

The Color of Emotions in Texts

Carlo Strapparava and Gozde Ozbal

FBK-irst

strappa@fbk.eu, gozbalde@gmail.com

Abstract

Color affects every aspect of our lives. There has been a considerable interest in the psycholinguistic research area addressing the impact of color on emotions. In the experiments conducted by these studies, subjects have usually been asked to indicate their emotional responses to different colors. On the other side, sensing emotions in text by using NLP techniques has recently become a popular topic in computational linguistics. In this paper, we introduce a semantic similarity mechanism acquired from a large corpus of texts in order to check the similarity of colors and emotions. Then we investigate the correlation of our results with the outcomes of some psycholinguistic experiments. The conclusions are quite interesting. The correlation varies among different colors and globally we achieve very good results.

1 Introduction

In our daily speech, we frequently make use of colors in order to increase our expressiveness by invoking different emotions. For instance, we usually stress the *redness* of someone's face for the implication of his/her anger or excitement, or we use phrases including the color *black* to refer to a depressed mood. On the other hand, the color *pink* is mostly used with positive connotations such as 'to see everything in pink light', where the meaning is related to optimism and happiness.

Actually, the parts of the nervous system which are responsible for emotional arousal are affected

as soon as a color is perceived. Thus, the term *color emotion* has lately been used to represent the emotions arousing in human beings when they percept a color (Xin et al., 2004).

The correlation of color and emotion has been the focus of a lot of psycholinguistic studies so far. In the experiments conducted by these studies, subjects have usually been asked to indicate their emotional responses to different colors so that some general results stating which color arouses what kind of emotion could be obtained.

Sensing emotion, or in other words, affective sensing in text by using Natural Language Processing (NLP) techniques is recently a very popular topic in computational linguistics. There exist several studies focusing on the automatic identification of emotions in text with the help of both knowledge-based and corpus-based methods. Thus it is conceivable to explore whether state-of-the-art corpus analysis techniques can give support to psycholinguistic experiments.

Considering that psycholinguistic experiments are very costly since a lot of resources are required for both the setup and evaluation phases, employing a corpus-based approach for affective sensing could be much more efficient for all analysis to be held in the future, if this technique was proven to give reasonable results.

In this paper, we employ a semantic similarity mechanism automatically obtained in an unsupervised way from a large corpus of texts in order to check the similarity of color and emotion via computational analysis methods. We adopt the psycholinguistic experiments as references, with which we compare our results to find out if there is a correlation between the two approaches.

The paper is organized as follows. In Section 2, we introduce some related work focusing on the association of color and emotion only from a psycholinguistic point of view, since this topic has not been addressed by computational analysis techniques so far. In Section 3, we describe the methodology for implementing a similarity between colors and emotions, in particular how to represent an emotion in a latent semantic space. We present the evaluation of our approach and make a comparison with the results of psycholinguistic experiments in Section 4. Lastly, we report the conclusions and possible future work in Section 5.

2 Background

As we mentioned previously, there has been a considerable interest in the psycholinguistic research area addressing the impact of color on emotions.

(Zentner, 2001) mainly addressed the question of whether young children could show reliable color preferences. This study also tried to make a comparison with the results obtained with adults and older children. Subjects' color preferences were obtained by asking them to choose the one that they prefer among an array of colored cardboard rectangles. As an alternative way to represent musical information for providing feedback on emotion expression in music, (Bresin, 2005) suggested to use a graphical non-verbal representation of expressivity in music performance by exploiting color as an index of emotion. And for the purpose of determining which colors were most suitable for an emotional expression, some experiments were conducted, where subjects rated how well several colors and their nuances corresponded to music performances expressing different emotions. (Kaya, 2004) tried to investigate and discuss the associations between color and emotion by conducting experiments where college students were asked to indicate their emotional responses to principal, intermediate and achromatic colors, and the reasons for their choices.

There exist also some research investigating whether the color perception is related to the region of the subject. For instance, (Gao et al., 2007) analyzed and compared the color emotions of people from seven regions in a psychophysical

experiment, with an attempt to clarify the influences of culture and color perception attributes on color emotions. This study suggested that it might be possible to compose a color emotion space by using a restricted number of factors. As for (Soriano and Valenzuela, 2009), this study tried to find out why there was often a relationship between color and emotion words in different languages. In order to achieve this, a new experimental methodology called the Implicit Association Test (IAT) was used to explore the implicit connotative structure of the Peninsular Spanish color terms in terms of Osgood's universal semantic dimensions explained in (Adams and Osgood, 1973). The research conducted by (Xin et al., 2004) tried to compare the color emotional responses that were obtained by conducting visual experiments in different regions by using a set of color samples. A quantitative approach was used in this study in an attempt to compare the color emotions among these regions. (Madden et al., 2000) focused on the possible power of color for creating and sustaining brand and corporate images in international marketing. This study tried to explore the favorite colors of consumers from different countries, the meanings they associated with colors, and their color preferences for a logo.

The study that we will use for evaluating our results is a work which focused on the topic "emotional responses to color used in advertisement" (Alt, 2008). During the experiments, this study conducted a survey where the subjects were required to view an advertisement with a dominant color hue, and then select a specific emotional response and a positive/negative orientation related with this color. More than 150 subjects participated in this study, roughly equally partitioned in gender. There are two main reasons why we preferred to use this study for our evaluation procedure. Firstly, the presentation and organization of the results provide a good reference for our own experiments. In addition, it focusses on advertisement, which is one of the applicative fields we want to address in future work.

3 Methodology

Sensing emotions from text is an appealing task of natural language processing (Pang and Lee,

2008; Strapparava and Mihalcea, 2007): the automatic recognition of affective states is becoming a fundamental issue in several domains such as human-computer interaction or sentiment analysis for opinion mining. Indeed, a large amount of textual material has become available from the Web (e.g. blogs, forums, social networks), raising the attractiveness of empirical methods analysis on this field.

For representing the emotions, we exploit the methodology described in (Strapparava and Mihalcea, 2008). The idea underlying the method is the distinction between *direct* and *indirect* affective words.

For direct affective words (i.e. words that directly denote emotions), authors refer to the WORDNET AFFECT (Strapparava and Valitutti, 2004) lexicon, a freely available extension of the WORDNET database which employs some basic emotion labels (e.g. anger, disgust, fear, joy, sadness) to annotate WORDNET synsets.

For indirect affective words, a crucial aspect is building a mechanism to represent an emotion starting from affective lexical concepts and to introduce a semantic similarity among generic terms (and hence also words denoting colors) and these emotion representations.

Latent Semantic Analysis is used to acquire, in an unsupervised setting, a vector space from the British National Corpus¹. In LSA, term co-occurrences in a corpus are captured by means of a dimensionality reduction operated by a singular value decomposition on the term-by-document matrix representing the corpus. LSA has the advantage of allowing homogeneous representation and comparison of words, word sets (e.g. synsets), text fragments or entire documents. For representing word sets and texts by means of a LSA vector, it is possible to use a variation of the *pseudo-document* methodology described in (Berry, 1992). This variation takes into account also a *tf-idf* weighting schema. In practice, each document can be represented in the LSA space by summing up the normalized LSA vectors of all the

¹BNC is a very large (over 100 million words) corpus of modern English, both spoken and written (see <http://www.hcu.ox.ac.uk/bnc/>). Other more specific corpora could also be considered, to obtain a more domain oriented similarity.

terms contained in it. Therefore a set of words (and even all the words labeled with a particular emotion) can be represented in the LSA space, performing the pseudo-document technique on them.

As stated in (Strapparava and Mihalcea, 2008), each emotion can be represented in various ways in the LSA space. The particular one that we are employing is the ‘LSA Emotion Synset’ setting, which has proved to give the best results in terms of fine-grained emotion sensing. In this setting, the synsets of direct emotion words, taken from WORDNET AFFECT, are considered.

For our purposes, we compare the similarities among the representations of colors and emotions in the latent similarity space.

4 Experiments

For the experiments in this paper, we built an LSA vector space on the full BNC corpus using 400 dimensions. To compare our approach with the psycholinguistic experiments reported in (Alt, 2008), we represent the following emotions: anger, aversion/disgust, fear, happiness/joy, and sadness. And we consider the colors Blue, Red, Green, Orange, Purple, Yellow. Table 1 reports the rankings of emotions according to colors from (Alt, 2008).

Color	Ranking of Emotions				
	Anger	Aversion/ Disgust	Fear	Joy	Sadness
Blue	5	2	4	1	3
Red	1	4	2	3	5
Green	5	2	3	1	4
Orange	4	2	3	1	5
Purple	5	2	4	1	3
Yellow	5	2	4	1	3

Table 1: Emotions ranked by colors from psycholinguistic experiments

In Table 2 we report our results of ranking emotions with respect to colors using the similarity mechanism described in the previous section. To evaluate our results with respect to the psycholinguistic reference, we use Spearman correlation coefficient. The resulting correlation between two approaches for each color is reported in Table 3.

We can observe that the global correlation is rather good (0.75). In particular, it is very high

Color	Ranking Emotions using Similarity with Color				
	Anger	Aversion/ Disgust	Fear	Joy	Sadness
Blue	4	2	3	1	5
Red	4	3	2	1	5
Green	4	2	3	1	5
Orange	4	2	3	1	5
Purple	5	2	3	1	4
Yellow	4	2	3	1	5

Table 2: Emotions ranked by similarity with colors

Color	Correlation
Blue	0.7
Red	0.3
Green	0.9
Orange	1.0
Purple	0.9
Yellow	0.7
Total	0.75

Table 3: Correlation

for the colors Orange, Green and Purple, which implies that the use of language for these colors is quite in accordance with psycholinguistic results. The results are good for Blue and Yellow as well, while the correlation is not so high for Red. This could suggest that Red is a quite ambiguous color with respect to emotions.

5 Conclusions

There are emotional and symbolic associations with different colors. This is also revealed in our daily use of language, as we frequently make references to colors for increasing our expressiveness by invoking different emotions. While most of the research conducted so far with the aim of analyzing the relationship between color and emotion was based on psycholinguistic experiments, the goal of this study was exploring this association by employing a corpus-based approach for affective sensing.

In order to show that our approach was providing reasonable results, we adopted one of the existing psycholinguistic experiments as a reference. Following that adoption, we made a comparison between the results of this research and our own technique. Since we have observed that these two results were highly correlated as we expected, we would like to explore further this direction. Cer-

tainly different cultures can play a role for variant emotional responses (Adams and Osgood, 1973). Thus, as a next step, we are planning to investigate how the perception of color by human beings varies in different languages by again conducting a computational analysis with NLP techniques. Employing this approach could be very useful and efficient for the design of applications related to the fields of multimedia, automatic advertisement production, marketing and education (e.g. e-learning environments)

In addition, based on our exploration about the color perception of emotions from a corpus-based point of view, we suggest that “visual” information regarding objects and events could be extracted from large amounts of text, using the same kind of techniques proposed in the present paper. This information can be easily exploited for creation of dictionaries or used in dynamic visualization of text such as kinetic typography (Strapparava et al., 2007). As a concrete example, our approach can be extended to discover the association of colors not only with emotions, but also with indirect affective words in various languages. We believe that the discovery of this kind of relationship would allow us to automatically build colorful dictionaries, which could substantially help users with both interpretation and memorization processes.

References

- Adams, F. M. and C. E. Osgood. 1973. A cross-cultural study of the affective meanings of colour. *Journal of cross-cultural psychology*, 4:135—156.
- Alt, M. 2008. Emotional responses to color associated with an advertisement. Master’s thesis, Graduate College of Bowling Green State University, Bowling Green, Ohio.
- Berry, M. 1992. Large-scale sparse singular value computations. *International Journal of Supercomputer Applications*, 6(1):13–49.
- Bresin, R. 2005. What is the color of that music performance? In *Proceedings of the International Computer Music Conference (ICMA 2005)*, pages 367–370.
- Gao, X.P., J.H. Xin, T. Sato, A. Hansuebsai, M. Scalzo, K. Kajiwara, S. Guan, J. Valldeperas, M. Lis Jose,

- and M. Billger. 2007. Analysis of cross-cultural color emotion. *Color research and application*, 32(223—229).
- Kaya, N. 2004. Relationship between color and emotion: a study of college students. *College Student Journal*, pages 396–405.
- Madden, T. J., K. Hewett, and S. Roth Martin. 2000. Managing images in different cultures: A cross-national study of color meanings and preferences. *Journal of International Marketing*, 8(4):90–107.
- Ortony, A., G. L. Clore, and M. A. Foss. 1987. The psychological foundations of the affective lexicon. *Journal of Personality and Social Psychology*, 53:751–766.
- Pang, B. and L. Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1–135.
- Soriano, C. and J. Valenzuela. 2009. Emotion and colour across languages: implicit associations in spanish colour terms. *Social Science Information*, 48:421–445, September.
- Strapparava, C. and R. Mihalcea. 2007. SemEval-2007 task 14: Affective Text. In *Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval 2007)*, pages 70–74, Prague, June.
- Strapparava, C. and R. Mihalcea. 2008. Learning to identify emotions in text. In *SAC '08: Proceedings of the 2008 ACM symposium on Applied computing*, pages 1556–1560, New York, NY, USA. ACM.
- Strapparava, C. and A. Valitutti. 2004. WordNet-Affect: an affective extension of WordNet. In *Proceedings of LREC*, volume 4, pages 1083–1086.
- Strapparava, C., A. Valitutti, and O. Stock. 2007. Dances with words. In *Proc. of the 20th International Joint Conference on Artificial Intelligence (IJCAI-07)*, Hyderabad, India, January.
- Xin, J.H., K.M. Cheng, G. Taylor, T. Sato, and A. Hansuebsai. 2004. A cross-regional comparison of colour emotions. part I. quantitative analysis. *Color Research and Application*, 29:451–457.
- Zentner, M. R. 2001. Preferences for colors and color-emotion combinations in early childhood. *Developmental Science*, 4(4):389–398.