Towards Multimedia Opinion Mining*

Giulia Boato[1], Valentina Conotter[1], Francesco G. B. De Natale[1],
Claudio Fontanari[2]

[1] Dept. of Information Engineering and Computer Science, University of Trento
[2] Dept. of Mathematics, University of Trento
Via Sommarive, 14, 38123 Trento, Italy
boato@disi.unitn.it, conotter@disi.unitn.it, denatale@ing.unitn.it, fontanar@science.unitn.it

Abstract. Both opinion mining and multimedia retrieval are active research areas with challenging applications, but as far as we know the present vision paper is the first attempt to integrate them into multimedia opinion mining. Here we address the specific case of satirical comments in politics by exploiting the presence of photomontage to infer a tendentially negative opinion. In order to do so, we introduce a novel digital forensics technique allowing source identification from a single image.

Keywords: Opinion mining, cross-media analysis, digital forensics, photomontage detection, sensor pattern noise.

1 Introduction

The availability of more and more devices that allow users to generate new multimedia content, by capturing their own experience in images and videos, mixing it with digital material collected from the web, and finally sharing it with other users, claims for a new paradigm of information extraction from digital data. Indeed, media search based on textual annotations seems to be intrinsically inadequate to access the richness of visual information; on the other hand, content-based image retrieval suffers from the so-called semantic gap between low level features and high-level semantics. A cross-media approach, exploiting both text and visual content, helps to bridge such a gap and provides more effective tools to information retrieval.
We point out that relevant information concerns not only facts, but also opinions. Extracting opinions from text documents is a very challenging but well-established discipline, known as opinion mining or sentiment analysis. However, ambiguity of text (especially in satirical or ironic comments) makes automated opinion extraction into a really non-trivial issue. Following the cross-media philosophy, we believe that images accompanying text provide valuable side information that should be exploited to accomplish such a task. To the best of our knowledge, this is the first contribution

* This work is partially funded by the European Commission in the framework of the Living Knowledge project.

towards an opinion mining based not only on text, but also on more general multimedia content.
Indeed, the idea of analyzing facts, opinion, and bias in large multidimensional data sets is the main goal of the European project Living Knowledge [1]. In such a framework, a first application of digital forensics techniques to investigate opinions conveyed by images is presented in [2].
Here instead we address a specific case study, namely, satirical comments in politics. A satirical text conveys a tendentially negative opinion about its subject, but it can be ambiguous enough to confuse an automated classification. However, luckily enough satirical comments about politicians appearing on the web are quite often accompanied by photomontages making their ironic purpose more easily detectable.
Our key idea is to apply digital forensics tools detecting image manipulations to classify as negative the opinion about a politician extracted from a text surrounding a photomontage. In order to do so, we need to distinguish (pieces of) pictures taken by different cameras. As we shall see, currently available tools based on sensor noise require either the devices which took the pictures or at least multiple images taken by each camera, which is clearly an infeasible assumption in the web context. We are able to overcome this point thanks to the recent work [3] about noise estimation from a single image.
The structure of the paper is the following: In Section 2 we describe the current state of the art on opinion mining and we outline the proposed cross-media approach. In Section 3 we report digital forensics tools currently available for image manipulation detection. The proposed method for photomontage detection and multimedia opinion mining from satirical comments on politics is detailed in Section 4. Finally, Section 5 reports concluding remarks and open problems to be addressed in future works on this new research topic.

2 Opinion Mining

In the last few years, opinion mining has attracted interest from different research areas including computational linguistics, artificial intelligence and computer science (see [4] for a comprehensive and updated survey). Given the increasing diffusion and popularity of user-generated content (e.g., blogs), opinion mining provides the opportunity to scan this information and to gain insights into the public's or an individual's perception of facts and products, and it is therefore appealing for marketers and analysts.
Existing methods focus on sentiment analysis on text, by determining whether a positive or negative sentiment is conveyed by a single words, a complete phrase, or a document. On one hand, the semantic orientation of single words is defined by looking at co-occurrence patterns with reference words (e.g., “excellent” and “poor”) [5-7] or by exploiting additional data such as word paraphrases [8]. On the other hand, semantic orientation of sentences and documents can be extracted by suitably modifying machine learning techniques [9-10]. Specific methods for product review are presented in the literature: the authors of [11-12] propose to associate the extracted opinion with particular characteristics of the product, while just a summary of the reviews is determined in [13-14], filtering out untruthful reviews that try to

manipulate the customer [15]. The problem of analyzing opinion on widespread news has been addressed only recently in [16], by exploiting comments and information reported on blogs.
We believe that state of the art techniques could strongly be improved by exploiting the integration of the semantic content of textual and non-textual data, thus allowing more accurate opinion extraction by detecting the characteristics of visual data that may alter the perception of a user and their relevant impact. In order to do so, we underline the need of opinion mining methods from multimedia data able to support current tools working on text. By linking associated text and images we may come to a cross-media characterisation and analysis. The selection and use of images for conveying a message or illustrating a textual message clearly have a strong potential for biasing due to the subtle message that can be conveyed by images.
In this paper we propose a first attempt focusing on satirical comments in politics, where multimedia data analysis can help disambiguating opinion extracted from text. In particular, satirical texts about politicians convey negative sentiments towards them, although irony may require an antiphrastic use of positive words therefore making an automatic text analysis very difficult. On the other hand, this kind of comments are commonly supported by photomontages, which can represent an anchor for multimedia opinion mining. Indeed, by detecting photomontages it will be possible to achieve a more accurate automatic analysis of cross-media comments about a politician. We propose to exploit the currently available techniques to detect image manipulations (described in details in the next Section) and to classify as negative the opinion about a politician extracted from a text surrounding a photomontage (following the algorithm presented in Section 4).

3 Digital Forensics

From a traditional point of view, a photograph is a trusty and close representation of a real scene. Notwithstanding, this is no longer true for digital images, nowadays widely used in several fields such as news, sports and information reporting, because of the ease of manipulation allowed by sophisticated photo editors (e.g., Photoshop). Doctored images cannot be admitted as a legal evidence, thus claming for advanced tools able to link the digital image to a specific camera and therefore demonstrating its integrity. Moreover, modified data may influence people opinions and even alter their attidudes in response to the represented event [17-18]. As a consequence, it is more and more important to be able to automatically verify the fidelity and authenticity of digital images in order to guarantee their truthfulness.
Digital watermarking [19] has been proposed as a valuable means to prove the content ownership and authenticity and to track copyright violations. Generally, a watermark (an imperceptible digital code) is embedded into a multimedia content and it is assumed to be modified whenever a tampering occurs. Authenticity can thus be demonstrated by comparing the extracted watermark with the original inserted code. The major drawback of this approach is that it requires the watermark to be embedded at the time of recording, thus limiting its application to specially equipped cameras. According to [20], digital watermarking is said to be an active forensics approach, in contrast to passive techniques which work in absence of any watermark or special

hardware.
Recently, the scientific community focused its attention on passive forensics techniques, whose aims can be primarily divided in three categories [21]:

- Image forgery detection, to prove that a a-posteriori manipulation has been applied to an image, e.g., moving or replacing an object within an image. Different tools, such as binary similarity measures, wavelet coefficient statistics, quality metrics, phase characteristic of the bicoherence spectrum, resampling, color filter array interpolation, and geometric optics can be used to this aim [22-27].
- Discrimination between synthetic and real images [28-29].
- Image source identification. All methods are based on the assumption that digital pictures taken by the same device are overlaid by a specific pattern, that is a unique and intrinsic fingerprint of the acquisition device. Each manufacturer selects specific hardware components for a given device model, thus different patterns can be present in the image, depending on the brand and on the model. These intrinsic characteristics allow linking images to a specific device for forensic purposes. Many techniques have been proposed in the literature to describe this unique pattern, each one analyzing different processing steps of the digital camera pipeline (i.e. demosaiking, CFA interpolation, lens radial distortion) [30-34]. The most promising approach belonging to this class of forensics techniques is based on the analysis of sensor imperfections. Two types of noise have been considered in forensics analysis. The first type is introduced by array defects and includes hot pixels, dead pixels, pixel traps and cluster defects [35]. The major drawback of these methods is that defect pixels are not very reliable since many cameras include in their hardware post-processing operations able to compensate such a noise. The second type of noise is called Patter Noise and indicates "any spatial pattern that does not change significantly from image to image" [36]. This reference pattern is known given the camera model that took the photo or is obtained by averaging the noise residuals of a set of available images, all taken from a specific camera [37-38].

4 Proposed Method

The main technical issue we have to face in our application is photomontage detection. In [39] sensor noise extraction is exploited to detect image forgeries, by computing correlation between the reference pattern and pattern extracted by local regions of the images. The main limitation of forensics techniques based on sensor noise remains the basic assumption of the availability of the device which took the image or, alternatively, of other images taken by the same device [39]. This may be a strong constraint, especially in those applications where only one image is available and its origin and integrity needs to be verified.

In our opinion, a possible solution could come from a recent work in a field different from digital forensics. Indeed, Liu et al. [3] perform noise estimation starting from a single image. This approach is based on a simple noise model of a CCD camera, namely, $I=f(L+n_s+n_c)+n_q$, where I is the observed image brightness, $f(.)$ is the camera response function (CRF), and $n_{s,}$ $n_{c,}$ n_q take into account different types of noise introduced in the image acquisition process. After a segmentation process (K-Means Clustering), each segment of image I is transformed by using the inverse of CRF (available at www.cs.columbia.edu/CAVE) to obtain a corresponding L in the irradiance plane. Such an L is then added with synthesized noises n_s and n_c (since n_q can be neglected) and direct CRF is applied again to return into the brightness domain. Next, a real camera demosaicing algorithm is reproduced. Through this process, a noisy image I_N can be obtained by adding to the original image I the synthesized CCD noise. At this point a noise level function (NLF) can be estimated, which is essentially a relation describing the noise level as a function of the image brightness. An example of estimated NLF curves is given in Fig. 1.

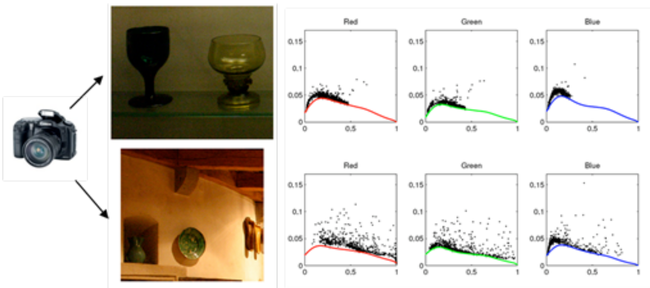

Fig. 1 Estimated NLF curves, one for each channel RGB, taken from [3]

Experiments show that the proposed method is efficient and is able to extract reliable noise from images. This algorithm has been successfully applied to adaptive bilateral denoising and canny edge detection, reporting very promising results.

In the context of image forensics, the most interesting aspect of this work is that authors claim that different images from the same camera give the same estimated NLF. Validation to this claim has been performed with success, as reported in Fig. 1. It is evident that two different images, taken by the same camera, results in a very similar NLF. Starting from this point, we could evince that different images taken by different cameras exhibit different NLF curves, thus resulting in a promising forensics technique to detect photomontage. Furthermore, with the described method it seems to be possible to reveal differences in images taken with the same camera but in different moments, thus leading to a complete forensics framework able to reveal any splicing.

Hence we stress two main advantages with respect to the technique presented in [38]. First of all, only one image is needed and the constraint to have available either the camera or a set of images is overcome. Second, photomontages deriving by the splice of two or more images taken by the same camera, but taken in different moments, can be revealed. On the contrary, forgeries detection based on sensor noise analysis only reveals if parts of the image are linked to a different camera, loosing its efficiency when splicing comes from images taken by the same camera.

The main idea of our contribution is to apply forensics techniques to the automatic analysis of cross-media comments about politicians. Since we assume that the opinion extracted from a satiric text surrounding a photomontage is negative, we detect photomontage of politicians in order to claim satirical purpose of both the image and the surrounding text.

The proposed algorithm is the following:

1. Exploit GOOGLE image search engine (based on textual annotations) to construct a database of websites about famous persons. In this case study we focus on images of the Italian Prime Minister Silvio Berlusconi.
2. A face detector (for instance [40]) is applied to isolate the face of the person, based on the assumption that typically a satirical photomontage is constructed by splicing the face with another image as in Fig. 2(a).
3. The noise estimation proposed in [3] is applied firstly inside and then outside the region extracted in step 2.
4. Based on the claim in [3] that different images taken by the same camera give the same estimated NLF and images taken from different cameras will exhibit different NLF, we check for image integrity as follows:
   - If the calculated NLFs are coherent, no photomontage has been applied;
   - Otherwise, if the calculated NLFs are not coherent, the image derives from a splicing operation of two different images. Thus, the non authenticity of the considered picture can be claimed.
5. According to a reasonable assumption, detection of a photomontage implies a tendentially negative opinion.
   We stress that a subjective analysis could be misleading, as demonstrated by the non tampered picture in Fig. 2(b).

Fig. 2 Examples of (a) photomontage and (b) authentic photograph

5 Conclusions

In this vision paper we have moved the first steps towards the integration of multimedia data in opinion mining. Here we have focused on the specific case of satirical comments in politics by exploiting the negative connotation implied by the presence of a photomontage. However, we believe that a cross-media approach may have a relevant impact on opinion mining by taking advantage of visual information also in a more general context. Indeed, sentiments induced by images strongly influence the opinion conveyed to users. From this perspective, textual analysis should be supported by a suitable multimedia understanding. In our specific application, we have reduced opinion extraction to photomontage detection and we have introduced a novel digital forensics tool based on noise estimation from a single image. This idea turns out to be innovative with respect to the state of the art and may find more general applications in the field of digital forensics.

References

[1] http://livingknowledge-project.eu/

[2] F. Uccheddu, A. De Rosa, A. Piva, and M .Barni, "Investigating Image Dependencies Through Image Forensics", GTTI 2009, Parma, Italy, June 2009.

[3] C. Liu, R. Szeliski, S.B. Kang, C.L. Zitnick and W.T. Freeman, "Automatic estimation and removal of noise from a single image", *IEEE Transactions on Pattern analysis and machine intelligence*, vol. 30, pp. 299-314, 2008.

[4] B. Pang, and L. Lee, "Opinion mining and sentiment analysis", *Foundations and Trends in Information Retrieval,* vol. 2, pp. 1-135, 2008.

[5] X. Ding, B. Liu and P. S. Yu, "A holistic lexicon-based approach to opinion mining", *Proc. of the ACM International Conference on Web Search and Web Data Mining,* 2008.

[6] V. Hatzivassiloglou and K. R. McKeown, "Predicting the semantic orientation of adjectives", *Proc. of the Annual Meeting of the Association for Computational Linguistics,* 1997.

[7] P. D. Turney and M. L. Littman, "Measuring praise and criticism: Inference of semantic orientation from association", *ACM Trans. Inf. Syst.,* vol. 21(4), pp. 315–346, 2003.

[8] A. Esuli and F. Sebastiani, "Determining the semantic orientation of terms through gloss classification", *Proc. of ACM International Conference on Information and Knowledge Management,* 2005.

[9] K. Dave, S. Lawrence and D. M. Pennock, "Mining the peanut gallery: opinion extraction and semantic classification of product reviews", *Proc. of WWW*, 2003.

[10] H. Yu and V. Hatzivassiloglou, "Towards answering opinion questions: separating facts from opinions and identifying the polarity of opinion sentences", *Proc. of the Conference on Empirical Methods in Natural Language Processing,* 2003.

[11] B. Liu, M. Hu and J. Cheng, "Opinion observer: analyzing and comparing opinions on the web", *Proc. of WWW,* 2005.

[12] A. M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews", *Proc. of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing,* 2005.

[13] P. Beineke, T. Hastie, C. Manning and S. Vaithyanathan, "An exploration of sentiment summarization", *Proc. of AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications,* 2003.

[14] M. Hu and B. Liu, "Mining and summarizing customer reviews", *Proc. of the ACM International Conference on Knowledge Discovery and Data Mining,* 2004.

[15] X. Ding, B. Liu and P. S. Yu, "A holistic lexicon-based approach to opinion mining", *Proc. of the ACM International Conference on Web Search and Web Data Mining,* 2008.

[16] M. Gamon, S. Basu, D. Belenko, D. Fisher, M. Hurst and A. C. König, "BLEWS - Using Blogs to Provide Context for News Articles", *Proc. of the International Conference on Weblogs and Social Media,* 2008.

[17] D. L. M. Sacchi, F. Agnoli, and E. F. Loftus, "Changing history: Doctored photographs affect memory for past public events," *Appl. Cognit. Psychol.*, vol. 21(8), pp. 1005-1022, 2007.

[18] H. Farid, "Digital Doctoring: Can we trust photographs?", in *Deception: Methods, Motives, Contexts and Consequences*, Stanford, CA: Stanford Univ. Press, 2007.

[19] S. J. Lee and S.H Jung, "A survey of watermarking techniques Applied to Multimedia", *Proc. of ISIE,* 2001.

[20] T.-T. Ng, S.-F. Chang, C.-Y. Lin, and Q. Sun, "Passive-blind image forensics,"

Multimedia Security Technologies for Digital Rights, Eds. New York: Elsevier, 2006.

[21] H. T. Sencar and N. Memon, “Overview of state-of-the-art in digital image forensics,” *Statistical Science and Interdisciplinary Research*. Singapore: World Scientific Press, 2008.

[22] I. Avcibas, S. Bayram, N. Memon, B. Sankur and M. Ramkumar, “A Classifier Design for Detecting Image Manipulations”, *Proc. of IEEE ICIP*, 2004.

[23] T. Ng, S. -F. Chang and Q. Sun, “Blind Detection of Photomontage Using Higher Order Statistics”, *Proc. of ISCAS,* 2004.

[24] S. Bayram, I. Avcibas, B. Sankur and N. Memon, “Image Manipulation Detection”, *Journal of Electronic Imaging*, vol. 15(4), 2006.

[25] A.C. Popescu, H. Farid, “Exposing Digital Forgeries by Detecting Traces of Re-sampling”, *IEEE Transactions on Signal Processing*, vol. 53(2), pp. 758-767, 2005.

[26] A.C. Popescu, H. Farid, “Exposing Digital Forgeries in Color Filter Array Interpolated Images”, *IEEE Transactions on Signal Processing*, vol. 53(10), pp. 3948-3959, 2005.

[27] M.K. Johnson, H. Farid, “Exposing Digital Forgeries in Complex Lighting Environments*”, IEEE Transactions on Infromation Forensics and Security*, vol. 2(3), pp. 450-461, 2007.

[28] S. Lyu and H. Farid, “How Realistic is Photorealistic?”, *IEEE Trans. on Signal Processing*, vol. 53(2), pp. 845-850, 2005.

[29] T.-T Ng, S. -F. Chang, J. Hsu, L. Xie, M. -P. Tsui, “Physics-Motivated Features for Distinguishing Photographic Images and Computer Graphics,” *ACM Multimedia,* 2005.

[30] M. Kharrazi, H.T. Sencar and N. Memon, “Blind source camera identification”, *Proc. of ICIP,* 2004.

[31] O. Celiktutan, I. Avcibas, B. Sankur and N. Memon, “Source cell-phone identification”, *Proc. of ADCOM*, 2005.

[32] S. Bayram, H.T. Sencar, N. Memon and I. Avcibas, “Source camera identification based on CFA Interpolation”, *Proc. of ICIP*, 2005.

[33] Y. Longand Y. Huang, “Image based source camera identification using demosaicing”, *Proc. of MMSP*, 2006.

[34] K. S. Choi, E. Y. Lam and K.K.Y. Wong, “Source camera identification using footprints from lens aberration”, *Proc. of SPIE*, *Electronic Imaging San Jose, CA* 2006.

[35] Z. J. Geradts, J. Bijhold, M. Kieft, K. Kurusawa, K. Kuroki and N. Saitoh, “Methods for Identification of Images Acquired with Digital Cameras”, *Proc. of SPIE Electronic Imaging San Jose, CA*, 2001.

[36] N. Khanna, G.T.C. Chiu, J.P. Allebach and E.J. Delp, “Scanner identification with extension to forgery detection”, *Proc. of SPIE Electronic Imaging San Jose, CA*, 2008.

[37] J. Lukas, J. Fridrich and M. Goljan, “Digital camera identification from sensor pattern noise”, *IEEE Transactions on Information Security and Forensics*, vol. 1(2), pp. 205-214, 2006.

[38] M. Chen, J. Fridrich, M. Goljan and J. Lukas, “Determining image origin and integrity using sensor noise”, *IEEE Transactions on Information Security and Forensics*, vol. 3(1), pp. 74-90, 2008.

[39] J. Lukas, J. Fridrich and M. Goljan, “Detecting digital image forgeries using sensor pattern noise”, *Proc. of SPIE Electronic Imaging San Jose, CA*, 2006.

[40] Jianxin Wu, S.C. Brubaker, M.D. Mullin, J.M. Rehg, “Fast Asymmetric Learning for Cascade Face Detection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 30(3), pp. 369–382, 2008.