Multimodal Communication on Tumblr: “I have so many feels!”

Elli E. Bourlai
Indiana University
ebourlai@indiana.edu

Susan C. Herring
Indiana University
herring@indiana.edu

ABSTRACT
We manually analyzed a corpus of Tumblr posts for sentiment, looking at images, text, and their combination. A dataset was constructed of posts with both text and images, as well as a dataset of posts containing only text, along with a codebook for classifying and counting the content in each. This paper reports on the construction of the overall corpus and the codebook, and presents the results of a preliminary analysis that focuses on emotion. Posts containing images expressed more emotion, more intense emotion, and were more positive in valence than posts containing only text. The study contributes a micro-level analysis of multimodal communication in a social media platform, as well as a gold standard corpus that can be used to train learning algorithms to identify sentiment in multimodal Tumblr data.

Categories and Subject Descriptors
K.4.m [Computing Milieu]: Computers and Society – *miscellaneous*.

General Terms
Measurement, Human Factors.

Keywords
Communication, GIF, image analysis, meme, multimodality, sarcasm, sentiment, social media.

1. INTRODUCTION
Increasingly, meanings are expressed through images in social media. The use of animated GIFs to express opinions and reactions, for example, is popular on sites such as 4chan and Tumblr, as well as in the comment sections of forums and blogs. On some sites, users continuously create new GIFs to express a range of attitudes and emotions, some of which become memes [11] and spread to other internet contexts [2]. What research methods should be used to analyze these new meaning units?

One possibility is to employ sentiment analysis and machine learning to analyze attitudes expressed through images. However, although sentiment analysis (SA) has been an active area of research in recent years [12], most SA research focuses on text; fewer studies have analyzed images, let alone animated images. This is because automating image SA is challenging; thus far, it has tended to rely on textual tags/metadata provided by users and low-level visual features [22]. Manual annotation of the expressive meanings conveyed in images would provide richer information and could raise the quality of sentiment studies that make use of machine learning, but it is a time-consuming process.

Moreover, it is not enough to analyze images alone. Images and text work together to create meaning on multimodal social media sites: The text often provides context for the images, indicating, for example, that the image’s apparent message is intended sarcastically. Sarcasm is especially difficult to detect using automated methods [7]. Furthermore, text is sometimes embedded directly into images [13], making it part of the image’s meaning. Thus text and image need to be analyzed in relation to one another. Few studies have done this in a systematic way.

In the study reported in this paper, we manually analyzed a corpus of Tumblr posts for sentiment, looking at images, text, and their combination. Two datasets were constructed – one of posts with both text and images, and the other of posts containing only text – along with a codebook for classifying and counting the content in each. Here we report on the construction of the two datasets and the codebook, and present the results of a preliminary study of a subset of the corpus that focused on emotion. The results indicate that posts with images express more emotion, more intense emotion, and are more positive in valence than posts containing only text. The study thus makes two contributions: It advances knowledge of multimodal (especially image) communication in social media, and it provides a manually-tagged gold standard corpus [20] that could be used to train learning algorithms to identify sentiment in multimodal Tumblr data.

2. BACKGROUND
2.1 Tumblr
Tumblr is a microblogging service that was founded in February 2007 by David Karp. As of 2014, it had 172.2 million blogs and 77.2 billion posts [19]. Tumblr users have their own individual blog(s), on which they can post new content or “reblog” content posted by other Tumblr users. Users can choose among seven rebloggable types of posts: text, photo, link, audio, video, chat, and quote. Tumblr also has a Private Messaging (PM) feature that lets users reply to messages received either privately or publicly. Most Tumblr users are female and under 34 years old [15].

Tumblr is especially known for its use of reaction GIFs: short clips of movies and television shows that communicate emotions (“feels”), reactions, and everyday events [4]. These images typically originate in Tumblr and spread to other electronic environments [6]; some of them become associated with specific meanings and functions, and acquire the status of internet memes. Despite Tumblr’s popularity and reputation as a source of reaction GIFs, however, we are not aware of any studies that have been conducted of such images.

2.2 Relevant Literature
Several studies have applied machine learning techniques to predict the sentiment of images in social media. For example, [17] analyzed a corpus of Flickr photographs using user-generated text tags and low-level visual features such as color to train a classifier. [21] leveraged mid-level attributes of Twitter images to predict their sentiment, including material (e.g., metal), function (e.g., playing), surface property (e.g., glossy), spatial envelope (e.g., man-made), and facial expression in images of people.

Multimodal online content has also been analyzed using discourse and content analysis methods, although few such studies have considered images. An exception is McDonald [13], who qualitatively analyzed interactive image-based communication on a community image blog. He identified four styles of visual 'conversation' in interactive image exchanges: positional play, image quote, text-in-picture, and animation. An image quote, in which a participant takes a picture posted by another participant, modifies it, and reposts the picture, is similar to a meme.

Internet memes have attracted research attention independent of their function in conversational exchanges. [11] used discourse analysis methods to identify key characteristics of successful internet memes reported in mainstream media venues between 2001 and 2005. The successful memes contained humor, wry intertextual cross-references to everyday and popular culture events, and/or anomalous juxtapositions. The study proposed a typology of meme purposes that includes social commentary, absurdist humor, and hoaxes. [2] investigated the spread of 150 famous Internet memes using time series data from Google Insights, Delicious, Digg, and StumbleUpon, and discovered that the temporal distributions that characterize meme popularity are mostly heavily skewed and long-tailed.

With the exception of [13], none of these studies analyzed image use in interactive online communication. Although some consider animated images, including YouTube videos, none include GIFs, whose popularity is a recent phenomenon. Moreover, we found no studies of Tumblr images. The present study contributes to filling these gaps.

2.3 Research Questions

This study addresses the following research questions:

- Do Tumblr users communicate differently in text vs. images?
- If so, how do textual and image communication differ?

Specifically, we focus in this paper on the expression of emotion in text versus images.

3. METHODOLOGY

3.1 Corpus Construction

In order to examine the use of the two modes on Tumblr, a systematic corpus with two datasets was constructed. The first (TXT dataset) comprises posts containing only text, and the second (IMG dataset) comprises posts with images and/or images and text. In order to collect user-generated data representing both text and image use efficiently, we sampled from popular user-generated tags. We first examined the 10 most popular user-generated tags [1] to identify tags that referenced posts with images (at least 10 of the 50 most recent posts). Four of the tags met these criteria, and a fifth tag that was very popular at the time of our data collection (December 2012) but was not mentioned in [1] was added. The first five tags are mostly used by fan communities ("fandoms"), the majority users of Tumblr. A sixth more generic tag (#feels) was added to diversify the sample.

The IMG dataset was collected first on five different days (three weekdays and two weekend days) during peak posting times on Tumblr (6:00pm-12:00am EST), then the TXT dataset was collected using the same systematic procedure. We manually examined the 100 most recent posts for the five fandom-based tags and the 200 most recent posts for the generic tag on each of the five days, selecting posts according to the following criteria:

a) The language used is English.
b) For the image dataset, each post includes at least one image used for discourse purposes (as opposed to simple reblogging of images). For the text dataset, posts contain only text.
c) Each post appears only once in the dataset (we excluded identical reblogged posts).

Table 1 shows the frequencies of the posts collected for each tag in each dataset.

Table 1. Post frequencies by tag and dataset

Tag	IMG Dataset	TXT Dataset	Total
#onedirection	56	103	159
#tomhiddleston	178	94	272
#legendofkorra	62	153	215
#loki	144	101	245
#supernatural	191	241	432
#feels	436	393	829
Total	1,067	1,085	2,152

The posts were imported as text into two Microsoft Excel spreadsheets (one for each dataset), preserving as much of their formatting as possible (e.g., italics and bold font for emphasis), where each row represents a functional move. Because Excel does not allow the insertion of pictures into cells, images were inserted as links in the location they would appear in the post. Posts on Tumblr can be ephemeral, either because users delete posts or change their usernames (which are part of their blogs' URL). This meant that some links saved in the spreadsheet would not work when we tried to access them to code the images. To address this problem, screenshots were taken of image posts during data collection and were saved in a separate folder for reference.

3.2 Content Analysis Codebook

The methodology employed for analyzing the posts is content analysis using a grounded theory approach, which allows coding categories to emerge from the data. The first step was to identify the coding unit. While some categories of interest can be applied to posts as a whole, the content of a Tumblr post is often heterogeneous across its different sections (title, body, tags) and/or within a section (e.g., a section may contain different modes and/or different emotion polarities). In addition, the images should be analyzed as part of the posts rather than as separate entities, because they play an integral role in the overall meaning of the post. We addressed these issues by subdividing posts into *functional moves*, a concept from linguistic text analysis that breaks down larger entities into a "schematic organization" or "conventional sequence of 'chunks'" [8]. Dividing a post into smaller, internally consistent parts allowed the same coding categories to be used for both modes (text and image).

For the corpus as a whole, since this is an exploratory study, the codebook includes multiple variables in three categories: demographics (gender and age of poster), structure, and function. The first is coded at the post level; the last two are coded at the post or functional move level, depending on whether the variable refers to the whole post or part of it, as described further below.

3.2.1 Demographics

Demographic information about the posters was coded during data collection. All posts were coded for gender and age of the users. Gender and age information was extracted from usernames, bio descriptions, and post content on the users' Tumblr blogs.

The users for 1,698 posts out of the total 2,152 were clearly identifiable as male or female. Females are equally distributed between the two datasets, while 60% of males use image communication and 40% use textual communication.[1]

3.2.2 Structure

Structural variables are identified from the format of the post itself. The variables and their code values are shown in Table 2.

[1] The overall posting patterns of the genders across the two datasets is significantly different, $\chi^2(1, N=1{,}698)=7.605$, $p=.006$.

Table 2. Codebook for structure variables

Unit of Analysis	Variable	Values
Post	Post Type	Text, Photo, Link, Audio, Video, Chat, Quote, Reply, N/A
Post	Interaction Type	PM + PR (rebloggable), PM + PR (non-rebloggable), PP + Reply/Comment, N/A
Post	Text in Post	Yes, No, N/A
Functional Move	Location of Text	Title, Body, Tag (+ Text-in-Image and N/A for IMG Dataset)

The variables coded at post level describe technically-specified characteristics of the posts' format. Post Type includes the seven types of posts offered by Tumblr and what users call a "reply" post, a private message (PM) with a public reply (PR). At the time of our data collection, replies could not be reblogged by other users. To overcome this limitation, Tumblr users took screenshots of private messages they received and embedded them into one of the seven rebloggable types of posts, adding their reply below the screenshot.[2] The Interaction Type of each post is thus coded according to whether it is a non-rebloggable Reply (PM + PR), a rebloggable Reply (PM + PR), or any other type of public post (PP) that was reblogged and contained a comment added by another user. The value N/A for Post Type and Interaction Type is used in cases where either is unclear or information is missing.

At the functional move level, the location of the text is coded for both datasets for the title, body, and tag sections. Two values were added for the IMG dataset: Text-in-Image for images that contain text and N/A for images that do not contain text.

3.2.3 Function

The codebook for function variables is shown in Table 3. Function is expressed mostly on the semantic and pragmatic levels of discourse and requires human interpretation. Because functions can be ambiguous, content analysis employs multiple coders to increase the reliability of functional codes. In the study described in section 4, the authors discussed all ambiguous cases in order to reach agreement for coding.

Table 3. Codebook for function variables

Unit of Analysis	Variable	Values
Post	Role-play	Yes, No
Post	Purpose	GIF Challenge, Personal Situation, Fandom-related Situation, Mixed, Other
Functional Move	Expression	Self-expression, Expression other than self, Mixed, N/A
Functional Move	Reaction	Yes, No, N/A
Functional Move	Presence of Emotion	Yes, No, Unclear
Functional Move	Polarity of Emotion	Positive, Negative, Mixed, N/A
Functional Move	Type of Emotion	Anger, Anticipation, Disgust, Fear, Joy, Sadness, Surprise, Trust, N/A
Functional Move	Intensity of Emotion	Extreme, Non-Extreme, N/A
Functional Move	Bona Fide	Yes, No, N/A
Functional Move	Person in Image Saying the Same as Text	Yes + Text-in-Image, Yes + Text-Outside-Image, No, N/A
Functional Move	Image Function	Mention, Use, Unclear, N/A
Functional Move	Image Format	Static, Dynamic, N/A

At the post level, the data are coded for role-playing, i.e., whether the user was posing as a fictional character from a fandom universe. The posts are also coded for purpose based on whether their content describes a personal situation, a fandom-related situation, a mixture of the two, or a GIF Challenge game, where users have to pick the n^{th} GIF from their personal GIF folder on their computer, which then represents their reaction to a situation described in the directions for the game.

Most of the function variables require separate analysis of each section of the post and are coded at the functional move level. Each functional move is coded for whether the text or image expresses the perspective of the user, another individual, or both (identified, e.g., by use of 1st vs. 3rd person pronouns); whether it is a reaction to something (as indicated by expressions such as "and then I was like", "I freaked out"); presence, polarity, and type of emotion[3] (as determined by words expressing emotion in text or facial expressions in images); intensity of emotion (as determined by capitalization of words, repetition of letters and punctuation, bold or italic fonts in text; intense movements and facial expressions in images); and sarcasm (whether the communication is bona fide/genuine or not).

The image dataset is additionally coded for three variables. Animated images featuring a person (or a personified object or animal) are coded for whether the person appears to say the same thing as the text in the image, the text around the image, or something different. The N/A code is used for images not featuring talking. Images are also coded for their function: as part of discourse ('use') or as a picture that was simply reblogged and commented upon ('mention'). Finally, images are coded as static (usually in JPG or PNG format) or dynamic (GIF format). Most of the images in the IMG dataset are dynamic GIFs.

4. PRELIMINARY ANALYSIS

4.1 Data

For the preliminary analysis, we coded a subset of the data in the corpus from the #tomhiddleston and #feels tags. The first 50 posts in each tag were coded from each of the datasets (100 posts from each dataset, for a total of 200 posts). Table 4 displays the frequencies of the posts and the functional moves contained in the subset. The IMG Dataset has more functional moves than the TXT Dataset, because its posts were considerably longer, and each image was coded as a separate functional move.

Table 4. Subset post and functional move (FM) frequencies

Tag	IMG Dataset	TXT Dataset	Total
Posts	100	100	200
Functional Moves	609	323	932

Analysis of the data subset using chi-square tests was conducted on variables relating to the expression of emotion. The variable tested at the post level was posting Purpose, and the variables tested at the functional move level were Emotion

[2] Tumblr recently modified public reply posts to make them rebloggable.

[3] The typology used for the types of emotions was taken from the NRC Word-Emotion Association Lexicon [14].

Presence, Emotion Intensity, Emotion Polarity, and Sarcasm (the Bona Fide variable). Most of the variables tested showed significantly different mode patterns across the two datasets.

4.2 Results

4.2.1 *Purpose*

The subset analyzed did not contain any GIF Challenge posts; therefore only the posting purposes personal situation, fandom-related situation, and mixed were tested. Because of the small number of posts (N=200), the results of the tests for significance were inconclusive. However, the proportions show a preference for image use when describing a fandom-related situation and for text use when describing a personal or mixed situation.

4.2.2 *Emotion*

Three variables related to emotion were tested for overall pattern differences and for pattern differences within each dataset. Excluding cases where emotion was not clearly present, the two datasets show differences overall in emotion presence, $\chi^2(1, N=899)=48.481$, $p<.001$. For cases where emotion is present, intensity of emotion is also significantly different overall, $\chi^2(1, N=660)=8.652$, $p=.003$. Finally, excluding cases where emotion polarity is mixed or unclear, there are significantly different distributions overall of negative and positive emotion in the two datasets, $\chi^2(1, N=532)=26.751$, $p<.001$. Binomial tests were conducted to examine differences within each dataset; the results are presented in Tables 5-7.

Table 5. Binomial test results for emotion presence

Dataset	Observed Prop. (Yes)	Observed Prop. (No)	*Significance*
IMG	.81	.19	<.001
TXT	.59	.41	.002

Although both datasets are emotional, posts containing images are much more emotional than plain text posts.

Table 6. Binomial test results for emotion intensity

Dataset	Observed Prop. (Extreme)	Observed Prop. (Non-Extreme)	*Significance*
IMG	.41	.59	<.001
TXT	.28	.72	<.001

Neither textual nor image posts usually convey extreme emotion. However, there is more extreme emotion in the image dataset.

Table 7. Binomial test results for emotion polarity

Dataset	Observed Prop. (Positive)	Observed Prop. (Negative)	*Significance*
IMG	.57	.43	<.001
TXT	.32	.68	<.001

Table 7 shows a reversal of proportions in the two datasets. Posts containing images are mostly positive, while posts with plain text are mostly negative in emotion polarity.

4.2.3 *Sarcasm*

Each functional move was coded for whether it showed genuine communication (bona fide) or sarcasm (non-bona fide). The patterns are overall significantly different between the two datasets, $\chi^2(1, N=803)=5.643$, $p=.018$, as well as within each dataset, as Table 8 shows.

Table 8. Binomial test results for sarcasm

Dataset	Observed Prop. (Bona Fide)	Observed Prop. (Non-Bona Fide)	*Significance*
IMG	.69	.31	<.001
TXT	.61	.39	<.001

Most functional moves in both datasets convey genuine messages. However, there is more sarcasm in purely textual communication than in communication using images.

4.3 Discussion

It appears that mode choice on Tumblr is not arbitrary. The results of our preliminary study suggest that Tumblr users communicate differently in text and image posts. Image communication is used more for describing fandom-related situations; it conveys more emotion than textual communication, greater intensity of emotion, and the emotion expressed in images is mostly positive. Text, in contrast, is used more to describe personal situations and express sarcasm, and it conveys more negative emotion compared to images. The greater negativity of text is consistent with what Herring and Demarest [9] found for text vs. audio and video posts on the multimodal commenting site Voicethread.com.

Communication in both Tumblr modes is mostly bona fide (although there is a considerable amount of sarcasm), and it is quite emotional overall. A possible explanation for the latter finding is that the #feels tag analyzed in the data subset conveys emotion by definition; it might be biasing the results. We checked for this possibility by analyzing the #tomhiddleston tag separately; the results were similar to those for both tags combined. Another possible explanation is that a majority of Tumblr users (as well as in our dataset) are female; previous research has found that females tend to express more emotion online, especially positive emotion, than males do [8][18]. Moreover, the tags used for data collection are mostly fandom-related and attract a younger audience, and young women in western cultures are socialized to be emotional [5]. These findings underscore the importance of taking user demographics and topic of communication into consideration when analyzing emotion in social media.

The differences between the two datasets can be explained in terms of the nature of the two modes. Image is a socially richer mode than text [16]: It can express facial expressions[4] and other iconic representations of emotion (e.g., a wave breaking over a person to show overwhelming emotion[5]). Moreover, images are more efficient at depicting humorous situations (e.g., boxes falling on top of a person[6]), which might explain the positive orientation of the image dataset. Image memes also tend to be rich in intertextuality [11], which may be why they are preferred when users post about fandom-related subjects. Previous studies [e.g., 11] have claimed that text is more distancing; this could help explain why the textual dataset has more negative polarity. Additionally, many of the textual posts that have negative emotion include the hashtags “#vent” and “#rant,” suggesting that the users wanted to “let out” their feelings quickly, rather than following the more time-consuming process of selecting and inserting an image in a post. Finally, a possible reason that textual communication is more sarcastic than image communication is

---

[4] See, e.g., http://media.tumblr.com/d5c26c47423209fa62a47391843972b6/tumblr_inline_mhtdvjq4Lq1qz4rgp.gif

[5] See, e.g., http://media.tumblr.com/1c63b3ba7349f2a9f2572ed3120e209a/tumblr_inline_mhtdtzBQew1qz4rgp.gif

[6] See, e.g., http://media.tumblr.com/f2a7bf22bc7228381063228a65a18b4a/ tumblr_inline_mhthw2aoTi1qz4rgp.gif

that sarcasm relies on ambiguity. Text, being less rich in social cues [16], may lend itself better to ambiguity.

5. CONCLUSIONS

This paper described the construction of two datasets to analyze Tumblr text and image content, as well as the creation of a content analysis codebook for analyzing the corpus manually. It further presented the results of a preliminary analysis of a subset of the corpus using code variables focused on emotion expression.

This research has several implications, especially for sentiment analysis. Image memes, both static and dynamic, are becoming more common and are spreading across the social web, for example, in the comment sections of blogs and forums. Our findings indicate that images convey more emotion than plain text. It follows that analyzing images along with text in multimodal environments should improve the performance and result in greater accuracy of sentiment analysis. Our hand-coded image dataset can also be used further as a training dataset for sentiment analysis using machine learning. As a research study, the paper contributes a micro-level analysis of multimodal communication in a social media platform, especially of image communication, which has been understudied. Finally, it sheds light on the production of Tumblr memes.

Our future work will follow the agenda set out in this paper. We are currently analyzing the whole corpus using the full codebook. We plan to separate dynamic from static images in future analyses; dynamic images should be socially richer than static images, in that they add movement [16], and therefore they can be expected to express meaning differently. Finally, we plan to collect additional Tumblr corpora from randomly-selected tags to evaluate the extent to which our preliminary findings, which are oriented towards fandom, can be generalized across the Tumblr platform.

6. REFERENCES

[1] Alfonso, F. 2012, June. The 10 most popular tags on Tumblr. *The Daily Dot.* Retrieved from http://www.dailydot.com/entertainment/10-most-popular-tags-tumblr/

[2] Bauckhage, C. 2011. Insights into internet memes. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media* (Barcelona, Spain, July 17-21). The AAAI Press, Menlo Park, CA, 42-49.

[3] Borth, D., J. Rongrong, Chen, T., Breuel, T., and Chang, S-F. 2013. Large-scale visual sentiment ontology and detectors using adjective noun pairs. *Proceedings of the 21st ACM International Conference on Multimedia* (Barcelona, Spain, October 21-25). ACM, New York, 223-232.

[4] Casserly, M. 2012, March. #WhatShouldWeCallMe revealed: the 24-year old law students behind the new Tumblr darling. *Forbes*. Retrieved from http://www.forbes.com/sites/meghancasserly/2012/03/29/whatshouldwecallme-revealed-24-year-old-law-students-tumblr-darling/

[5] Currie, D. H., Kelly, D. M., and Pomerantz, S. 2006. The geeks shall inherit the earth. *Journal of Youth Studies*, 9, 4 (Sep. 2006), 419-436.

[6] Faircloth, K. 2012, October. Tumblr is down, so how're we supposed to get our GIFs now? *BETABEAT*. Retrieved from http://betabeat.com/2012/10/tumblr-is-down/

[7] González-Ibáñez, R., Muresan, S., and Wacholder, N. 2011. Identifying sarcasm in Twitter: A closer look. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (Portland, OR, June 19-24). *Scopus*®, 581–586.

[8] Herring, S. C. 1996. Two variants of an electronic message schema. In *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives,* S. C. Herring, Ed. John Benjamins, Amsterdam, 81-101.

[9] Herring, S. C., and Demarest, B. 2011. Mode choice in multimodal comment threads: Effects on participation and language use. Unpublished ms.

[10] Kiesler, S., Siegel, J., and McGuire, T. W. 1984. Social psychological aspects of computer-mediated communication. *American Psychologist*, 39 (Oct. 1984), 1123-1134.

[11] Knobel, M., and C. Lankshear 2006. Online memes, affinities, and cultural production. *A New Literacies Sampler*, M. Knobel and C. Lankshear, Eds. Peter Lang, 199-227.

[12] Liu, B., and Zhang, L. 2012. A survey of opinion mining and sentiment analysis. *Mining Text Data*, C. C. Aggarwal and C. Zhai, Eds. Springer US, 415-463.

[13] McDonald, D. 2007. Visual conversation styles in web communities. *Proceedings of the 40th Hawaii International Conference on System Sciences*. IEEE Computer Society Washington, DC, 76.

[14] Mohammad, S., and Turney, P. 2010. Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. *Proceedings of Workshop on Computational Approaches to Analysis and Generation of Emotion in Text* (Los Angeles, CA, June 5). ACL, Stroudsburg, PA, 26-34.

[15] Pingdom, 2012. Social network demographics in 2012. *Royal Pingdom*. Retrieved from http://royal.pingdom.com/2012/08/21/report-social-network-demographics-in-2012/

[16] Short, J., Williams, E., & Christie, B. 1976. *The social psychology of telecommunications*. London: Wiley.

[17] Siersdorfer, S., and Hare, M. 2010. Analyzing and predicting sentiment of images on the social web. *Proceedings of the 18th ACM International Conference on Multimedia* (Firenze, Italy, October 25-29). ACM, New York, 715-718.

[18] Thelwall, M., Wilkinson, D., and Uppal, S. 2010. Data mining emotion in social network communication: Gender differences in MySpace. *Journal of the American Society for Information Science and Technology,* 61(1), 190–199.

[19] Tumblr. (2014). *About* page. Retrieved February 19, 2014 from http://www.tumblr.com/about

[20] Wiebe, J., Bruce, R., and O'Hara, T. 1999. Development and use of a gold-standard data set for subjectivity classifications. *Proceedings of the 37th annual meeting of the Association for Computational Linguistics* (College Park, MD, June 20-26). ACL, Stroudsburg, PA, 246-253.

[21] Yuan, J., Mcdonough, S., You, Q., and Luo, J. 2013. Sentribute: Image sentiment analysis from a mid-level perspective. *Proceedings of the Second International Workshop on Issues of Sentiment Discovery and Opinion Mining* (Chicago, IL, August 11-14). ACM, New York.

[22] Zha, Z-J., Wang, M., Shen, J., and Chua, T-S. 2010. Text mining in multimedia. *Mining Text Data*, C. C. Aggarwal and C. Zhai, Eds. Springer US, 361-384.