

NewsViz: Emotional Visualization of News Stories

Eva Hanser, Paul Mc Kevitt, Tom Lunney and Joan Condell

School of Computing & Intelligent Systems

Faculty of Computing & Engineering

University of Ulster, Magee

Derry/Londonderry, BT48 7JL

Northern Ireland

`hanser-e@email.ulster.ac.uk,`

`{p.mckevitt, tf.lunney, j.condell}@ulster.ac.uk`

Abstract

The NewsViz system aims to enhance news reading experiences by integrating 30 seconds long Flash-animations into news article web pages depicting their content and emotional aspects. NewsViz interprets football match news texts automatically and creates abstract 2D visualizations. The user interface enables animators to further refine the animations. Here, we focus on the emotion extraction component of NewsViz which facilitates subtle background visualization. NewsViz detects moods from news reports. The original text is part-of-speech tagged and adjectives and/or nouns, the word types conveying most emotional meaning, are filtered out and labeled with an emotion and intensity value. Subsequently reoccurring emotions are joined into longer lasting moods and matched with appropriate animation presets. Different linguistic analysis methods were tested on NewsViz: word-by-word, sentence-based and minimum threshold summarization, to find a minimum number of occurrences of an emotion in forming a valid mood. NewsViz proved to be viable for the fixed domain of football news, grasping the overall moods and some more detailed emotions precisely. NewsViz offers an efficient technique to cater for the production of a large number of daily updated news stories. NewsViz bypasses the lack of information for background or environment depiction encountered in similar applications. Further development may refine the detection of emotion shifts through summarization with the full implementation of football and common linguistic knowledge.

1 Introduction

News reports are regarded as objective facts, commonly delivered in an objective, unbiased manner and represented in a neutral and formal format: typically a static headline, a summarizing paragraph with one image and eventually the body text with one to three more images. Even though reporters find the content of news stories worth mentioning for emotional reasons and the content often affects readers emotionally, story brevity, scarce background information and poor combination of visual and verbal information hinders learning and feeling by viewers. In order to reach the audience emotionally, to educate and to entertain, emphasis on visual elements is important as they tend to be more memorable than verbal ones. The emphasis of NewsViz lies on expression, impacting on the reader's understanding of the article and making it more memorable. The software prototype, NewsViz, automatically creates animations from news articles. Abstract design elements show emotions conveyed in the stories. The main objective of NewsViz remains information provision and thus our focus is emotion extraction which is universally applicable and without opinion bias. NewsViz is an efficient software tool for designers to be able to build daily updated animations. Input for NewsViz is natural language text. Multimodal systems automatically mapping text to visuals face challenges in interpreting human language which is variable, ambiguous, imprecise and relies on the communicative partners possessing common knowledge. Enabling a machine to understand a natural language text involves feeding the

machine with grammatical structures, e.g. part-of-speech, semantic relations, e.g. emotion value and intensity, and visual descriptions, e.g. colors and motion direction, to match suitable graphics.

2 Background and Related Research

Text-to-visual mapping relates to the areas of natural language processing (NLP) and multimodal storytelling which attempt to enable computers to interpret and generate natural human language and mental images. Text-to-visual mapping starts with linguistic analysis of the text. Despite variability, ambiguity and imprecision, syntactic analysis tools achieve mostly reliable results, such as trainable part-of-speech tagger software tools which identify parts of speech with 97% accuracy. For example, Qtag (Mason, 2003) attaches a tag to each word labeling it as noun, verb, adjective or other.

Semantic interpretation and actual understanding of the meaning of a text is more difficult, because it depends largely on commonsense knowledge. Commonsense knowledge and mental images need to be structured, related through logical rules and entered into databases before computational text interpretation is possible. WordNet (Miller, 1995) determines semantic relations between words and is an extended dictionary specifying word relations such as similarity, part-of relations, hierarchy or manner. Story segmentation is performed by e.g. SeLeCT (Stokes, 2003), an example application based on semantic analysis to find story or subtopic changes within a text. Groups of semantically related words called cohesive 'lexical chains' are extracted from a text. They are determined through WordNet's semantic relations and additionally through statistically acquired co-occurrences (e.g. Diego Maradona, Hand of God). Their starting and end points indicate topical unit boundaries.

Sensing emotions from multimodal input has mainly been investigated with the objective of developing human-like agents. The football commentary system, Byrne (Binsted and Luke, 1999), includes a commentator with emotions influenced by his personality and intentions. SOCCER (Retz-Schmidt, 1988) analyses football scenes visually in order to simultaneously add linguistic descriptions of the events. SOBA (Buitelaar et al., 2006) ex-

tracts information from soccer match reports, annotates relevant expressions (e.g. players, teams, goals.) and generates knowledge base entities. The collected football knowledge can set preconditions and context to consequently evaluate current events and assign appropriate emotions. The MoodNews website (Mitchell, 2005) demonstrates a very simple linguistic method to distinguish positive, negative and neutral content in BBC news headlines. It effectively ranks them on a color scale between good to bad. The three kinds of emotions are appointed through keyword scoring based on a small vocabulary of 160 words and phrases. The Emotion Sensitive News Agent (ESNA) (Shaikh et al., 2007) categorizes news stories from different RSS sources into eight emotion categories according to their emotional content, determined through a cognitive evaluation and user preferences.

Automated story visualization systems deliver initial results for object and action depiction, as in WordsEye (Coyne and Sproat, 2001), creating static 3D images from written descriptions. Additionally, automated camera and character animation, interaction and speech synthesis is realized in CONFUCIUS (Ma, 2006). ScriptViz (Liu and Leung, 2006) renders 3D scenes from NL screenplays immediately during the writing process, extracting verbs and adverbs to interpret events and states in sentences. The Unseen Video (Scheibel and Weinrother, 2005), is a good example of abstract mood visualization. Local weather data is automatically retrieved from news websites and influences the look and feel of the Flash animation through shapes, colors and images. The Story Picturing Engine (Joshi et al., 2004) visualizes texts selecting and matching pictures and their annotations from image databases.

The work discussed here demonstrates that sufficient subsets of the English language can be mapped to computer understandable language for the visualization of stories.

3 The NewsViz System

NewsViz takes online news articles as input and outputs animations reflecting the content of these news stories. NewsViz consists of three main components: the linguistic analysis, the animation composer and an interface for editing text and animations (Figure

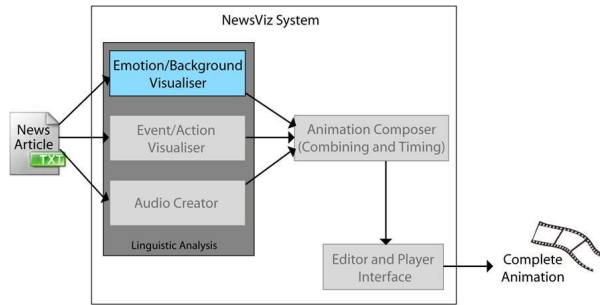


Figure 1: NewsViz System Architecture.

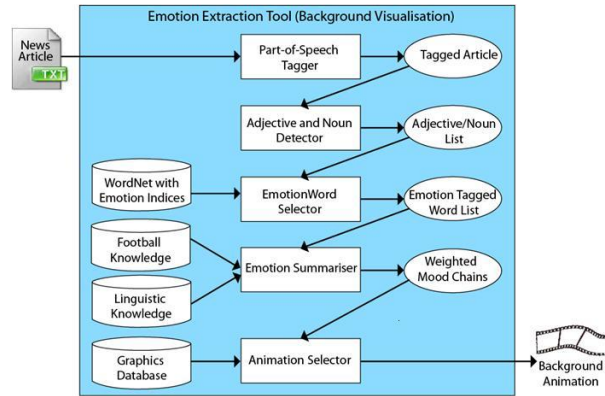


Figure 2: Emotion Extraction Component.

1). The *linguistic component* constructs three elements of the animation in different processes. The *emotion extraction tool* creates atmospheric background visuals, the *action visualizer* depicts people, objects and their actions and the *audio creator* selects music and sound effects. The *composer* synchronizes the different outputs. Here, we focus on the emotion extraction component (Figure 2) developed in Flash MX and Photoshop. Emotional aspects within the news story are identified and linked to appropriate presets of background animations.

3.1 Emotion Extraction

The first step in processing the text is to tag parts of speech for all words. The part-of-speech tagger, Qtag (Mason, 2003), attaches tags to nouns, verbs, adjectives and other parts of speech. The tagged text is sent on to the adjective and noun detector. These

two types of words are selected for further processing because they are most central to conveying emotional meaning and sufficient for the visualisation of the emotional content. Nouns and adjectives are the parts of speech which represent the highest number of affective words as found in WordNet-Affect (Strapparava and Valitutti, 2004). Verbs and adverbs will be addressed in future work to increase sensitivity and precision, but their impact on the resulting animations may not be as significant. Next, the emotion word selector checks the adjectives and nouns in the emotion dictionary and attaches emotion tags indicating their kind of emotion and intensity. The dictionary holds manually created emotion-indices and default intensity values of all affective words.

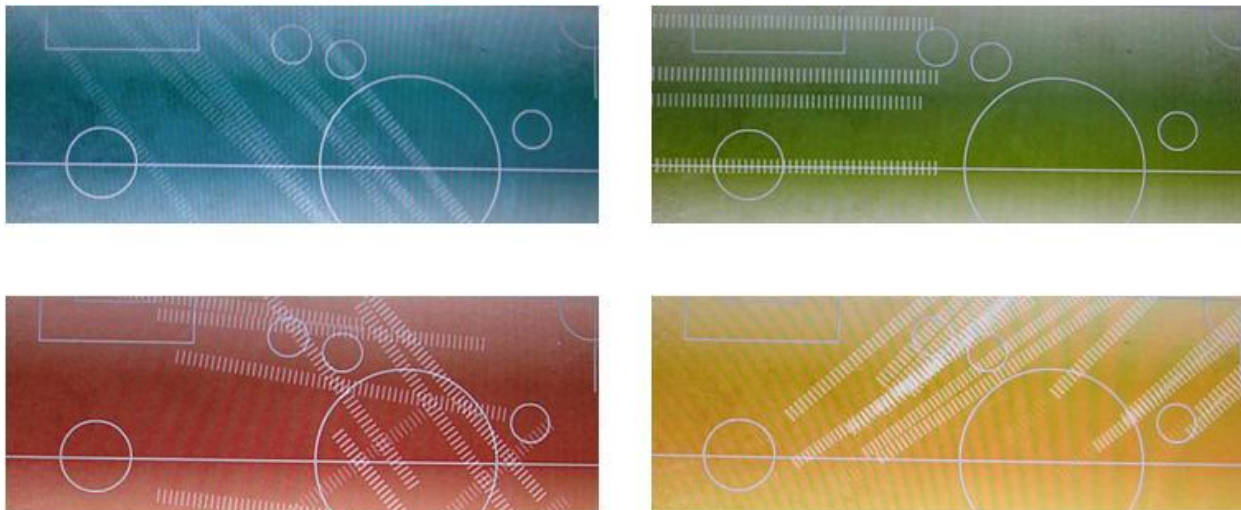


Figure 3: Animations for Sadness (blue), Boredom (green), Tension (red) and Happiness (yellow).

Four emotions have been found relevant in relation to football matches - happiness, sadness, tension and boredom. Words with a neutral emotion index do not describe football relevant emotions. To achieve a coherent course of emotion and animation, neutral phrases are replaced by the previous mood with decreasing intensity. The list of emotion tagged words is handed to the emotion summarizer. During the summarization process subsequent emotions of the same type are combined to form one longer-lasting mood. Each mood is labeled with its type, average intensity and display duration. With the 'word-by-word' summarization method mood boundaries appear as soon as the emotion type of the next word differs. In order to reduce error and excessive mood swings, the minimum threshold method sets a minimum number of words required to represent a mood. Alternatively, the sentence-based method assumes that one sentence conveys one idea and consequently one emotion. Hence, it calculates an average emotion for each sentence, before combining identical emotions. A chronological list of mood chunks is created.

3.2 Animation Construction

The animation selection component loads the individual animation elements from the graphics database and combines them in a 30 seconds long animation. The graphics database contains prefabricated graphics sorted by an emotion index which are combined and adjusted according to mood intensities. Based on the weighted mood list, the emotion sequence order, the type of graphic element, its display duration, and the background color are determined. The intensity value specifies the element size and the number of objects loaded. An emotion change causes the current animation elements to fade out and to load different elements. Animation examples are shown in Figure 3.

3.3 User Interface

NewsViz provides users with options to load or type news stories into the text editor. The options menu offers different emotion extraction and mood summarization methods. By pressing the 'run' button the visualization can be watched in the preview window. The text processing runs 'on the fly' in the background. If the user is satisfied they can save

the animation. If the user prefers to alter the animation manually, they have the option to edit the original text or the animation elements frame by frame. Figure 4 shows the user interface with animation player. The final animations are integrated at the top of the news article's internet page (Figure 5).

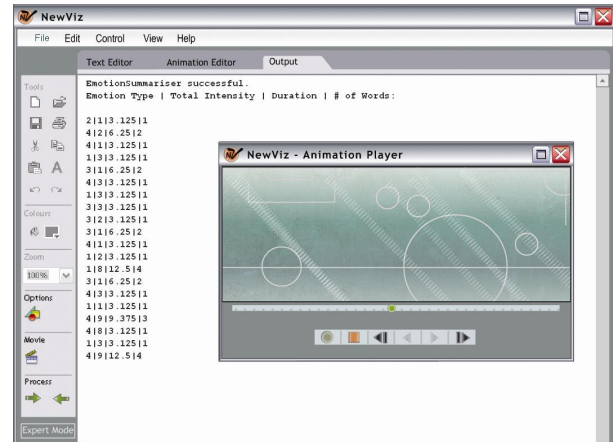


Figure 4: NewsViz User Interface.



Figure 5: Animation Integrated into Website.

4 Evaluation and Testing

NewsViz was tested on a set of four news articles related to the same news domain - football match reports. The articles were taken from BBC and FIFA online describing the same two World Cup 2006 matches. The three different emotion extraction methods, word-by-word, sentence-based and

Method \ Wordtype	Word by Word		Sentence based		Threshold				total
	correct	grain	correct	grain	2		3		
adjectives	3.125	12	3.25	7.5	2.375	5	1.25	2.3	2.5
nouns	3.875	31	2.625	9.3	2.875	14	2	4.8	2.844
both	4	33	2.75	9.5	3.5	18	1.5	10	2.938
total	3.667	25	2.875	8.8	2.917	12	1.583	5.7	

Figure 6: Results Analysis of all Test Texts.

threshold were run on these news stories with varying word types or word type combinations. The output of NewsViz is evaluated against two forms of human interpretation of the articles. A short manual description outlines the general course of emotion of a match as reported in each article naming three to five emotions. A second more fine grained interpretation assigns one (or two) emotions to each sentence. In correspondence to Beeferman’s probabilistic error metric (Beeferman et al., 1999) three types of emotion extraction error are distinguished. Falsely detected emotions are rated with zero points. Missing emotions were assessed depending on their significance in the text. If the overall feeling of the match was represented, two to three points would be given, but if the main emotions were missing, no points were assigned. Very close, but not exact emotions, got a value of four. A correct representation of the course of emotion received five points. The grain counts the number of the extracted emotions per text. The results for correctness of emotional findings and amount of emotions detected (grain) of each method run on each part-of-speech or word type combination are presented in Figure 6.

The results analysis shows that the effectiveness of adjectives or nouns varies from text to text, but generally the best results are achieved with the extraction of both kinds of words. On average the word-by-word method produces emotion sequences with the closest correctness, but unfortunately its output is too fine grained for visualization. Thirty second long animations are best visualized with two to ten mood swings. This means that some form of summarization is needed. Combining emotions of logically structured chunks of text, namely sen-

tences, in the sentence-based summarization method achieved better results than the minimum subsequent occurrence of two or three emotions with the threshold method. The sentence-based summarization as well as the threshold method with a minimum value of 3 produce the most appropriate grain/number of emotions. Some misinterpretation is due to false part-of-speech tagging by Qtag which has particular trouble with proper nouns. More accuracy can be achieved through training Qtag on football reports. Overall the results for NewsViz are satisfactory and it demonstrates that it is possible to extract emotions from news texts. The generally different sensations of the two described football matches are distinguishable. Three of the four test texts show good results, but for one article the extracted emotions do not seem to match the human sensation.

5 Relation to Other Work

NewsViz uses natural human language as input to create animated output. NewsViz aims to solely reflect emotions as they are mentioned in the news article to keep the objective and formal character of news reporting. Therefore, NewsViz applies a reduced, universal and ‘personality-free’ version of existing concepts for emotion and mood construction. Instead of facial expressions and gestures NewsViz combines and illustrates emotions with design principles. NewsViz offers manual reediting of the automatically created animations.

6 Conclusion and Future Work

NewsViz extracts emotion-bearing words from online football news reports based on an extended dictionary with emotion-indices assigned to each en-

try. The extracted emotions are processed and illustrated in abstract background animations. Results from initial testing demonstrate that this automated process has satisfactory performance. Technologically, NewsViz is viable for the fixed domain of football reports and offers a sound basis for more affective text-to-visual mapping. Future work will aim to improve the linguistic and semantic processing of emotions. This involves the extension of the parts of speech selection to include verbs and adverbs, assuming that more input data will lead to better results. Rules for common and linguistic knowledge will be integrated. Linguistic knowledge identifies emotions in context applying language rules to emotion interpretation, i.e. it solves negation by inverting emotions. With the integration of a dependency parser, which relates words according to their sentence structure, emotions of related words can be found and their average emotion determined. Domain-specific knowledge (e.g. football) provides background information including match statistics, players' and teams' names, team colors and league tables. It also accommodates game rules or match situations with their emotional consequences. The mood list is refined through moods discovered with commonsense knowledge and football facts which set pre-conditions and context representing long-term moods influencing current event-based emotions. Future work will reveal whether NewsViz is feasible when extended to different domains. The emotion database could be extended through the WordNet-Affect dictionary (Strapparava and Valitutti, 2004). NewsViz enriches standard news websites with attractive and informative animations and can track emotional aspects of people's views on world events. NewsViz brings news reported on the internet closer to readers, making it more easily understood and memorized which is much appreciated by online users overloaded with information. NewsViz assists animation designers in the production of daily updated visualizations creating initial scenes.

References

- D. Beeferman, A. Berger and J. Lafferty. 1999. Statistical models for text segmentation. *Machine Learning*, 34:177–210. Springer Netherlands.

- K. Binsted and S. Luke. 1999. Character Design for Soccer Commentary. *Lecture Notes in Computer Science. RoboCup-98: Robot Soccer World Cup II*, 1604:22–33. Springer-Verlag, London, UK.
- P. Buitelaar, T. Eigner, G. Gulrajani, A. Schutz, M. Siegel, N. Weber, P. Cimiano, G. Ladwig, M. Mantel, H. Zhu. 2006. Generating and Visualizing a Soccer Knowledge Base. *Proceedings of the EAACL06 Demo Session*, 4/2006:123–126.
- B. Coyne and R. Sproat. 2001. WordsEye: an automatic text-to-scene conversion system. *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, 487–496. ACM Press, Los Angeles, USA.
- D. Joshi, J. Z. Wang and J. Li. 2004. The Story Picturing Engine: Finding Elite Images to Illustrate a Story Using Mutual Reinforcement. *Proceedings of the 6th ACM SIGMM International Workshop on Multimedia Information Retrieval*, 119–126. ACM Press, New York, USA.
- Z. Liu and K. Leung. 2006. Script visualization (ScriptViz): a smart system that makes writing fun. *Soft Computing*, 10(1), 34–40. Springer Berlin/Heidelberg, Germany.
- Minhua Ma. 2006. Automatic Conversion of Natural Language to 3D Animation. *Ph.D. Thesis*. School of Computing and Intelligent Systems, University of Ulster, UK.
- O. Mason. 2003. Qtag. <http://phrasys.net/uob/om/software>.
- G. A. Miller. 1995. WordNet: a lexical database for English. *Communications of the ACM*, 38(11):39–41.
- Davy Mitchell. 2005. MoodNews. <http://www.latedecember.com/sites/moodnews>.
- G. Retz-Schmidt. 1988. A REPLAI of SOCCER Recognizing intensions in the domain of soccer games. *Proc. European Conf. AI (ECAI-88)*, 8:455-457.
- Daniel Scheibel and Ferdinand Weinrother. 2005. The Unseen Video. <http://www.theunseenvideo.com>.
- Mostafa Al Masum Shaikh, Helmut Prendinger and Mitsuru Ishizuka. 2007. Emotion Sensitive News Agent: An Approach Towards User Centric Emotion Sensing from the News. *Proceedings 2007 IEEE/WIC/ACM International Conference on Web Intelligence (WI07)*, 614–620. Silicon Valley, USA.
- N. Stokes. 2003. Spoken and Written News Story Segmentation Using Lexical Chains. *Proceedings of HTL-NAACL 2003*, 49–54. Edmonton, Canada.
- C. Strapparava and A. Valitutti. 2004. WordNet-Affect: an Affective Extension of WordNet. *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004)*, 1083–1086.