

IMAGE DIVERSITY ANALYSIS: CONTEXT, OPINION AND BIAS

P. Zontone, G. Boato, F. G. B. De Natale¹
A. De Rosa, M. Barni, A. Piva²
J. S. Hare, D. Dupplaw, P. H. Lewis³,

¹ Dep. of Information Engineering and Computer Science, University of Trento, Trento ITALY

² CNIT - National Inter-University Consortium for Telecommunications, Firenze/Siena ITALY

³ School of Electronics and Computer Science, University of Southampton, Southampton, UK

Abstract. The diffusion of new Internet and web technologies has increased the distribution of different digital content, such as text, sounds, images and videos. In this paper we focus on images and their role in the analysis of diversity. We consider diversity as a concept that takes into account the wide variety of information sources, and their differences in perspective and viewpoint. We describe a number of different dimensions of diversity; in particular, we analyze the dimensions related to image searches and context analysis, emotions conveyed by images and opinion mining, and bias analysis.

1 Introduction

With the advent of digital media, the number of images in the Web has rapidly increased and consequently the role of images in the communication process has gained importance. It is well known that an image can capture the attention of a viewer more than a long sentence. Perhaps the most powerful and meaningful way to inform, educate and persuade an individual is through the combination of memorable visual messages with text [1]. A single image may not always convey precise information or detailed data on a given subject, but an image can often transmit, in a more effective and immediate way, a message or an emotion. The use of visual data joined to textual data powerfully enriches the communication process that the writer is performing. In other instances, photography is a means for the faithful and true reproduction of real events and photographic images are used for documenting facts. In this scenario, it is worth mentioning two aspects: on one hand, photographers by taking pictures choose their own way of reporting an event (as the writers do). On the other hand, pictures may be manipulated before their use, thus conveying different information with respect to their original intent; therefore, the value of photography as a record of events must be established carefully.

To summarize, images can have three main roles within a communication process. Pictures can be used to: i) attract the attention of observers: a picture can be included in a document for attracting attention and making the document more appealing; ii) convey opinions and emotional messages: an image can be used for conveying an emotional message with a positive or negative implication; iii) convey information for documenting a given claim: images can be used to reproduce and document a claim.

Diversity plays an important role on the Internet and in all scenarios characterized by a large amount of information input from different sources. The information derived from multimedia content is the result of a clear diversity in cultural backgrounds, religious beliefs, political beliefs, ideologies and temporal contexts, and has an evident effect on opinions and bias of every person using such content. We will consider different dimensions of diversity, and the effect of image diversity on opinions and bias.

From the previous considerations it is clear that one of the dimensions of diversity on images is related to the intent or role of images within a communication process. Considering the impact of this type of diversity on bias, the use of images for conveying a message or for illustrating a textual message has a strong potential for bias due to the subtle message that can be conveyed.

In the following, we introduce the concept of diversity and present its image-related dimensions. We show how the results of image searches and context analysis might be analyzed for diversity (Section 2.1). Then, we provide an overview of the research activity in the area of opinion mining and sentiment analysis (Section 2.2). Finally, we describe bias analysis methods that allow the detection of bias in images (Section 2.3).

2 Diversity analysis

Let us define *diversity* as the co-existence of contradictory opinions and/or statements (typically with some being non-factual or referring to opposing beliefs/opinions) [2]. There are several forms and aspects of diversity to be considered:

- the existence of opinions with different polarity¹ about the same entity, e.g., at different times;
- diversity of themes, speakers, arguments, opinions, claims and ideas /frames;
- diversity of norms, values, behavior patterns, and mentalities;
- diversity in terms of geographical (local, regional, national, international, global focus of information), social (between individuals, between and within groups), and systemic (organizational and societal) aspects in media content;
- static (at one point in time) and dynamic (long-term) diversity;
- internal diversity (within one source) and external diversity (between sources).

Regarding dimensions of diversity that can be distinguished in images, we can list a set of dimensions, which are also applicable to text, that are: diversity of sources (e.g., suppliers in commercial search); diversity of resources (e.g., images, text); diversity of topic; diversity of viewpoint; diversity of genre (e.g., blogs, news, comments); diversity of language; geographical diversity; and temporal diversity. In addition, other dimensions specifically for images include:

- author/holder (person or professional agency who took the picture);
- time (date and time the picture was taken);
- location (where it was taken);
- source (where it was published, e.g., web site, blog, forum, PDF document);

¹ The polarity of an opinion is the degree to which a statement is positive, negative or neutral.

- source producing the picture (if the picture is computer generated or natural, if it comes from a digital camera or a scanner);
- intent (it is used to attract the attention of observers, to convey emotional messages, to give information for documenting a given claim);
- sentiment/opinion (positive, negative or neutral);
- context (characteristics of the text surrounding the picture, e.g., background of author, considered aspect, theme of the text);
- subject (words that describe what the picture shows and that can be linked to the same-similar terms contained in the surrounding text);
- time (date and time the picture was taken, night/day, summer/winter);
- style (words describing the style of the photos, e.g., photorealistic, pictorial);
- pure visual diversity (how visually similar or dissimilar images are).

Some values for these dimensions can be directly extracted from the EXIF information in the picture, which may have been inserted automatically (by the digital camera) or manually (by the photographer). If EXIF tags are unavailable, some features can be derived using image retrieval techniques [3], forensic techniques [4], and algorithms for automatically annotating images with high-level semantic concepts [5]. An example for extracted values along these dimensions can be seen in Fig. 2. This is a clear example of temporal diversity. These pictures have been extracted from a PDF document entitled ‘Global Warming’s Increasingly Visible Impacts’². For these pictures no EXIF information is available and so the features reported below are the ones that could be derived using the algorithms introduced above. Another example is shown in Fig. 3 where the subject of global warming is used in a very different way from the Italian design company DIESEL during its advertising campaign of 2007: the picture is a composition of computer generated data and photography, with the intent of attracting attention (a common intent in advertisement images) and incorporating a glamorous style.



Fig. 2: Diversity dimensions: *source*: PDF document, *context*: climate change, *topic/theme*: mountain, lake, rocks, etc.

² Environmental Defense Fund:
http://www.edf.org/documents/4891_GlobalWarmingImpacts.pdf



Fig. 3: Diversity dimensions: *author:* Terry Richardson, *source:* web site, *source producing the picture:* digital camera + computer generated data, *sentiment:* positive, *content:* buildings, water, woman, man, etc.

In the following subsections we will focus on some diversity dimensions, analyzing the state-of-the-art and possible research directions.

2.1 Diversity in image search and context analysis

Images can play several roles in the analysis of diversity. They might be used along with text to try to distribute documents along a diversity axis where the documents are primarily text based and the images play secondary roles both in the context of the document and in the analysis of their diversity. However, in some searches, image retrieval may be the main goal and in this section we show how the results of image searches might be analyzed for diversity.

Diversity in image search is usually considered as a problem of result diversification. Image search engines on the web, since they are based on exploiting textual information associated with an image, often do not care about the diversification of final visual results. Instead, a user's information need is often better satisfied when the result set for a particular query shows many different aspects of that query; this is especially important when the query is poorly specified or ambiguous [6, 7].

The diversification of search in image search engines is a relatively new area of research. In terms of image search, one particular way of increasing diversity is to ensure duplicate, or near-duplicate images in the retrieved set are hidden from the user [8].

We have been considering how to make use of semantic web technologies to help increase diversity of search results. Using the Yahoo BOSS (Build your Own Search Service) API, we have developed a tool that is capable of providing image search with different axes of diversity. The tool requires a user to input a query in the form of a subject (i.e., “David Beckham”), context ([optional] i.e. “football”), and axis of diversity (i.e., “football clubs”). Currently the axis must be specified as a DBpedia resource URI (i.e., “<http://dbpedia.org/ontology/clubs>”), however that constraint will be relaxed in future versions. The search engine works by using DBpedia to infer a list of topics along the diversity axis that are related to the subject. These topics (both the English name of the topic, and synonyms are considered) are then combined with the subject and context to generate a (potentially large) number of queries that can be fed to BOSS and structured into results. The results are presented as columns corresponding to the particular topics discovered during the semantic inference.

Context analysis is also considered for the study of diversity [9, 10]. It involves investigating the relevant information behind the content in order to better understand the context in which it was created. In fact, from a diversity point of view, we may wish, for example, to identify the location of events referred to in documents and if it is not explicit, related documents or contextual information may give the information necessary to find the location. Considering another example, we may wish to classify the documents in terms of the location of the writer as views may vary geographically, and although the writer's geographical location may not be explicit in the article, secondary searches or contextual information analysis (such as a semantic web search) may provide this information. The same is true for many other dimensions of diversity, such as time and general political affiliation. This could be a way not only of deriving opinions and sorting by diversity, but also a way of determining possible bias in documents.

2.2 Diversity on opinions and emotions conveyed by images

As described previously, an important role of images in the communication process is to convey opinions and emotional messages. With the growing availability of images and opinion-rich resources, such as online review sites and web blogs, the area of opinion mining and sentiment analysis has recently enjoyed a huge burst of research activity [11]. The activity in this area deals with the computational treatment of opinion, sentiment, and subjectivity in text and images. In particular, research on opinion mining that refers to opinions and sentiments expressed in images is still at the primary stage. The key problem is to select meaningful features that have a close relationship with human emotions and to convert them into numerical features. Some features (e.g., color, hue, luminance, saturation etc.) have been proposed but their effectiveness has not yet been evaluated. Emotional semantic image retrieval is a new and promising research direction in this field. Emotional semantics refer to the highest level of abstract semantics, i.e., the semantics that describe intensity and type of feelings, moods, emotions evoked in humans when they are viewing images [12]. One of the first emotional image retrieval systems was designed by Colombo et al. [13]. They proposed an innovative method to obtain a high-level representation of art images, which allowed the derivation of emotional semantics such as action, relaxation, joy and uneasiness. Since then, other research approaches and emotion-based retrieval systems have been proposed. In [14] a novel scheme to automatically annotate the image emotional semantics and realize emotional image retrieval using semantic words is described. In [15] the authors present an emotion categorization system, trained by ground truth from psychological studies and applied to a collection of masterpieces. In [16] only one of the aspects of aesthetic appeal is instead analyzed. The authors consider harmony, i.e., the pleasing or congruent arrangement of parts producing internal calm or tranquility. They conducted a series of experiments to identify what low level features could predict harmony in an image. However, emotional semantic image retrieval research is still at its primary stage because emotion is a subjective characteristic, i.e., it is strongly linked to the concept of human personality, and because it is difficult to find relations between the features and emotions.

From our point of view, it is crucial to develop a system that allows the categorization of pictures in terms of distinct emotions, taking into account the subjective

characteristics of emotions (the same image can lead to different emotions based on the cultural background of the viewer) and the differences in the levels of intensities of emotions (e.g., happiness, joyfulness). It is also important to find out what are the features to be extracted from an emotional perspective that best represent emotional semantics.

2.3 Bias on images

Let us define bias as a correlation between the polarity of an opinion and the context of the opinion holder. Focusing on images, we believe that in order to understand how the use of a given image within a given context can have influence on bias, it is crucial to know the history of the image itself. In particular, important historical aspects include the type of device used for producing the digital content, and, whether and what kind of tampering the image or its sub-parts suffered. For instance, discovering the semantic information within an image derived from a photomontage may highlight how the exploitation of a particular image in a communication process aims to polarize opinions, and may provide evidence that a biased view is being projected.

Recently, image forensics has been largely proposed as a valid technological means for ensuring the credibility of digital images, by both extracting knowledge about the origin of the content and detecting the application of a wide variety of manipulations [4]. Image forensics is based on the idea that inherent traces (like digital fingerprints) are left behind in a digital media during both the creation phase and any other successive processes [17]. By resorting only to analyzed data, digital forensic techniques can be seen as 'forensic blocks' taking as input an image and providing as output intrinsic information carried out by the digital asset, which permits better evaluation, understanding and validation of pictures used in the communication process.

Regarding the information on the content origin, the aim of a forensic block is to identify the source that produced the picture, e.g. the forensic block can determine whether the picture is computer generated or natural [18, 19], or whether the picture comes from a digital camera or a scanner [20]. The exploitation of such information can be used to validate a picture as an accurate and trustworthy representation of reality. An example of this is the case of computer-generated images.

Regarding the detection of a wide variety of manipulations, different forensic blocks are able to distinguish different processing operations, for example:

- re-sampling operation: when geometric transformations are applied (e.g. rotation, scaling) a re-sampling of the origin image to a new sampling grid is comprised [21, 22];
- double JPEG compression: when creating a digital forgery, it is often necessary to resave the modified image, so often the tampered image suffers a double JPEG compression [23];
- copy-move forgery: a part of the image is copied and pasted on another part of the same image [24].

Such forensic blocks can also be applied block-wise in order to spatially localize specific characteristics that could be different from one block to another. If we are studying some features that should be coherent in the overall image (e.g. source, JPEG compression, etc.), inconsistency of such features infers that some processing

has been locally applied to the content. The most common example of this is the creation of photomontages that are usually considered as a cut-and-paste composite of fragments coming from different images. This functionality could be very useful for understanding what semantic information has been altered.

A new way of exploiting image forensic technologies would be through their application to groups of images instead of single images, with the aim to discover dependencies between different images, used in different places, representing similar or equal contents, thus constructing a graph that describes picture relationships. By focusing on two images, the idea is to understand if one image comes from the other and the processing which possibly produced such a transformation. Knowing how a set of images are related to each other could allow the clustering of images sharing the same root image. In this way, we could discover that several images regarding an event have been actually produced from a limited set of source images, thus permitting isolation of the original information. In other situations, knowing how a few source images have evolved into a large set of derived pictures could allow us to reconstruct how the usage of the information contained in the original images has evolved in time and space. For instance, this could permit us to identify how these images have been used by groups of people with different opinions about the original event.

3 Conclusions

This paper gives an overview of the role of images in the analysis of diversity. We have considered how the results of image searches might be analyzed for diversity, and how context analysis can be used to better understand the context in which some information is created. Opinion mining that has recently attracted interest from different research communities has also been introduced. Finally, some methods to investigate the impact of diversity on bias have been presented.

Acknowledgements

The authors wish to thank the European Union, which supported this work under the Seventh Framework project LivingKnowledge (IST-FP7-231126).

References

1. Lester, P.M., Visual Communication Images with Messages, Fourth Edition, 2006, <http://commfaculty.fullerton.edu/lester/>.
2. Living Knowledge, <http://livingknowledge-project.eu/>.
3. Datta, R., Joshi, D., Li, J., Wang, J.Z., Image Retrieval: Ideas, influences, and trends of the new age. ACM Comput. Surv., 40(2):1-60, 2008.
4. Digital Forensics. IEEE Signal Processing Magazine, 26(2), 2009.

5. Tsai, C.-F., Hung, C., Automatically annotating images with keywords: A review of image annotation systems. *Recent Patents on Computer Science*, 1:55–68, 2008.
6. Tian, S. K., Gao, Y., Huang, T., Diversifying the image retrieval results. In *proc. ACM Multimedia '06*. Santa Barbara, CA, USA. Oct 23-27 2006. pp 707-710, 2006.
7. Chen, H., and Karger, D. R., Less is more: probabilistic models for retrieving fewer relevant documents. In *proc. ACM SIGIR '06*. Seattle, Washington, USA. Aug 06-11 2006. pp 429-436, 2006.
8. Arni, T., Clough, P., Sanderson, M., and Grubinger, M. (2008) Overview of the ImageCLEFphoto 2008 Photographic Retrieval Task. Retrieved 18-06-2009. http://www.clef-campaign.org/2008/working_notes/ImageCLEFphoto2008-final.pdf
9. Smith, A., Carr, L., Hall, W. (2005) An Opportunistic Approach to Adding Value to a Photograph Collection. *Proceedings of the 4th International Semantic Web Conference*. November 2005, Galway Ireland.
10. Tuffield, M., Harris, S., Dupplaw, D P., Chakravarthy, A., Brewster, C., Gibbins, N., O'Hara, K., Ciravegna, F., Sleeman, D., Wilks, Y., Shadbolt, N R. (2006) Image Annotation with Photocopain. *Proceedings of the First International Workshop on Semantic Web Annotations for Multimedia*. WWW2006, May 2006, Edinburgh.
11. Pang, B., Lee, L., *Opinion Mining and Sentiment Analysis*. *Foundations and Trends in Information Retrieval*, Vol. 2, 2008.
12. Wang, W., He, Q., A Survey on Emotional Semantic Image Retrieval. *IEEE International Conference on Image Processing*, San Diego, California, U.S.A., October 12–15, 2008.
13. Colombo, C., Bimbo, A. D., Pala, P., *Semantics in Visual Information Retrieval*. *IEEE Multimedia*, Vol.6, No.3, pp.38-53, 1999.
14. Wei-ning, W., Ying-lin, Y., Sheng-ming, J., *Image Retrieval by Emotional Semantics: A Study of Emotional Space and Feature Extraction*. *IEEE International Conference on Systems, Man, and Cybernetics*, Taipei, Taiwan, October 8-11, 2006.
15. Yanulevskaya, V., van Gemert, J.C., Roth, K., Herbold, A.K., Sebe, N., Geusebroek, J.M., *Emotional Valence Categorization Using Holistics Image Features*. *IEEE International Conference on Image Processing*, San Diego, California, U.S.A., October 12–15, 2008.
16. Fedorovskaya, E., Neustaedter, C., Hao, W., *Image Harmony For Consumer Images*. *IEEE International Conference on Image Processing*, San Diego, California, U.S.A., October 12–15, 2008.
17. Swaminathan, A., Wu, M., Liu, K.J.R., *Digital Image Forensics via Intrinsic Fingerprints*. *IEEE Transactions on Information Forensics and Security*, 3(1), pp. 101-117, 2008.
18. Lyu, S., *Natural Image Statistics for Digital Image Forensics*. Ph.D. Dissertation, Department of Computer Science, Dartmouth College, 2005.
19. Ng, T.T., *Statistical and Geometric Methods for Passive-blind Image Forensics*. PhD Dissertation, Graduate School of Arts and Sciences, Columbia University, 2007.
20. McKay, C., Swaminathan, A., Gou, H., Wu, M., *Image acquisition forensics: forensic analysis to identify imaging source*. *IEEE International Conference on Acoustic, Speech, and Signal Processing*, Las Vegas, NV, March 2008.
21. Mahdian, B., Saic, S., *Blind authentication using periodic properties of interpolation*. *IEEE Transactions on Information Forensics and Security*, 3(3), pp. 529-538, 2008.
22. Kirchner, M., *Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue*. In *Proc. 10th ACM Workshop on Multimedia and Security*, Oxford, UK, pp. 11-20, September 2008.
23. Farid, H., *Exposing digital forgeries from JPEG ghosts*. *IEEE Transactions on Information Forensics and Security*, 4(1), pp. 154-160. 2009.
24. Bayram, S., Sencar, H.T., Memon, N., *An Efficient and Robust Method For Detecting Copy-Move Forgery*. In *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2009*, 2009.