
Generating Narratives from Personal Digital Data: Triptychs

Matthew P. Aylett
University of Edinburgh
School of Informatics
10 Crichton Street
Edinburgh, EH8 9AB, UK
matthewa@inf.ed.ac.uk

Thomas Dickinson
The Open University
Knowledge Media Institute
Walton Hall
Milton Keynes, MK7 6AA, UK
thomas.dickinson@open.ac.uk

Elaine Farrow
University of Edinburgh
School of Informatics
10 Crichton Street
Edinburgh, EH8 9AB, UK
Elaine.Farrow@ed.ac.uk

Larissa Pschetz
University of Edinburgh
Edinburgh College of Art
Evolution House, 78 Westport
Edinburgh, EH1 2LE, UK
larissa@entretags.de

Abstract

The need for users to make sense of their growing mass of personal digital data presents a challenge to Design and HCI researchers. There is a growing interest in using narrative techniques to support the interpretation and understanding of such data. In this early study we explore methods of selecting images from personal Instagram accounts in the form of a *triptych* (a sequence of three images) in order to create a sense of narrative. We present a brief description of the algorithms behind image selection, evaluate how effective they are in creating a sense of narrative, and discuss the wider implications of our work. Results show that semantic tagging, a dynamic programming algorithm, and a simple narrative structure produced triptychs which were significantly more story-like, with a significantly more coherent order, than a random selection, or a neutral sequence of images.

Author Keywords

social media, narrative, triptych, multi-media

ACM Classification Keywords

H.5.1 [Multimedia Information Systems]: Evaluation/Methodology.

Permission to make digital or hard copies of part or all 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 third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).
CHI'15 Extended Abstracts, April 18–23, 2015, Seoul, Republic of Korea.
ACM 978-1-4503-3146-3/15/04.
<http://dx.doi.org/10.1145/2702613.2732702>

Introduction

Industrialized societies are immersed in a sea of digital data, and a common challenge for HCI and Design researchers is to help users make sense of it. A typical approach treats it as database to be queried, allowing users to perform selection and summarization tasks such as finding all emails with the phrase 'flight reservation', retrieving date and time information from a friend's baby pictures, or discovering how many Twitter users chose the hash tag #justinbieber in June 2014.

However, we often interpret our own experiences, desires, and motivations using stories not statistics, and the use of narrative to make sense of our everyday lives is considered a fundamental human behavior [7]. The exploration of narrative formats for data presentation can help Designers and HCI researchers understand how meaning is constructed through stories and how stories can be used to interpret data.

The work presented here is part of the ReelLives project (EPSRC Reference EP/L004062/1) which attempts to build an automatic narrative generation system for personal digital data in a film-like medium. The project's aim is to explore the human response to such systems, to develop tools for editing and interacting with the output, and by doing so to empower users, demonstrating how their personal digital data may be used by third parties and returning control over how they present themselves.

We investigate methods for selecting sequences of three images. We present a brief description of the algorithms behind image selection and evaluate how effective they are in creating a sense of narrative for experimental participants who have never seen the images before. This is a first step in validating our approach before we carry out further work to investigate the effect of presenting

users **own** content to themselves and is a validation of the technique, rather than a user experience study. We conclude by discussing the wider implications of our work.

Background

Social Media and Narrative

Our initial study is limited to a structural interpretation of narrative. This approach (and its study, narratology), considers narrative to be reducible to different components that can be analyzed and modeled. A useful introductory text is Bal [1] which lists the elements of the underlying story (or fabula) as events, actors, times, and locations. Events are defined as "the transition of one state to another experienced by actors", distinguishing mere chronological sequence from a true narrative *event*, which must also contain a change.

There have been a few popular examples of automatic narrative generation from personal social media data. The Intel *Museum of Me* (intel.com/museumofme) aggregates individuals' public Facebook data into a virtual exhibition. Facebook's own *A Look Back* (facebook.com/lookback) application compiles users' most popular posts into a narrative that aims to represent their life events. In both these examples, however, the means used to detect and model events is proprietary. Furthermore, as noted in previous research, e.g. [8, 10, 9], the result is very much a finished product. While *A Look Back* allows users to replace chosen posts with others from a limited selection, users have very little control over the resulting film, and cannot choose different events, reorder material, or control the presentation style. Therefore, in order to produce an open system that can support users to interpret their data, we explore the process of building basic primitive narrative units using semantic analysis and dynamic programming.



Figure 1: The triptych form arose in early Christian art and was popular for altar paintings from the middle ages onwards.

Triptychs

In order to pursue our long-term goal of producing film-like narrative, we limit ourselves to visual presentation and concentrate on a simple method of communicating narrative using images: the *triptych*. A triptych is a sequence of three pictures, traditionally panel paintings, where the two side panels fold in to cover the middle painting to allow for easy transportation (Figure 1). The three pictures allow the artist to explore a sequential or thematic narrative.

Three is also considered to be the minimum number of components needed to construct an argument (thesis, antithesis, synthesis) and events to construct a story (beginning, middle, end) [5], represented in the classic three-act structure (setup, confrontation, resolution). Modern artists have explored triptychs, most famously Francis Bacon, and the form has become popular for wedding and graduation memorabilia.

ReelOut

The study presented in this paper is a first step in constructing ReelOut, a system that takes as input digital data preprocessed into a specific format, and automatically outputs a sequence of images. In the final system, the software will use presentation techniques appropriate to the data to generate a film-like output from the images. It will also allow users to edit the result to suit their own vision of the narratives they see within their lives.

The core system takes a set of RLUnits, consisting of marked-up text and semantic tags, and uses dynamic programming to fit these to a narrative structure. See Figure 2 for an example RLUnit. The sentiment of each unit is calculated using Sentistrength (sentistrength.wlv.ac.uk), a popular sentiment analysis

```
<rlunit id="10003" src="FB"
  time="20140420:064319">
  <image>image URL here</image>
  <messageText>
    #Easter
    #sunrise from Calton Hill.
  </messageText>
  <sentiment>neutral</sentiment>
  <location>Calton Hill</location>
  <theme>Easter</theme>
  <theme>sunrise</theme>
  <nextunit>10004</nextunit>
  <prevunit>10002</prevunit>
</rlunit>
```

Figure 2: An example RLUnit.

tool for short web texts. Hash tags found in the text are used as themes, and locations are identified using Apache OpenNLP (opennlp.apache.org). The narrative target is represented as an ordered collection of slots with associated semantic tags which constrain their contents.

Units are selected using a Viterbi algorithm, where two cost functions are optimized to produce an output. The first, *join cost*, represents how connected two adjacent units are: units which share similar semantic tags and appear in the correct temporal order will have a low join cost. The second, *target cost*, represents the fit with the narrative target: units which share semantic tags with the target slot will have a low target cost.

We produced three types of triptychs for evaluation:

1. **Random** We selected three random units (no narrative target, no Viterbi search). The only constraint was that units were not repeated.
2. **Neutral Narrative** The narrative target consisted of three slots, each with a neutral sentiment.
3. **Emotional Narrative** A narrative target with a positive trajectory: the first slot had a negative sentiment and the third had a positive sentiment, with the middle slot left unconstrained.

Evaluation

We used the Websta Instagram Web Viewer (websta.me) to identify Instagram profiles using the “life” keyword and not marked as private. 26 profiles, each with more than 30 posts containing both an image and text, were used to generate our three types of triptychs, resulting in 78 stimuli in total. 20 participants, none of which had produced or seen the pictures before, were recruited, all native or near-native speakers of English, with normal vision.

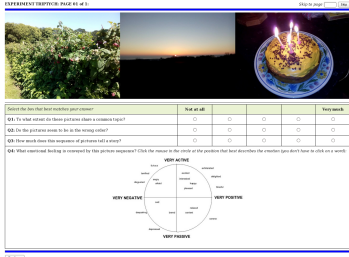


Figure 3: Thumbnail of experiment web interface (original Instagram images replaced with similar images to protect privacy)

<p>Q1: To what extent do these pictures share a common topic?</p> <p>Q2: Do the pictures seem to be in the wrong order?</p> <p>Q3: How much does this sequence of pictures tell a story?</p> <p>Q4: What emotional feeling is conveyed by this picture sequence?</p>
--

Table 1: Narrative evaluation questions

Participants were asked four questions about each triptych (Table 1). The stimuli were presented via a web interface (Figure 3). The first three questions asked for a rating on a 5-point scale from “Not at all” to “Very much”, with the options presented as radio buttons.

The fourth question required the participant to click a location in a two-dimensional emotional continuum based on [6](Figure 4). This approach to evaluating emotional responses regards emotions as varying in *activation*, from passive to active, and *evaluation* from negative to positive. To help participants interpret these two dimensions the space is labeled with emotions such as ‘bored’, ‘angry’, ‘pleased’. A 2D space is an over-simplification of human emotions and cannot take into account richer views of emotions such as OCC [4], or those explored in [2]. However, it is a useful pragmatic device for assessing emotional *differences* between stimuli.

Narrative is often characterized as being thematically coherent and expressing a sequence of events (Q1 and Q2). In addition we ask the participant directly if the pictures seem to tell a story (Q3). To discover whether the triptych encourages an emotional response, another common feature of narratives, we ask Q4.

Our hypotheses were as follows:

- H1:** Participants could discern narrative differences between the three types of triptychs.
- H2:** The ReelOut algorithm created a deeper sense of narrative for **Neutral Narrative** and **Emotional Narrative** conditions compared to **Random**.
- H3:** The ReelOut algorithm created the deepest sense of narrative for **Emotional Narrative**.

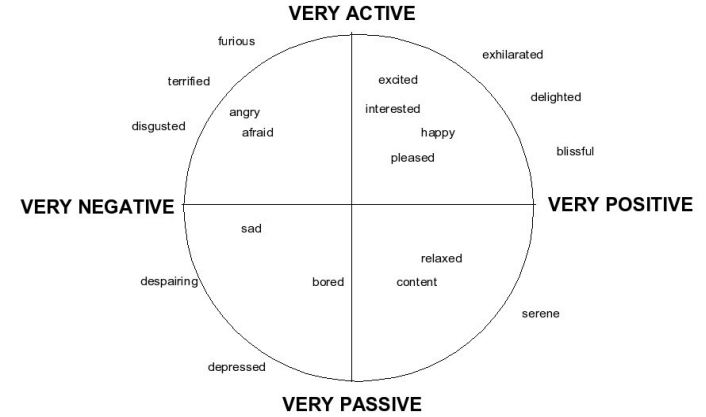


Figure 4: Closeup of 2D emotional evaluation/activation space

The data was analyzed using a repeated measures MANOVA, with stimulus type as a within-subjects factor where the values are the average of the subjects’ responses to each stimulus type for each question. Parametric analysis of interval data is valid on averages of six or more results [3]. In order to avoid a type I error, a traditional approach was taken in requiring significant results in the MANOVA to legitimize a repeated measures subject ANOVA, where in turn a significant result was required to make legitimate the post-hoc paired t-test analysis.

Results

A MANOVA was carried out with 5 dependent variables (a by-subject average by condition, separating the two dimensions in the emotional continuum, Evaluation and Activation), and stimulus type as the within-subjects factor. Results were highly significant (Wilk’s lambda 0.22, $F(10, 10) = 44.542$, $p < 0.001$).

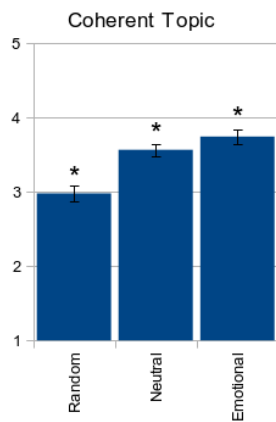


Figure 5: Coherent Topic

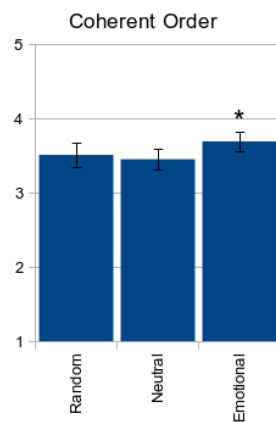


Figure 6: Coherent Order

Sphericity held for all dependent variables except Evaluation (Mauchly's Test of Sphericity, $p < 0.05$). Subsequent by-subjects ANOVA with stimulus type as the within-subjects factor were significant for all 5 dependent variables (results shown for sphericity assumed except Evaluation where a Greenhouse-Geisser correction is applied). Q1 (Coherent Topic) ($F(2, 38) = 113.854$, $p < 0.001$), Q2 (Coherent Order) ($F(2, 38) = 7.002$, $p < 0.005$), Q3 (Story) ($F(2, 38) = 13.477$, $p < 0.001$), Q4a (Evaluation) ($F(1.47, 27.12) = 15.112$, $p < 0.001$), Q4b (Activation) ($F(2, 38) = 10.289$, $p < 0.001$).

Post-hoc tests

Post-hoc paired t-tests with Bonferroni correction (3 multiplier) were carried out between the difference-means for each dependent variable. Results did not follow a simple pattern. For clarity we report t-test results for the most significant 3 way comparison only. For Coherent Topic, all three conditions were significantly different from each other, with **Emotional Narrative** rated the most topically coherent ($t(19) = 18.414$, $p < 0.001$, Figure 5). For Coherent Order, **Emotional Narrative** was rated significantly more coherent than the other conditions ($t(19) = 4.150$, $p < 0.005$, Figure 6). For Story, **Neutral Narrative** and **Emotional Narrative** were rated significantly higher than **Random** ($t(19) = 4.358$, $p < 0.001$, Figure 7). For Evaluation, **Neutral Narrative** and **Emotional Narrative** were rated as engendering a significantly more positive emotional response than **Random** ($t(19) = 4.710$, $p < 0.001$). In contrast, for Activation, **Neutral Narrative** was rated as producing a significantly more active emotional response than **Random** and **Emotional Narrative** ($t(19) = 4.892$, $p < 0.001$). Figure 8 shows means by condition for each subject on the 2D emotional continuum (zoomed to the area of interest). These results are interpreted in the following section.

Discussion

Data from Instagram public accounts was readily-available and allowed evaluation without compromising privacy. However, it was not rich in narrative events, and the associated text was of limited extent and complexity, leading to limited extraction of the semantic features which are used to search and select the triptychs. Nevertheless, the evaluation has shown that participants were very sensitive to narrative content and there were significant differences in participants' responses by stimulus type. We can therefore accept hypothesis **H1**, that participants were sensitive to the different selection criteria used by ReelOut.

Results for Coherent Topic and Story showed that stimuli produced by selecting temporally-ordered sequences with similar features (**Neutral Narrative**, **Emotional Narrative**) resulted in a greater sense of narrative. The sequenced conditions also engendered a significantly stronger positive emotional response than the random images. We note that the means for Story were all below 3, suggesting that the level of perceived story remained low. This may be due to asking participants to rate images which have no connection with them. However, overall we can accept hypothesis **H2**, that the selection algorithm does choose images which convey a greater sense of narrative than a random selection.

More exciting is that the results also allow us to accept hypothesis **H3**. By searching for an emotional change across the images, the sense of narrative is increased. This is supported by the result that **Emotional Narrative** scored significantly higher for Coherent Topic than both **Random**, and more critically, **Neutral Narrative**. In addition, **Emotional Narrative** was also regarded as producing pictures with a greater Coherent Order than **Neutral Narrative**, even though the requirement for

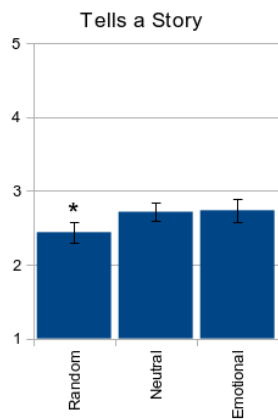


Figure 7: Story

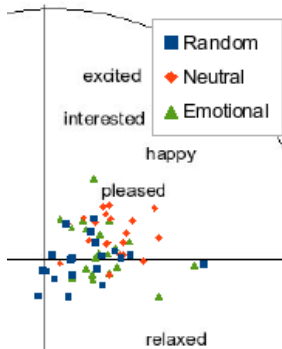


Figure 8: Emotion

emotional change means the selected images are less likely to be temporally sequential. This shows that the sense of narrative is influenced more by whether participants perceive a Coherent Order than the images having been taken one after the other in a similar context.

Conclusion and Future Work

The triptych form has allowed us to lay rigorous and testable groundwork for the ReelLives project. By limiting our media format in this way, we can test our search algorithm and explore the effect of narrative presentation of image data. However, the triptych is far removed from our final objective, a film-like medium. Although the emotion results are significant, Figure 8 indicates considerable variance between subjects. This wide variance is also present by materials. Thus although the method was useful in demonstrating differences between our conditions, further work developing a more consistent emotion metric is required.

In future work we will extend the system to richer sources of narrative data, gathered from Facebook and other social media. The ReelLives project has commissioned six experienced film-makers to create films using social media data. Techniques used in these films such as transitions, text, and music, will be adapted to enrich our output. Two key aspects of future work will be to: 1) present users' own content to them and explore the effect familiarity and personal meaning has on the perception of narrative, and 2) enable selection and curation of the generated narrative. A prototype editor which offers participants alternative images based on the underlying Viterbi search will be developed and evaluated.

The work presented here demonstrates the feasibility of automatically generating narrative from images. It throws

some light on mechanisms that a system can use to foster an increased sense of narrative. Such insights are not only relevant for the continuation of the ReelLives project, but also suggest ways HCI researchers can help users make sense of the sea of data in which they are immersed.

References

- [1] Bal, M. *Narratology: Introduction to the theory of narrative*. University of Toronto Press, 2009.
- [2] Boehner, K., DePaula, R., Dourish, P., and Sengers, P. How emotion is made and measured. *International Journal of Human-Computer Studies* 65, 4 (2007), 275–291.
- [3] Jr, H. B., and Boone, D. A. Analyzing likert data. *Journal of Extension* 50 (2012).
- [4] Ortony, A., Clore, G., and Collins, A. *The Cognitive Structure of Emotion*. CUP, Cambridge, 1988.
- [5] Prince, G. *A Grammar of Stories: An Introduction*. De Proprietatibus Litterarum. Series Minor. De Gruyter, 1974.
- [6] Russell, J. A. A circumplex model of affect. *Journal of personality and soc. psych.* 39, 6 (1980), 1161.
- [7] Sarbin, T. R. *Narrative psychology: The storied nature of human conduct*. Praeger Publishers/Greenwood Publishing Group, 1986.
- [8] Schwanda Sosik, V., Zhao, X., and Cosley, D. See friendship, sort of: How conversation and digital traces might support reflection on friendships. In *CSCW '12*, ACM (New York, 2012), 1145–1154.
- [9] Thiry, E., Lindley, S., Banks, R., and Regan, T. Authoring personal histories: Exploring the timeline as a framework for meaning making. In *CHI '13*, ACM (New York, NY, USA, 2013), 1619–1628.
- [10] Zhao, X., and Lindley, S. E. Curation through use: Understanding the personal value of social media. In *CHI '14*, ACM (New York, 2014), 2431–2440.