short clips, but there was

no observed effect of study condition on the average length of shots per story (raters: $\chi^2(1) = 0.8767, n.s.$; expert: $\chi^2(1) = 2.4231, n.s.$). This may be due to the fact that participants from both conditions tended to narrate their thoughts in almost every shot they recorded, even for prompts that did not ask for narration. Evaluators did not see a difference in the quality of narration between conditions (raters: $\chi^2(1) = 0.8891, n.s.$, expert: $\chi^2(1) = 3.2124, n.s.$).

Motif videos ranked higher in quality overall

The expert evaluator placed participant videos in an overall ranking according to narrative and videographic quality. According to a Mann-Whitney U test, videos created by participants in the Motif condition received significantly higher ranks than videos created by participants in the control condition ($U = 38, p < 0.05$).

DISCUSSION

Through observations of participants creating video stories with and without Motif, we found that patterns extracted from expert work are able to successfully support novices in making creative decisions such as deciding on a narrative structure and making judgments about the kinds of content to capture. Further, we found that Motif videos were rated as higher quality than the control videos. In this section, we discuss the strengths and limitations of the strategy based on videos created by participants.

Supporting the what and the how

Because we generated patterns in terms of story function rather than visual aesthetic, the patterns we extracted from expert work were largely structure-oriented. That is, patterns provided templates for the kinds of shots to capture and how these shots are grouped. It is unsurprising that the areas in which Motif supported novices was in creating story structure and a unified theme in video stories.

Patterns provided by Motif did not seem to increase novice ability in *composing* the types of shots to capture; videos from both study conditions contained similar violations of videographic principles. For example, most professional videos convey a coherent picture of a subject through a series of several short shots that hold still and let subjects move in and out of the frame. However, in this study, participants employed amateur techniques such as capturing a location through 360-degree panoramas and moving the camera through a space for a long, continuous amount of time. In other words, participants knew *what* to capture, but reverted to natural habits with respect to *how* to capture.

We may be able to use the same strategy of extracting patterns from expert work to make examples of a different characteristic of video stories (such as shot composition) accessible to novices. For example, populating story blocks with visual examples rather than just text prompts may provide a more suitable scaffold for both story structure *and* creative execution. Another approach might be to apply basic machine learning techniques to detect common shot missteps (e.g., “Try holding your camera still while you take multiple shots!”) and provide users with tips for improvement.

Being mindful of discovery and emotion

Though we provided participants with maps and suggested points of interest in the area about which they were making a video story, some participants were less ready to create content about the area than others. Participants found themselves sometimes unsure of how to describe a place or thing, leading some to suggest that Motif suggest not just the types of clips to capture but also the kinds of information that might be interesting to a viewer. As one participant put it, this was a clear gap in the approach Motif used (structuring novice creative work around expert patterns) and the way most people experience new places during travel or events:

The process of creating a video is forcing you to think about creating a video, you know, you have to think about what does the audience want to see, what's the story, what's the explanation... it's quite a different experience of being a person that's just going to somewhere for the first time and seeing it in person – you don't know what you're going to see and you're discovering things.

Participant 6, Motif condition

Instead, we may be able to use expert patterns to structure “micro-moments” in addition to the overall story. Currently, Motif depends on the participant either having some time beforehand to plan out the story blocks they anticipated using in a story, or being familiar with the types of shots Motif would ask them to capture. However, travel often includes unexpected events, changes in plans, and serendipitous meetings. Beyond just utilizing patterns evident in the end result of expert work, Motif could also embed some of the strategies experts use to anticipate and ready themselves for capturing such events.

For example, the user could indicate to the system that a certain type of event is currently happening or that they have a certain idea they want to convey. Motif could then make suggestions about how to prepare for this idea (e.g. “Is there an informational plaque nearby to help you figure what to say to viewers?”) or how to prioritize capture (e.g. “Make sure you capture the audience’s reaction before the performance ends!”). After the moment of discovery is over, Motif could then use its knowledge of what patterns work well together to help the user situate their captured moment in the right place within the larger story being developed.

The novice’s audience

Our paper focuses on evaluation criteria surrounding narrative structure and videographic quality. However, it is likely that the social aspect of novices’ storytelling goals are significantly different from that of experts. Amateurs may place significantly more weight on creating stories that are shareable with family and friends, while placing less weight on creating a video with high-quality videography that may appeal to a larger audience.

We did ask participants about shareability of the videos they created, satisfaction with the video creation process, and satisfaction with the end result through surveys and interviews, but saw no significant difference in responses between the two conditions. This makes sense: we designed a study with

a realistic video capture environment, but participants were ultimately paid to make the video. As a result, participants’ motivations were artificial – they created a video in a content-oriented way (“create a video about a certain place”) rather than because it was an event or location they truly wanted to remember or share. While this paper focuses on aiding a novice in a certain creative task, we do note that it is important to consider how an anticipated audience affects the creative process and evaluation criteria for a novice storyteller, and leave this as future work.

Limitations

Common video editing tasks such as trimming video clips were not supported in either study condition. For similar technical reasons, Motif participants did not have access to the feature of automatically arranging their content using the timeline arrangements we generated for each pattern (Table 1). Though this resulted in video stories that were less concise than if one were to use existing video editing software, we were interested less in the ability of novices to edit video and more in how patterns might play a role in helping novices make creative decisions.

The patterns we generated in this paper are not meant to comprehensively represent all strategies used by videographers; rather, we wanted to demonstrate that extracting patterns from expert work is possible and that these patterns can be made accessible to novices and support their creative work.

CONCLUSION AND FUTURE WORK

In this paper, we sought to tightly weave together *capture* and *construction* for storytelling novices. Toward this end, we identified 24 storytelling patterns such as *Setting Out*, *Key Moment* and *Travel Break* through an analysis of expert examples. Each pattern includes a set of constituent elements as well as a video timeline describing how these elements are arranged. We embedded these patterns into Motif, a mobile application that scaffolds novice creative work using story blocks and prompts. In a field experiment, we found that Motif videos were significantly better than videos created by participants without access to patterns in terms of narrative structure and overall videographic quality.

Stories are inherently social; a storyteller often develops their understanding of a story as they tell it to more and more people. Suggesting captured pictures and video that might be appropriate to share next improves narrative engagement [9]. It would be interesting to see if patterns could play a role in facilitating joint meaning-making through story with groups of people. In the domain of video production, one can imagine a group of friends acting as an informal film crew, with a system dividing creative responsibilities across each person in the group. For example, one person could be assigned to capture ambient sound in each place, another person could be assigned to capture interviews, and another person could be in charge of managing the overall story structure. Expert patterns, if broken down by these responsibilities, may be able to guide groups of users as well as individuals.

With Motif we demonstrated how pattern languages can be used in the design of end-user creative tools. While we ex-

plored supporting novices in creating short video stories in this paper, this approach could apply to other domains as well. Imagine being able to build an essay out of patterns seen in the New York Times or compose a song by mixing and matching musical strategies used by your favorite artists. Examples of creative work made by experts are everywhere; by breaking down these examples into accessible patterns, we can open up creative opportunities for all.

ACKNOWLEDGMENTS

We would like to thank our study participants for their time and valuable feedback, as well as Sebastien Robaszkiewicz and Michael Rubin for their expert perspectives on video creation. Thanks to our colleagues that participated in pilot studies and tested the Motif app during development. This material is based upon work supported by Adobe Research and the National Science Foundation Graduate Research Fellowship under Grant No. DGE-114747.

REFERENCES

1. Animoto. <http://animoto.com/>.
2. Explory. <http://www.explory.com/>.
3. iMovie. <https://www.apple.com/mac/imovie/>.
4. Memoir. <http://www.yourmemoir.com/>.
5. Timehop. <http://timehop.com/>.
6. Alexander, C., Ishikawa, S., and Silverstein, M. Pattern languages. *Center for Environmental Structure 2* (1977).
7. Balabanović, M., Chu, L. L., and Wolff, G. J. Storytelling with digital photographs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '00, ACM (New York, NY, USA, 2000), 564–571.
8. Barry, B. A. *Mindful Documentary*. PhD thesis, Cambridge, MA, USA, 2005. AAI0808503.
9. Chi, P.-Y., and Lieberman, H. Intelligent assistance for conversational storytelling using story patterns. In *Proceedings of the 16th International Conference on Intelligent User Interfaces*, IUI '11, ACM (New York, NY, USA, 2011), 217–226.
10. Duyne, D. K. V., Landay, J., and Hong, J. I. *The Design of Sites: Patterns, Principles, and Processes for Crafting a Customer-Centered Web Experience*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2002.
11. Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., Smyth, G., Kapur, N., and Wood, K. Sensecam: A retrospective memory aid. In *Proceedings of the 8th International Conference on Ubiquitous Computing*, UbiComp'06, Springer-Verlag (Berlin, Heidelberg, 2006), 177–193.
12. Lindley, S. E., Durrant, A. C., Kirk, D. S., and Taylor, A. S. Collocated social practices surrounding photos. In *CHI'08 Extended Abstracts on Human Factors in Computing Systems*, ACM (2008), 3921–3924.
13. Petrelli, D., van den Hoven, E., and Whittaker, S. Making history: Intentional capture of future memories. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, ACM (New York, NY, USA, 2009), 1723–1732.
14. Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., and Kafai, Y. Scratch: Programming for all. *Commun. ACM* 52, 11 (Nov. 2009), 60–67.
15. Rubin, M. *The little digital video book*. Pearson Education, 2008.
16. Stokes, P. D. *Creativity from constraints: The psychology of breakthrough*. Springer Publishing Company, 2005.
17. Thiry, E., Lindley, S., Banks, R., and Regan, T. Authoring personal histories: Exploring the timeline as a framework for meaning making. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, ACM (New York, NY, USA, 2013), 1619–1628.
18. Van House, N., Davis, M., Takhteyev, Y., Good, N., Wilhelm, A., and Finn, M. From what? to why?: the social uses of personal photos. In *Proc. of CSCW 2004*, Citeseer (2004).
19. Wines-Reed, J., and Wines, J. *Scrapbooking for Dummies*. John Wiley & Sons, 2011.