Visual Trend Analysis Method for Ontology Based Opinion Mining on Movie Reviews

Seongmin Mun [*] , Ginam Kim [†] , Gyeongcheol Choi [‡] , Kyungwon Lee [§]

[*‡] Life media Interdisciplinary Programs [†] Naver corp [§] Department of Digital Media, Ajou University

[*]skdkflak@naver.com, [†]mocca002@naver.com, [‡]ckc6842@gmail.com, [§]kwlee@ajou.ac.kr

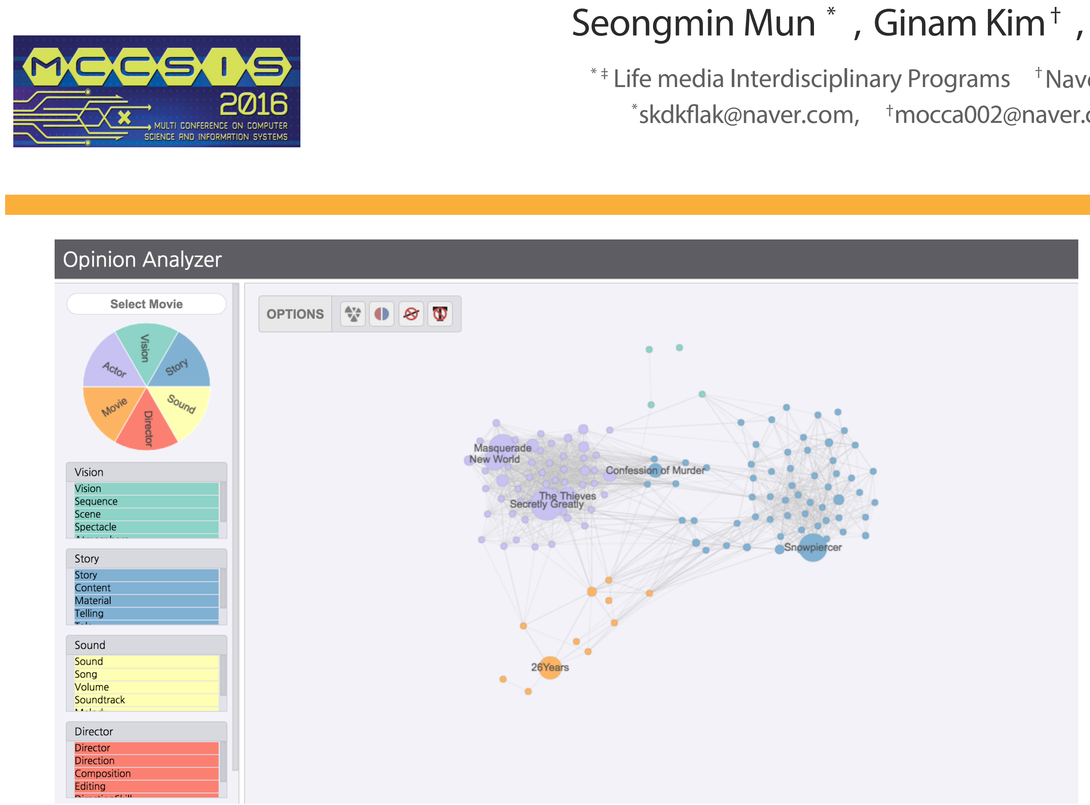

Figure 1. Over view (Demo site URL : http://54.255.190.140/index/v0#)

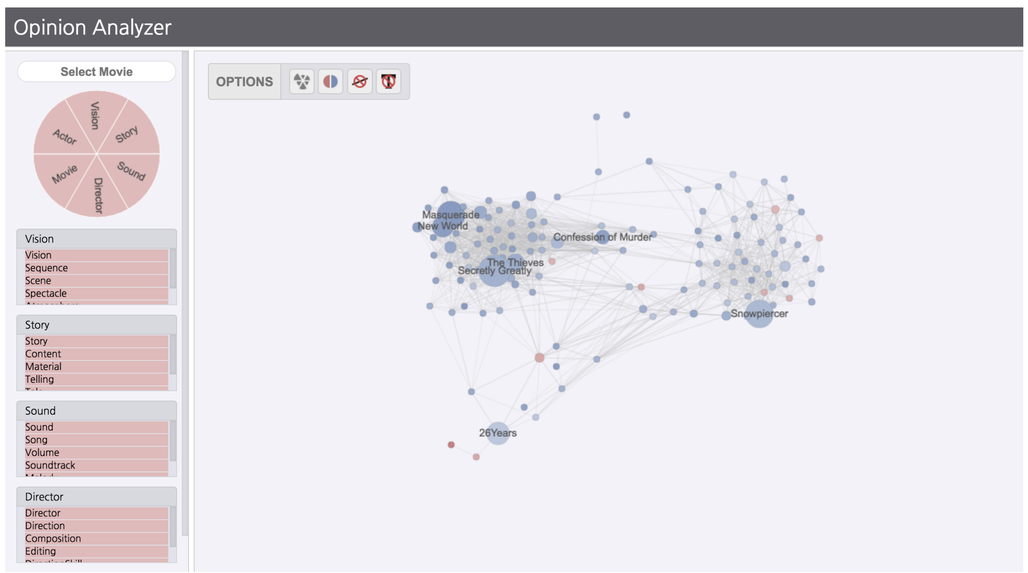

Figure 2. Over view colored by Polarity (Positive : Blue, Negative : Red)

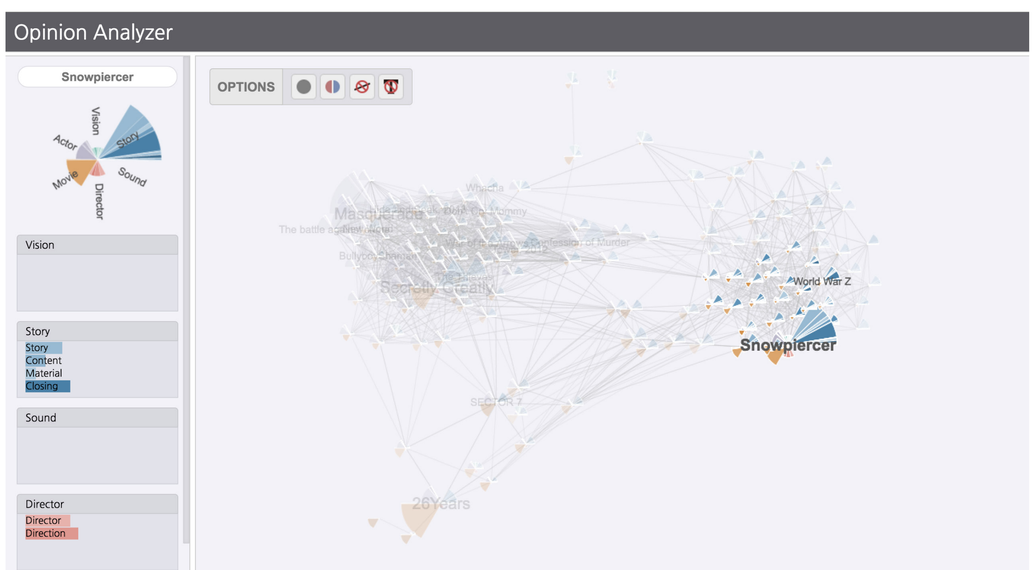

Figure 3. The clustering group centered by “Snowpiercer”

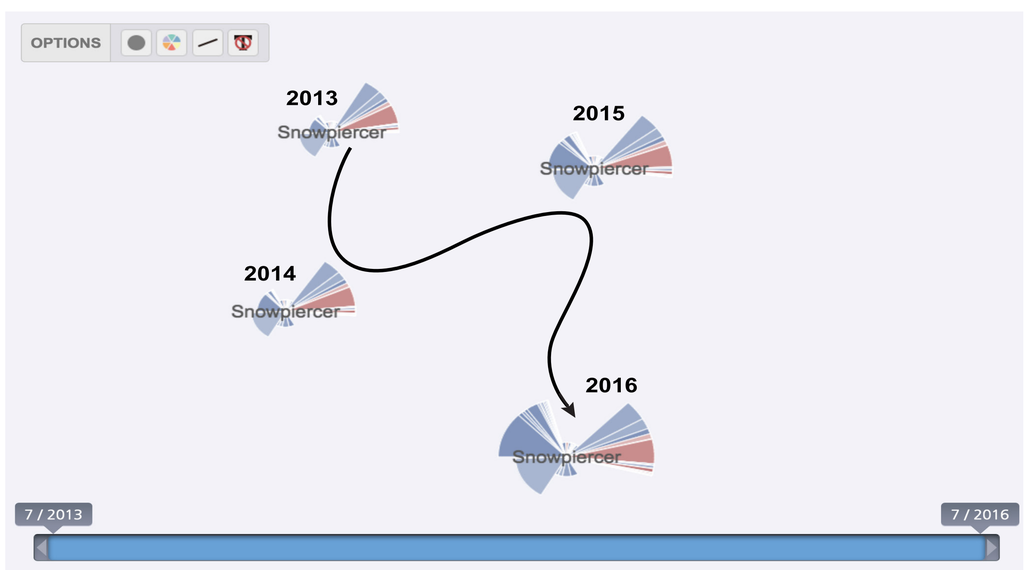

Figure 4. Changes of reviews on “Snowpiercer” over time

Introduction

As film industry is growing every year with the advance of media technology, online movie reviews are getting becoming predominant on potential audience behaviors to decide whether to be a consumer. We thus designed the ontology based upon such audience reviews in order to propose a methodology to analyze movie reviews in two ways; overall reviews regardless of time flow, and trending reviews reflecting how the opinions have changed over time. To demonstrate, we first constructed ontology to reflect the contents of individual reviews. Second, we conducted opinion mining based on the ontology. We next visualized the results and analyzed such results, presenting a scenario through the methodology of this research to analyze the changes of opinions in a movie review.

Ontology Construction

In order to construct an ontology for opinion mining, we first selected the categories of ontology and its evaluation keywords associated with each category. After extracting the index terms with high frequency in review data, we screened them as potential keywords and classified the finally selected terms. We further verified such process with topic modeling. Keywords utilized in our ontology in consequence of this progress are presented in below-ing Table.

Element class	Feature words
Director	Director, Direction, Composition, Editing, DirectionSkill, Attraction
Movie	Movie, Work, Overall, Scale
Actor	Performance, Actor, Character, MainCharacter, PerformanceSkill, Lines, SupportingActor, Casting, Voice, Style, Person, Appearance, Dubbing, Actress, Performer, Role, CastMember
Vision	Vision, Sequence, Scene, Spectacle, Atmosphere, Expression, Screen, Visual, Background, Graphic, MakeUp, Visuality, ImageBeauty
Story	Story, Content, Material, Telling, Tale, Closing, Ending, Scenario, Probability, Synopsis, Setting, OriginalWork, Finish, Subject, Accident, Structure, ContentTelling, StoryLine, Reading, StoryTelling, CastMember, StoryComposition
Sound	Sound, Song, Volume, Soundtrack, Melody, Music

Visualization

As well as providing the positive/negative tendency of each movie review, our new visualization method enables users to compare the evaluations of different movies since it forms clusters between movies with similar reviews, by comparing most heavily evaluated elements of each movie. After positive/negative tendency information visualization was performed, we further visualized clustering data in order to form the clusters of movies with similar reviews. We also designed an interface to facilitate more various visualization analyses through interaction with users.

Mining results of individual movies in this visualization are presented as one set of a pie chart and a bar graph. A pie chart, containing six directions of 12, 2, 4, 6, 8 and 10 o’ clock, indicates the frequency and polarity of each evaluation element. Size of a sector would increase when the elements are considered important as evaluation elements depending on its frequency, and vice versa.

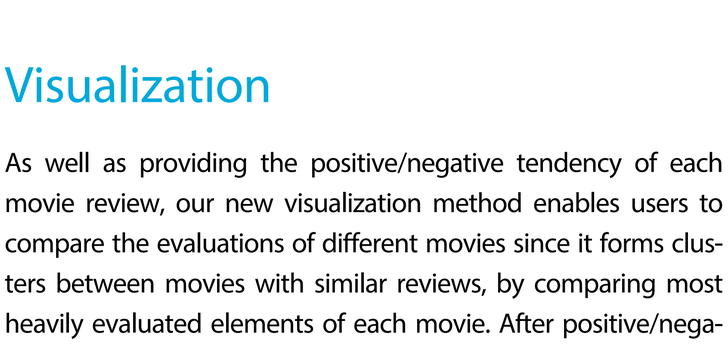

Figure 5. Opinion Mining Visualization for the each movie

Our finding that the location of nodes is determined according to the weight of evaluation elements when visualizing the clusters in Opinion Analyzer also implies that the location of nodes provides information on the weight of each element suggesting its importance.

Since it is general to analyze the relationships between the pre-clustered data groups if the amount of data is too large, sample movies of this research were divided into three groups.

Network Analysis represents a method to signify similar nodes, determining which node shares similar data with other nodes located in different clusters.

Trend Analysis

In addition to analysis and visualization of the overall reviews of a movie, this methodology can reflect the time flow as suggested in Figure which is a pilot data visualization of a movie Snowpiercer. In 2013 when the movie was first released, its plot was a major evaluation element, positive and negative reviews accounting for approximately equally. However it is found that Actors/Actresses gained more popularity in the movie review as time went on, with high positivity. It is also noticed that locations of nodes thereby changed according to the importance weight of evaluation elements.

Conclusion

This study aimed to propose a methodology to analyze and visualize the movie reviews in two ways; overall and trending. We thus constructed an ontology of movie reviews and suggested a visualization analysis methodology in order to analyze the movie reviews. Major conclusions of this research are as follows. First, visualization analysis on each movie revealed that overall tendency of reviews of each movie differ from each other. Second, 130 films as samples in this research were clustered as three groups depending largely on the feedback from the audience. Third, it is suggested that this visualization analysis methodology can be applied to review trend analysis that reflects the changes of audiences’ opinion over time.

In addition to these findings, this research also proposes a new framework of ontology on movie domain applying keyword extraction and topic modeling, and that argument structure identification was applied in opinion miming.

References

[1] Hui Yang, Minjie Zhang. Ontology-based resource descriptions for distributed information sources. Information Technology and Applications, 2005. ICITA 2005. Third International Conference on, pages 143-148. IEEE Computer Society. 2005.

[2] Chao He, Yu-feng Zhang. Research on semantic association pattern mining model based on ontology. Advanced Computer Theory and Engineering (ICACTE), 2010 3rd International Conference on, volume 1, pages 497-501. IEEE Computer Society, 2010.