Procesamiento del Lenguaje Natural, Revista nº 44, marzo de 2010, pp 35-42 recibido 12-01-10 revisado 19-02-10 aceptado 05-03-10

An initial study on text summarisation in film stories

Un estudio inicial sobre el resumen de argumentos de películas

Yan Xu
University of Sunderland
DCET, DGIC, St. Peter's Campus, Sunderland
SR6 0DD, England
Yan.Xu-1@sunderland.ac.uk

Michael P. Oakes
University of Sunderland
DCET, DGIC, St. Peter's Campus, Sunderland
SR6 0DD, England
Michael.Oakes@sunderland.ac.uk

Resumen: El objetivo de nuestra investigación es el de generar resúmenes de películas a partir de textos colaterales, capturando el contenido semántico, estructura narrativa y líneas clave de los diálogos de la película. Nuestra hipótesis es que se pueden generar de forma eficiente resúmenes en texto de películas mediante el empleo de técnicas de resumen automático sobre textos colaterales: subtítulos, descripciones del audio y guiones de postproducción. En caso de disponer de códigos de tiempo, entonces podemos generar también resúmenes en vídeo a partir de dichos resúmenes en texto. En este estudio inicial construimos los resúmenes seleccionando las diez tomas de la película original que contienen la mayor proporción de palabras clave. Definimos las palabras clave de dos formas: como palabras de frecuencia media, ya que son las palabras de frecuencia media en un texto las que contienen la mayor parte de la información acerca del contenido de dicho texto; y como entidades nombradas derivadas del reparto de la película. Ye et al. (2007) sostienen que la calidad de un resumen puede evaluarse en base a cuántos de los conceptos principales del texto original se conservar en el resumen. Hemos comprobado que esta aproximación a la evaluación de resúmenes obtiene resultados más favorables que la aproximación ROUGE (Lin and Franz, 2004) basada en comparar el número de correspondencias de secuencias de caracteres entre los resúmenes generados automática y manualmente.
Palabras clave: Resumen automático de textos, argumentos de películas

Abstract: The objective of our research is to produce summaries of films from collateral text which capture the semantic content and narrative structure and key lines of dialogue in film. Our hypothesis is that text summaries of movies can efficiently be produced by the use of text summarisation techniques on collateral texts: subtitles, audio descriptions and post-production screenplays. If time codings are available, then video summaries can be constructed from these text summaries. In this initial study, we form summaries by selecting the ten shots from the original film which contain the highest proportion of keywords. We define keywords in two ways: as mid-frequency words, since it is the mid-frequency words in a text which contain most information about the content of that text, and as named entities derived from the cast list of the film. Ye et al. (2007) argued that the quality of a summary can be evaluated based on the how many concepts in the original text are preserved after summarisation. We found that this approach to summary evaluation gave more favourable results than the ROUGE approach (Lin and Franz, 2004) of comparing the number of matching character sequences in an automatic and a human-generated summary.
Keywords: Text summarisation, movie scripts

ISSN 1135-5948

1 Introduction

Through the mass media people are deluged with information and would greatly benefit from automatic systems which would provide very concise versions of information, relevant to them as individuals, relevant to their needs at the time of delivery. We are working to develop methods which will allow potential subscribers of film repositories effective access to the films which will most interest them. With a large volume of films to choose from, presenting the user with a summary of each one greatly speeds up the process of selection. For example, the Internet Movie Database (IMDb) (2009) is the biggest source for movie related content in the world, so they have an interest in ensuring that for every title that they have listed, they have a plot outline, summary or synopsis. Commercial applications would include recommender systems, where vendors of online DVD catalogues can help potential customers find the films they would like.

The field of video summarisation integrates techniques for language engineering, specifically text summarisation and corpus-based text analysis, video summarisation, and knowledge representation, to specifically deal with film characteristics. Video summarisation is especially difficult to achieve for story-oriented video. Current research focuses on the selection of key frames and shots to generate a summary, but such approaches do not summarise the plot of a film adequately (Smeaton, 2007).

Our approach is to use relatively simple text summarisation techniques on various kinds of collateral text, including film scripts/screenplays and audio description scripts. Audio description scripts are scripts described by trained audio describers to describe the visual features which visually impaired people would miss when they watch video (TV or films). An advantage of audio description is that it is time-coded, so the scripts are mapped onto film playing time. This directly links video features with text annotation.

1.1 Review of video summarisation techniques

A number of researchers have focused on content-based analysis of video, where events are defined by visual properties such as the colour, texture, shape of individual frames and differences between adjacent frames. Much work to date in event detection is either domain specific or related to news topics. For example, Hakeem and Shah (2007) used training data and clustering techniques to detect events in sports, railway monitoring and surveillance. Lin and Liang (2008)'s SToRe topic detection and tracking system gathers together news events under a single item, and composes a summary so that the readers can quickly appraise the evolution of events in that news item. Aner-Wolf and Kender (2004) detected scenes of sitcoms. With their mosaic-based representation, where an image is generated from a sequence of frames of a single shot, scenes are clustered into physical settings such as an apartment or coffee house. The chosen representative mosaics for each physical setting produced summaries and served as a basis for indexing video libraries. Money and Agius (2009) traced physiological responses such as respiratory rate, respiratory amplitude and heart rate in viewers of videos. Times of increased physiological response were taken to be the most entertaining segments of the videos, which were then incorporated into what they called affective summaries.

1.2 Why text retrieval is the key to video retrieval

Content-based multimedia information retrieval tends to be based on the units of shots and keyframes (Zhang et al., 1997). Unfortunately, shots are too small to bear enough content, while a combination of shots making up a scene describes the story element much better (Smeaton, 2007). Our project work focuses more on semantic events such as scenes. Our link between the low level features of video images and the high level semantic features of scenes, the so-called semantic gap, will be collateral text. Collateral text is any text accompanying a still or moving image, sound, or even another text: the important

point is that it is possible to extract information from a collateral text which cannot easily be extracted from original sources. Examples of collateral texts accompanying a film are subtitles, audio descriptions, or post-production screenplays.

1.3 Review of text summarisation approaches

We need to make the very most of existing conventional techniques, such as the use of mid-frequency words, title words and words in the first paragraph as candidates for extraction. Paice (1990) gives an excellent summary of these tried and trusted techniques. We also wish to explore original and innovative approaches to summarisation. In previous work by Teufel and Moens (2002), sentence features are first collated, and in a training phase, associations between those features and human-provided target categories are learnt. Other authors have brought together complementary work on probabilistic models such as Hidden Markov Models (HMMs) for Bayesian learning and template extraction, and stochastic parsing. Gervás et al. (2004) represented stories with case-based reasoning and a concept ontology. Aliguliyev (2009) uses a sentence similarity measure to cluster sentences, so that related sentences can be joined together in a briefer representation of the original text. Similarly, Antiqueira et al. (2009) built networks of sentences based on the number of related nouns in each sentence, then performed various mathematical characteristics of those networks to produce summaries.

Fattah and Ren (2008) compared a number of machine learning approaches to content selection for summaries, finding the Gaussian mixture model the most successful. Yeh et al. (2008) select the best sentences for a summary using a PageRank-like algorithm. To make more use of the context of sentences, Ko and Seo (2008) first consider sentence pairs for extraction, and then separate them into single sentences. Yeh et al. (2005) use a trainable summariser, trained on a corpus of text, which takes account of traditional extraction clues such as word position, cue and stigma words, topic centrality and degree of match with the title. Hidekazu et al. (2004) identified four sentence types (sales, specifications, structure and function) to summarise promotional literature about new products. Ko et al. (2004) use automatic summarisation as a precursor of text classification. The sentences appearing in the summary are regarded as more important than the others, and so the text features appearing in them are weighted more highly in the classification phase. Zhu and Chen (2007) identify key players, organisations and stories by applying co-occurrence analysis (specifically latent semantic analysis) of named entities and importance measures such as "degree centrality" and "betweenness centrality" – this overcomes the weaknesses of pure entity frequency counting.

2 Initial Analysis

In the approaches described here, we aim for fully automatic summarisation with minimal post-editing of the produced summaries. Our summaries are extractive, based only on the words which appear in the collateral texts. In this paper we produce summaries for single films in English, though many of our techniques will be applicable to films in other languages. We obtained high quality, post-production collateral texts, together with "gold standard" summaries written for specific purposes from IMDb.

These collateral texts were summarised using an established generic text summarisation heuristic described by Paice (1990) and Kupiek et al. (1995). This is to select the sentences (or in our case the film shots) containing the highest proportion of "keywords", which may be mid-frequency words, words from a gazetteer such the "dramatis personae" in the credits of a film, or words which appear in the title.

In our initial study, the screenplays of two movies, *Casablanca* (1942) and *The Abyss* (1989) were chosen. The screenplays are available online at http://www.weeklyscript.com/Casablanca.txt and http://www.dailyscript.com/scripts/abyss.html. According to the nature of the screenplay, they either have shot IDs or

camera motion terms (such as *fade in*, *cut to*) to identify fragments/shots of the story. First, camera motion terms and location description are excluded for each shot or the whole screenplay.

To identify the mid-frequency keywords, we first removed all the "stop" words appearing on a list of words which provide grammatical information, but no information about the content of a text, such as "the", "and" and "of". The 200 most frequent non-stop words were then taken to be the mid-frequency words which typically tell us most about the content of the text, such as *Rick*, *café*, *Casablanca*, *German*, *letters*, *America*, *transit* and *beautiful*. A second list of keywords was then derived from the list of characters' names. Only credited proper names or nicknames were used and first names and surnames are separated. For example, *Virgil 'Bud' Brigman* is a valid name and produced three terms: *virgil*, *bud* and *brigman,* while *USS Montana Executive Officer* is not a valid name. There were 18 and 35 keywords from characters' names for *Casablanca* and *The Abyss* respectively. A third list of keywords was formed by the union of the other two keyword lists.

Top200 keywords		Cast keywords		Combined keywords	
shots	score	shots	score	shots	score
34	0.727	57	0.364	57	0.773
57	0.727	51	0.324	34	0.727
49	0.700	38	0.316	49	0.700
17	0.697	55	0.304	17	0.697
22	0.696	18	0.299	22	0.696
55	0.696	64	0.294	55	0.696
45	0.692	24	0.292	45	0.692
46	0.692	56	0.291	46	0.692
65	0.692	19	0.286	65	0.692
62	0.687	36	0.286	62	0.687

Tabla 1: Different sets of keywords result in different rankings of the top 10 shots for the film *Casablanca*.

For each keyword list, the 10 shots which contained the highest proportion of keywords were extracted to produce a summary of the original film. The results for *Casablanca* are shown in Table 1, where in each case the score is the number of keywords found in the shot divided by the total number of words in that shot.

It can be seen that using cast keywords alone leads to lower scores, since there are relatively few cast keywords, whereas using the mid-frequency keywords and the combined keyword list lead to similar results. When using cast keywords alone, 8 out of 10 of the extracted shots are different when compared to the set of shots extracted using the other two lists of keywords, which are identical except for the different ranking of the top 2 shots. In the *Abyss*, 2 out of 10 shots are different when comparing the mid-frequency keywords with the combined set of keywords, but 7/10 and 8/10 are different comparing the mid-frequency or combined lists with cast keywords. An example of a shot in the automatic summary of *Casablanca*, derived from the combined list of keywords, is as follows. The capitalised names are followed by the named person's speech.

<shot=22>
INT. RICK'S PARIS APARTMENT – DAY
Ilsa fixes flowers at the window while Rick opens champagne. She walks over and joins him.
RICK Who are you really? And what were you before? What did you do and what did you think? Huh?
ILSA We said "no questions."
RICK Here's looking at you, kid.
They drink.
</shot=22>

The plot summaries created by IMDb users were considered as gold standard film summaries against which the machine summaries were compared. The chosen plot summary of *Casablanca* was:

In World War II Casablanca, Rick Blaine, exiled American and former freedom fighter, runs the most popular nightspot in town. The cynical lone wolf Blaine comes into the possession of two valuable letters of transit. When Nazi Major Strasser arrives in Casablanca, the sycophantic police Captain Renault does what he can to please him, including detaining Czech underground leader Victor Laszlo. Much to Rick's surprise, Lazslo arrives with Ilsa, Rick's one time love. Rick is very bitter towards Ilsa, who

ran out on him in Paris, but when he learns she had good reason to, they plan to run off together again using the letters of transit. Well, that was their original plan…

Two evaluation methods were used. One was ROUGE (Lin and Franz, 2004), which counts the number of adjacent character sequences which are found both in the automatic and the gold standard summaries. According to Dunlavya et al. (2007), ROUGE is the best known automatic metric for summarisation system evaluation. This is the metric used at the Document Understanding Conference (DUC) which has moved to the Text Analysis Conference (TAC). In this initial analysis, we used the BLEU and N-gram co-occurrence version of ROUGE. First, all punctuation and spaces between sentences are removed and replaced by one white space. The texts of both the automatic and gold standard summaries are segmented into 2, 3, and 4-grams of characters. C_{n-r} is the average n-gram coverage score in which co-occurring n-grams in both automated and gold standard summaries are divided by the number of n-grams in the gold standard summary, the "recall" metric. C_{n-p} is the average n-gram coverage score in which co-occurring n-grams are divided by the number of n-grams in the automatic summary, the "precision" metric. The n-gram co-occurrence statistics are based on the following formula:

$$Ngram_{(2,4)} = \exp\left(\frac{1}{3}\log C_2 + \frac{2}{3}\log C_3 + \log C_4\right)$$

So depending on whether the C_2, C_3, C_4 are used as recall or precision metrics, the results are different. The $Ngram_{(2,4)-r}$ results are higher than those for $Ngram_{(2,4)-p}$. The results for the film *Casablanca*'s automatic summaries produced by mid-frequency, cast and combined keywords respectively are: $Ngram_{(2,4)-p}$ =0.022, 0.001, 0.022 and $Ngram_{(2,4)-r}$ =0.367, 0.772, 0.367; whereas the results for the film *the Abyss* using mid-frequency, cast and combined keywords are: $Ngram_{(2,4)-p}$ = 0.009, 0.006, 0.010 and $Ngram_{(2,4)-r}$ = 0.350, 0.354, 0.384. We tested the formulas by comparing identical documents (the plot summary of *The Abyss* with itself) and random documents (the plot summary of *The Abyss* with an article about the South African runner Caster Semenya) . The formulas returned '1' for identical documents, but higher than '0' for random documents (0.03 and 0.19 for precision and recall respectively). The precision score for random documents is quite similar to the scores for summary comparison, but the recall scores for the automatic summaries were significantly better than random.

A second method of summary evaluation, based on Ye et al. (2007), also involved a comparison between a machine summary and a human-produced "gold standard". Here the quality of the summary is determined according how many concepts in the original text are preserved after summarisation.

For example, the IDMb summary of *Casablanca* was manually parsed into the following list of discrete facts, and the machine summaries were examined to see how many of these facts were retained.

1. Time is World War II.
2. Location is Casablanca.
3. Rick Blaine is exiled.
4. Rick Blaine is American.
5. Rick Blaine was a freedom fighter.
6. Rick runs the most popular nightspot in town.
7. Blaine possesses two valuable letters of transit.
8. Blaine is the cynical lone wolf.
9. Strasser is a Nazi Major.
10. (+)Strasser is in Casablanca.
11. (+)Renault is the sycophantic police Captain.
12. (+)Renault pleases Strasser.
13. (+?)Renault detained Laszlo.
14. Victor Laszlo is the Czech underground leader.
15. Rick is surprised.
16. Lazslo arrived with Ilsa.
17. (+?)Ilsa was Rick's one time love.

18. Rick is very bitter.
19. Ilsa ran out on him.
20. (+)They were in Paris.
21. Ilsa had a good reason to leave him.
22. Rick learns the reason.
23. Rick and Ilsa plan to run off together.
24. Rick and Ilsa want to use the letters of transit.

There were 24 such facts for *Casablanca* and 10 such facts for *the Abyss*. For the film *Casablanca*, we found 16 facts (3 of them doubtful) were found in the summary produced by cast keywords, but just 6 with the summaries produced by mid-frequency keywords or the combined list. The facts found in the combined list produced summary are marked with "+"(confident match) or "+?"(doubtful match). For film *the Abyss*, we found 3 facts (2 of which were doubtful), 3 facts (1 of which was doubtful), and 4 facts (of which 3 were doubtful) were found in the summaries produced by mid-frequency keywords, cast keywords and the combined list of keywords respectively. As an example of a confident match, fact 11 could be confirmed with the following dialogue in the automatic summary.

Fact 11: Renault is the sycophantic police Captain.
Automatic summary:
RENAULT Take him quietly. Two guards at everydoor.
OFFICER Yes, sir. Everything is ready, sir.
The officer salutes and goes off to speak to the guards. Renault walks over to Strasser's table as Rick comes down the stairs.
RENAULT Good evening, gentlemen.
STRASSER Good evening, Captain.

Examples of less confident matches were as follows:

Fact 13: Renault detained Laszlo.
Automatic summary: (RENAULT talks to the German Major) Especially so tonight, Major. In a few minutes you will see the arrest of the man who murdered your couriers.

Fact 17: Ilsa was Rick's one time love.
Automatic summary: INT. RICK'S PARIS APARTMENT – DAY Ilsa fixes flowers at the window while Rick opens champagne. She walks over and joins him.

The first example indicates that Renault was going to arrest Laszlo, whereas the second shows that Rick and Ilsa used to be lovers in Paris in the past. So doubts exist here as to whether the facts are confirmed by the descriptions in the automatic summary.

3 Future Work

In this paper we have used the heuristic that the film shots containing the highest proportion of keywords should be included in the summary. We have examined the use of mid-frequency terms, but we will also try to identify high information keywords using the tf-idf measure. This will identify words which are common in a particular film script, but rare in film scripts generally. We would like to try other heuristics provided by Paice (1990) and Kupiek (1995) including the identification of key sentences (in our case, key lines of dialogue), those which more than 5 words long, and/or contain upper case words such as the names of people, places and organisations. Sentences can also be selected based on location clues: the first and last sentences in a paragraph tend to contain most information, as do the first 10 and last 5 paragraphs in a section (e.g. dialogue turns in a scene). We will then go on to use the text cohesion technique of Hoey (1991), which produces a network of interrelated sentences. We will examine the relations between these networks and the story ("plot unit") networks originally devised by Lehnert (1981) and developed by Xu (2006). Interviews will be held with the writers of existing IDMb summaries, to elicit rules for producing high quality human summaries which could be encoded in an automatic summariser.

The NetFlix Prize (www.netflixprize.com) exists as a

competition to find the best video recommender system. Many users also wish to view personalised, interactive video *summaries* (Money and Agius, 2008). Research to date has focused on generating summaries in different lengths according to user needs, but we wish to provide tailored summaries in terms of their content. This will involve the building of profiles of user's interests and previous choices, or requesting users to formulate their specific requirements in the form of a search engine query. Many techniques for user-focusing summaries are given by Liang (2008), in her work on query-biased summaries for search engines. We will use a metric widely used for matching queries and documents in search engines, the cosine similarity measure, for matching individual user queries and the lines of the film script most relevant to those queries.

Both the generic and user-tailored summaries will be evaluated by means of user satisfaction questionnaires. Liang (2008) distinguishes intrinsic and extrinsic evaluations. An intrinsic evaluation is made relative to the original – how well does the summary capture the information present in the original? An extrinsic evaluation is concerned with how well a summariser enables the user to perform a given task, in this case to judge whether or not a film should be viewed in its entirety.

Having produced and evaluated text summaries from collateral text, the next step will be to use the time-codings provided with the collateral texts to construct the video correlates of the text summaries. Time codings in audio descriptions will enable the key portions of the collateral text to be mapped onto the corresponding most important sections of the video. Another source of transcripts of multimedia content will be made available by EBU Technical's project P/SCAIE, who have made a public call for content summarisation tools and technologies (http://tech.ebu.ch/pscaie_rfts). EBU will evaluate the resulting summaries according to their own subjective evaluation of the product quality and the calculation complexity of the methods used.

In this paper we have only considered two films, *Casablanca* and *The Abyss*. In future we would also like to compare the effectiveness of automatic summarisation methods for other films in different genres, such as children's films, horror, or science fiction.

References

Aliguliyev, R. M. (2009) A new sentence similarity measure and sentence based extractive technique for automatic text summarization, *Expert Systems with Applications* 36(4), pp. 7764-7772, May 2009.

Aner-Wolf, A. and Kender, J. R. (2004) Video summaries and cross-referencing through mosaic-based representation, *Computer Vision and Image Understanding* 95(2), pp. 201-237, August 2004.

Antiqueira, L., Oliveira Jr., O. N., da Fontoura Costa, L. and das Graças Volpe Nunes, M. (2009) A complex network approach to text summarization, *Information Sciences* 179(5), pp. 584-599, 15 February 2009.

Dunlavya, D M., O'Leary, D. P., Conroy, J. M. and Schlesinger J. D. (2007) QCS: A system for querying, clustering and summarizing documents, *Information Processing & Management* 43(6), pp. 1588-1605, November 2007. Text Summarization

Fattah, M. A. and Ren, F. (2008) GA, MR, FFNN, PNN and GMM based models for automatic text summarization, *Computer Speech & Language* 23(1), pp. 126-144, January 2009.

Gervás, P., Díaz-Agudo, B., Peinado, F. and Hervás, R. (2004) Story plot generation based on CBR, *Knowledge-Based Systems* 18(4-5) pp. 235-242, August 2005.

Hakeem, A. and Shah, M. (2007) Learning, detection and representation of multi-agent events in videos, *Artificial Intelligence* 171(8-9), pp. 586-605, June 2007.

Hidekazu, T., El-Sayed, A., Masao, F., Kazuhiro, M., Kazuhiko, T. and Aoe, J-i. (2004) Estimating sentence types in computer related new product bulletins using a decision tree, *Information*

Sciences 168(1-4), pp. 185-200, 3 December 2004.

Hoey, M. (1991) "Patterns of Lexis in Text", Oxford University Press: Oxford, 1991.

IMDb (2009) Internet movie database. http://www.imdb.com last accessed 12 Dec. 2009.

Ko, Y. and Seo, J. (2008) An effective sentence-extraction technique using contextual information and statistical approaches for text summarization, *Pattern Recognition Letters* 29(9), pp. 1366-1371, 1 July 2008.

Ko, Y., Park, J. and Seo, J. (2004) Improving text categorization using the importance of sentences, *Information Processing & Management* 40(1), pp. 65-79, January 2004.

Kupiek, J., Pedersen, J. O. and Chen, F. (1995) A trainable document summarizer, In *Proceedings of the 18th Annual International Conference in Research and Development in Information Retrieval (SIGIR-95)*, pp 68-73, 1995.

Lehnert, W. G. (1981) Plot Units and Narrative Summarization, *Cognitive Science* 5(4), pp. 293-331, October – December 1981.

Lin, C-Y., and Franz, J. O. (2004) Automatic Evaluation of Machine Translation Quality Using Longest Common Subsequence and Skip-Bigram Statistics, *In Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL 2004)*, Barcelona, Spain, July 21 - 26, 2004.

Liang, S-F. (2008) "Triangulated Formulaic Measurement and Subjective Impression in Automatic Summarisation", *PhD Thesis*, University of Sunderland, UK.

Money, A. G. and Agius, H. (2008) Video summarisation: A conceptual framework and survey of the state of the art, *Visual Communication and Image Representation* 19(2), pp. 121-143, February 2008.

Money, A. G. and Agius, H. (2009) Analysing user physiological responses for affective video summarisation, *Displays* 30(2), pp. 59-70, April 2009.

Paice, C. D. (1990) Constructing Literature Abstracts by Computer: Techniques and Prospects, *Information Processing and Management*, 26(1): pp 171-186.

Smeaton, A. F. (2007) Techniques used and open challenges to the analysis, indexing and retrieval of digital video, *Information Systems* 32(4), pp. 545-559, June 2007.

Teufel, S. and Moens, M. (2002) Summarizing Scientific Articles: Experiments with Relevance and Rhetorical Status, *Computational Linguistics* 28(4), pp. 409-445, 2002

Xu, Y. (2006) "Representation of Story Structures for Browsing Digital Video", *PhD Thesis*, University of Surrey.

Ye, S., Chua, T-S., Kan, M-Y. and Qiu, L. (2007) Document concept lattice for text understanding and summarization, *Information Processing & Management* 43(6), pp. 1643-1662, November 2007. Text Summarization

Yeh, J-Y., Ke, H-R., Yang, W-P. and Meng, I-H. (2005) Text summarization using a trainable summarizer and latent semantic analysis, *Information Processing & Management* 41(1), pp. 75-95, January 2005.

Zhang, H. J., Wu, J., Zhong, D. and Smoliar, S. W. (1997) An integrated system for content-based video retrieval and browsing, *Pattern Recognition* 30 (4), pp. 643-658, April 1997.

Zhu, W. and Chen, C. (2007) Storylines: Visual exploration and analysis in latent semantic spaces, *Computers & Graphics* 31(3), pp. 338-349, June 2007.