

Assessing the Impact of Image Manipulation on Users' Perceptions of Deception

Valentina Conotter^a, Duc-Tien Dang-Nguyen^a, Giulia Boato^a, María Menéndez^a
Martha Larson^b

^aDISI - University of Trento, Italy

^bDMIR - Delft University of Technology

ABSTRACT

Generally, we expect images to be an honest reflection of reality. However, this assumption is undermined by the new image editing technology, which allows for easy manipulation and distortion of digital contents. Our understanding of the implications related to the use of a manipulated data is lagging behind. In this paper we propose to exploit crowdsourcing tools in order to analyze the impact of different types of manipulation on users' perceptions of deception. Our goal is to gain significant insights about how different types of manipulations impact users' perceptions and how the context in which a modified image is used influences human perception of image deceptiveness. Through an extensive crowdsourcing user study, we aim at demonstrating that the problem of predicting user-perceived deception can be approached by automatic methods. Analysis of results collected on Amazon Mechanical Turk platform highlights how deception is related to the level of modifications applied to the image and to the context within modified pictures are used. To the best of our knowledge, this work represents the first attempt to address to the image editing debate using automatic approaches and going beyond investigation of forgeries.

1. INTRODUCTION

The lifecycle of media contents is becoming more and more complex, with people increasingly accessing, sharing, modifying, adapting and re-using the digital information. In this scenario, the trustworthiness of the information conveyed by media networks is becoming one of the key challenges of the information society, strongly affecting the success and penetration of future multimedia applications. Indeed, digital information can be easily manipulated, thus generating various possible misuses, which may be intentional or accidental, malicious or honest, legal or illegal, manifest or hidden. Our understanding of the implications related to the use of a manipulated data is lagging behind, as new image editing technology becomes more sophisticated and gains wide-spread use. Indeed, modified data may influence people's opinions, for example, altering their attitudes to past events depicted in images.¹ As a consequence, the image editing debate is growing in importance, and it has a significant impact on how we communicate information both in the public and the private sphere.

Our work represents an initial investigation analyzing how different types of image manipulation impact users' perceptions and how the context in which a modified image is used impacts human perception of image deceptiveness. We are motivated by the observation that expectations of images truthfulness are in some sense related to context. In Figure 1, we provide an example of a manipulated image and its corresponding original version, illustrating the intuition behind this work: the use of this modified image as a representation of a business summit may be considered inappropriate in a newspaper, but it may be totally suitable as an advertisement. The extreme cases are clear: expectations of truthfulness of images are radically different if the image is hanging in an art gallery or being used as evidence in a court case. Verifying the trustworthiness of multimedia data thus becomes a key pre-requisite, also for content recommendation and the assessment of the reputation of content generators. Multimedia forensics techniques are particularly relevant in this field, as they deal with the recovery of information that can be directly used to authenticate and estimate the trustworthiness of digital multimedia contents, i.e., if it is genuine, authentic, non manipulated, non duplicated.² A critical shortcoming of existing

Further author information: Dr. Conotter: E-mail: conotter@disi.unitn.it, Telephone: +39 0461 285268



(a) Original Image



(b) Manipulated Image

Figure 1. An example of image manipulation which may be considered unfair in some contexts (like journalism) but acceptable in others (like advertisement).

image forensic approaches is that they attempt to assess the authenticity of an image independently from its contextual information, without being able to retrieve any knowledge about its suitability for the context of use.

Very little effort have been devoted so far to understanding the specific characteristics of images and contexts that influence users perceptions of deception and how strong and consistence these perceptions are. This paper reports the results of a study that provide evidence that different types of manipulation and different contexts have different influences on users perceptions of the deceptiveness of modified images. This evidence is collected with a study that applies crowdsourcing methodologies³ to elicit the opinion of a large number of people about different types of manipulations, different image contents, and different reasons for motivations, and different contexts of use. To the best of our knowledge, this work represents a first approach to gain more insights on the relation between the level of modification in images and its perception on users. Currently very little related work is present in the literature, mainly involving small scale user studies focused on the impact of a specific editing techniques.^{4,5}

The structure of the paper is the following: Section 2 briefly reviews the state of the art in the field of crowdsourcing and human perception of image deceptiveness; in Section 3 the conducted crowdsourcing study is described; Section 4 shows the obtained results; finally in Section 5 conclusion and future work are drawn.

2. RELATED WORK

2.1 Crowdsourcing

Crowdsourcing can be defined as “the act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call”.³ Examples of crowdsourcing platforms are Amazon Mechanical Turk* (AMT) and CrowdFlower† (CF). These platforms are usually referred to as microtask markets, where users (i.e., workers) can find small tasks published by micro-employers (i.e., requesters) in return for a micro-payment (i.e., small award). On AMT, a microtask is referred to us a Human Intelligence Task (HIT). A more general view of crowdsourcing refers to human computation, which is defined as problems solved by human participation that is either directed by or intended to replace a computational process.⁶ The authors of⁶ emphasize that social computing is distinct from human computation in that it involves facilitation of natural human behavior that is relatively natural, but happens to be mediated by technology. Human computation is not “natural” behavior in that the reason for the behavior is created or driven by the computational system. Crowdsourcing tasks range from objective to very subjective tasks. Objective tasks are “mechanical” tasks for which is possible to distinguish correct from incorrect answers. Such tasks include image annotation. Existing research demonstrates that labels provided by the crowd are as reliable as those provided by experts.⁷ The use of crowdsourcing for “subjective” tasks has been shown to be

*www.mturk.com

†www.crowdflower.com

effective in the areas of human decision-making and human behaviour^{8,9}. Recently, the multimedia community has begun to use crowdsourcing to elicit information from users about their reactions and interpretations of multimedia^{10,11} as well as the motivations for using multimedia systems.¹² Subjective crowdsourcing tasks can yield rich information for multimedia access and authoring, however, they are difficult to design, because there is no way to check correct answers. Instead, the tasks must gauge user seriousness and engagement.

To the best of our knowledge, this work is the first attempt to use this technique to gain more insight on the relation between modification on images and the perception of user in terms of differences between original and modified image, intention behind different kind of modification, possible usage of fake data, etc.

2.2 Human perceptions of image deceptiveness

The possibility of image manipulation has existed since the invention of the analogue camera. However, the rise of digital photography has created the possibility for a wider range of manipulations and has also made the technology necessary to manipulate images available to a broader public. These developments have increased concerns about image manipulation both among professionals and with the broader public.¹³ In some contexts, users of images have a precisely formulated notion of what constitutes image deception. Journals sometimes codify the definition of deceptive images into a set of explicitly formulated standards and rules for image integrity, e.g., *Nature*[‡]. In the area of visual journalism, professional associations also spell out what is understood to be image deception, e.g., the code of ethics of the US-based National Press Photographers Association[§]. In other domains, the attitudes of the users of images are less clearly defined. Currently, there is an ongoing trend towards the blurring of boundaries between domains of news and advertising and also an increase in the speed in which news media is produced and the number of sources that it integrates.¹³ This trend contributes to making the question of whether users will perceive a particular image manipulation in a particular context as deceptive more complex. The public outcry that has arisen when high-profile cases of the discovery of manipulation of images by the press support the view that image manipulation is and will remain an issue of vital concern. This paper provides a contribution towards understanding the context-sensitive nature of that concern.

As mentioned above, previous work on the subject of the impact of image modification on human perceptions has been limited. In much work, such as,⁵ the impact of image modification is presumed, and the question addressed is whether humans can perceive modifications. Other work looks at the impact of modification, but does not directly address perceptions of deception. For example, the importance of the impact of image modifications on opinions and interpretation of image meaning was highlighted by.¹⁴ In,⁴ the focus is on the impact of image modifications on human feeling, and not specifically on feelings of having been deceived.

3. PROPOSED APPROACH

We aim at studying how people perceive the manipulations in the created images. In order to do that and collect users' opinions about the images present in our dataset, we carefully designed a Human Intelligent Task (HIT) and published it on Amazon Mechanical Turk. In the following we will describe this in details.

3.1 Dataset collection

We collected/created 275 pairs of original/manipulated images. The applied manipulations reflect possible forgeries to which we may be exposed in our daily-life, ranging from common and popular image processing operators to more technically complex and time-consuming image editing. The images in the dataset fall into four categories based on the level of the applied manipulation: Similar (25 pairs): images are visually similar, with differences only in resolution or format, Enhanced (84 pairs): images are slightly modified on global low-level attributes (e.g., contrast enhancement, light balancing), Retouched (86 pairs): images are not only different on global characteristics but also altered locally, Doctored (80 pairs): images are manipulated via advance editing techniques (e.g., splicing) changing the semantics of the original picture and/or creating something different from the original data. Some examples of these pairs are illustrated in Figure 2. Notice that the number of images in different categories are almost balanced, except for category "similar". A smaller number of similar images was

[‡]www.nature.com/authors/policies/image.html

[§]https://nppa.org/code_of_ethics

included in the study as we expect them to be less informative to our goal of understanding users’ perception of deception than enhanced, retouched, and doctored images.

3.2 Task design

The task started with a brief introduction describing the context of the study and presenting an example of an original/manipulated image pair. The main body of the task was divided into three parts. In the first part, people were asked to rate the level of difference between the pair of original/manipulated images in a 7-point Likert scale.¹⁵ They were also asked to imagine possible motivations of the editor for transforming the original image into the provided manipulated version. The second part dealt with potential uses of the manipulated image envisioned by the workers. People were asked to select a role and to answer the next set of questions playing that role. The proposed roles were photographer, blogger, journalist, advertiser and “other” and were exemplified as cartoon characters (see Figure 3) so to invite people interpreting the task as a game to be played. In particular, the main reason for asking workers to play a role was to encourage them letting go their inhibitions, making them feel free to play as in a game so to possibly chose a non-honest use of an image, which might be otherwise unlikely. The last part of the task dealt with the perceived level of honesty of the manipulation. Workers were asked to stop playing the role and answer a set of questions on differences, in terms of honesty, between original and manipulated images. As an example, the asked question was *“In your opinion how honest is the use you chose of the manipulated image?”*, and multiple choice answers were provided, such as *“The image does not distort reality”*, *“The image makes things prettier than they are”*, *“The image can be considered to ‘tell a white lie’”*, *“The image can be considered to seriously distort reality”*, and *“The image can be considered to directly manipulate people’s feelings or opinions”*. The order of the answers indicated, in our understanding, a descending level of honesty. Moreover, workers were asked whether the use of the original image would have been more honest than the use of the manipulated image. Finally, a set of questions invited people to provide demographic and profile information such as age, gender, country, education level, and familiarity with manipulated images.

4. MAIN RESULTS

In this section we provide the results of our analysis based on the answers provided by the workers in the described crowd-sourcing experiment. The study was run in AMT. The data was split into 5 batches, each containing 55 pairs of images, and each pair of images was assigned to 10 workers. Rejected records were resent to other workers to complete. In total, 553 workers participated in the validation study, performing 2934 records. After quality control, 2503 over 2934 received records ($\sim 85\%$) were accepted. In average workers spent 6 minutes 52 seconds on each image pair over all five batches and each batch was completed in around three days.

An important consideration is that workers became very engaged with the task: 76% chose *“yes”* for the question *“Would you like to perform more HITS of this type?”* and we collected also numerous positive feedbacks, e.g., *“It is really an interesting subject and I welcome such initiatives in the future.”*. In the following we detail the analysis on quantitative and qualitative (free text) answers.

4.1 Perceived difference between original and manipulated images

The workers were asked to rate the difference they perceive between each of the 275 pairs of original/manipulated images, by using a 7-point Likert scale (1 indicates that images are perceived as very similar while 7 stands for very different images).

For every image pairs present in each group of our dataset, we calculated the average value of all the collected evaluations. In order to deal with outliers, we discarded those ratings whose difference from the mean exceeds the threshold $T > 2$. As a result, out of 2503 collected ratings for all the 275 image pairs, we rejected 199 evaluations as considered outliers. We then recomputed the total mean of users ratings and plotted them in Figure 4(a). The built image groups for modification in our dataset track the decreasing similarity of the manipulated image with respect to the original image, measured here in terms of the well-known objective image quality metric Structural Similarity (SSIM) index.¹⁶ This full-reference metric has been proved to successfully measure the image quality consistently with subjective judgments, as opposite to traditional methods like Peak-Signal-to-Noise-Ratio (PSNR). Moreover, (SSIM) index brings image quality measurement from a pixel-based analysis (PSNR) to a structural information-based stage. The choice of SSIM metric was primarily due to its widespread



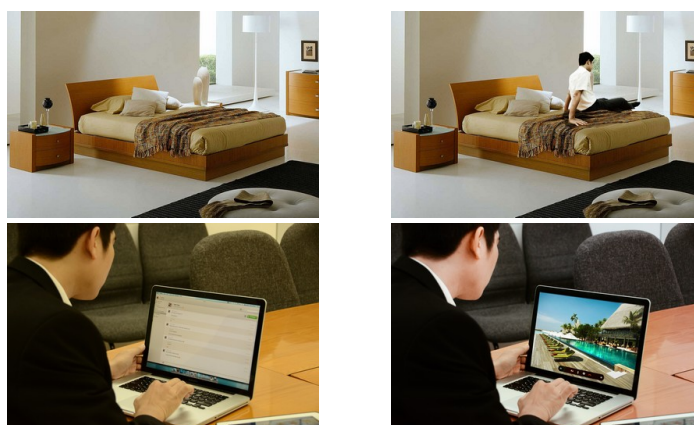
(a) Similar. These images are not different.



(b) Enhanced. The modified images were sharpened and enhanced in contrast.



(c) Retouched. The modified images were cropped, sharpened, enhanced in contrast, and some color effects were applied.



(d) Doctored. The modified images were tampered with advanced editing techniques (e.g., splicing in these examples).

Figure 2. Examples of collected pairs of images for the different types of manipulation: for each pair, on the left the original image and on the right the modified one.

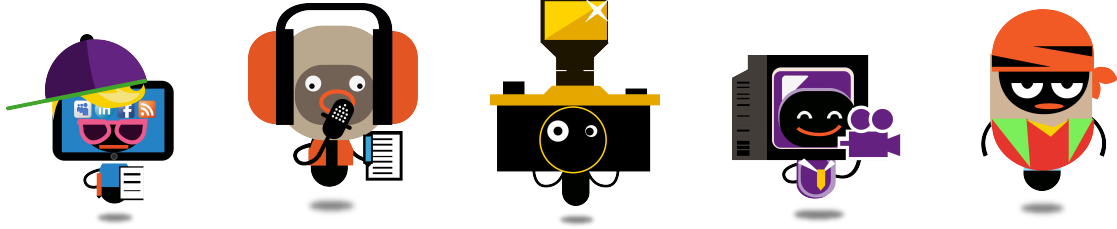


Figure 3. The five roles used in the final HIT: (from left to right) blogger, journalist, photographer, advertiser, and “other”.

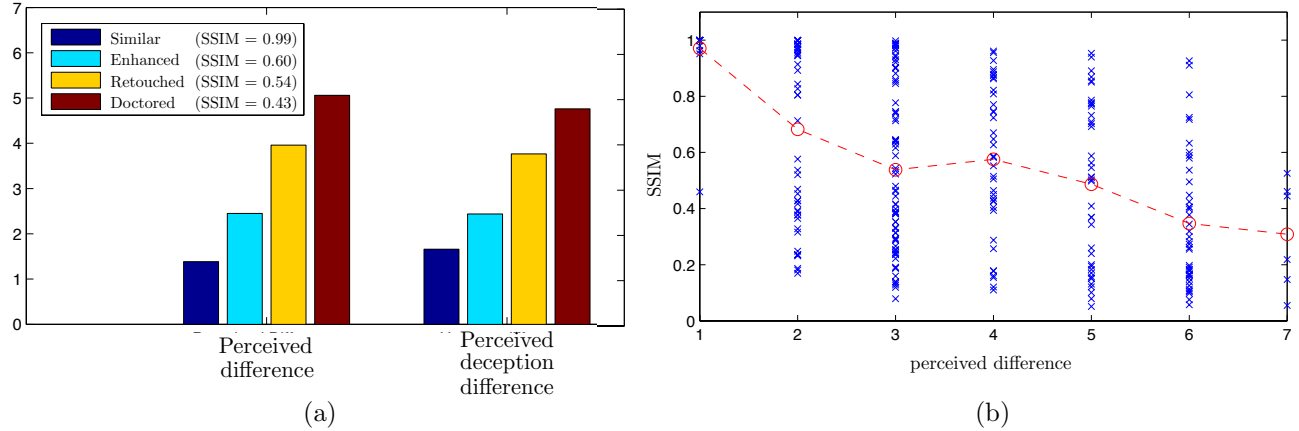


Figure 4. Given the images present in the dataset, divided in 4 manipulation categories, panel (a) shows the average difference between the original and the manipulated image perceived by the user, together with the perceived deception difference between the two versions of the images. Panel (b) shows how the difference perceived by users tracks SSIM quality metric.

use as one of the most accepted metric for image quality assessment. However, we conducted the here presented analysis also with different metrics (i.e., PSNR, Weighted-PSNR¹⁷), leading to similar results as SSIM. Due to space constraints, we are here presenting results only with one metric, i.e., SSIM. As reported in Figure 4(a), we can see, for example, that images falling in the category “similar” present, on average, a higher value of SSIM when compared to the other categories. As a first result, we deduce that the perceived difference tracks the increasing level of manipulation (Figure 4(a)).

Moreover, to further demonstrate that the adopted quality metric reflects users’ perception of difference between the image pairs, we plot the perceived difference with respect to the calculated SSIM. In particular, Figure 4(b) shows, for each considered image pairs (blue-crosses), the calculated average users’ evaluation (x-axis) with respect to the SSIM between considered original and modified image (y-axis). The red line represents the average SSIM of the images falling within the same ranking. This plot further confirms that the perceived difference tracks also the objective quality of the considered images, measured with a traditional quality metric.

4.2 Perceived deception of the manipulation

In order to get insights about the perceived deception, users were asked to evaluate how honest they think would be to use the manipulated version of the image. Answers go from 1 indicating that “*The image does not distort reality*”, increasing up to 5 where “*The image can be considered to directly manipulate people’s feelings or opinions*”. Moreover, users were asked to rank the difference in terms of perceived deception between the original and manipulated image, using a 7-point Likert scale. We verified that both the perceived level of deception and the perceived deception difference between un-modified and modified images track the level of manipulation (Figure 4(a)).

In general, most of the answers perceived the chosen use of the manipulated image as honest. 26% of the answers indicated *“the image does not distort reality”* and 36% indicated *“the image makes things prettier than they are”*. In order to investigate pairs which were perceived as less honest, we analyzed the answers to image manipulations *“it can be considered to directly manipulate people’s feelings or opinions”* or *“the image can be considered to seriously distort reality”*. In this case, the categories are evenly distributed. When looking at the type of modification, most of the answers related to image pairs involving doctored images (49%), followed by retouched (31%), enhanced (15%) and similar (5%) images.

On average, annotators ranked the difference in terms of honesty between member of the pairs to be more precise than between the pairs with 3.49 points (SD=1.95) in a 7-point Likert scale. An ANOVA was run to analyse the effect of perceived honesty of the manipulated pictures on perceived difference between pairs in terms of honesty. Results indicate an effect of perceived honesty on perceived difference of honesty ($F(4, 2498) = 180.386$, $p < .001$, $\mu p2 = .22$). Post-hoc analysis indicates that the effect is due to a scale with significant intervals between pairs, starting from manipulated images annotated with *“the image does not distort reality”*, where the difference in honesty between the original and the manipulated is perceived as small, to *“the image makes things prettier than they are”*, *“the image can be considered to tell a harmless lie”*, *“the image can be considered to directly manipulate people’s feelings or opinions”* and *“the image can be considered to seriously distort reality”*, where the difference between original and manipulated was perceived as the largest. These results suggest that annotators perceived images which *“can be considered to seriously distort reality”* to differ more from the original picture, in terms of honesty, than images which *“can be considered to directly manipulate people’s feelings or opinions”*. There is a significant positive strong correlation between the general perceived difference and the perceived difference in terms of honesty between two pictures ($p(985) = .78$; $p < .01$).

4.3 Roles

Because we are interested in determining if context impacts users perception of the deceptiveness of modified images, we invited workers to select a role from a list (i.e. photographer, blogger, journalist, advertiser, and other) and to play that role while answering the questions about image deception. We expected that asking workers to play a role might encourage them to answer freely concerning possible non-honest use of an image, disregarding hesitation that they might feel if we were to ask their opinion directly.

Workers who chose to play the role of a photographer were 45% of the participants to the HITs, while the distribution among the other roles is almost uniform: 17% blogger, 12% journalist, 14% advertiser, 12% other.

The results of the study show that the choice of the photographer role has significant correlation with the perception of an image manipulation as being honest. Our observations provide significant insight onto the extent to which people understand the motivation that drives an editor to change images (the motivations were clearly different from one editor “role” to another).

ANOVA indicates an effect of the selected role on perceived difference between pairs of images in general ($F(4, 2498) = 27.069$, $p < .001$, $\mu p2 = .01$) and honesty terms ($F(4, 2498) = 22.679$, $p < .001$, $\mu p2 = .01$). Post-hoc analysis indicates that the effect is due to people who chose the role of photographer perceiving less difference between the pairs of pictures, in general and honesty terms, than those who chose to be a blogger, journalist, advertiser, or other. These results suggest that people select to be a photographer when the manipulated picture is perceived similar to the original.

4.4 Motivation behind manipulation

Considering the question *“Do you have an intuition about why the editor might have changed the images?”*, we manually analyzed all answers, grouped them into four different classes and removed unserious answers (i.e., not related to the question): (i) 12% reported that they did not perceive a difference between the images nor could they conceive what the difference might be; (ii) 29% commented about differences on the global appearance and/or basic attributes of pictures, e.g., *“He probably might have thought that more brightness would make the image more interesting.”*; (iii) 36% commented about differences on the image content and emotion, e.g., *“Changing the umbrella color from blue to red warms up the image and makes it seem more inviting.”*, *“Maybe to show that the girl is enjoying her holidays on a beach.”*; (iv) 24% commented about the purposes of the creator or on the effects on the observers, e.g., *“This image after manipulation is good for newspaper advertising.”*

Table 1. Distribution of targeted people on different roles.

	Type (i)	Type (ii)	Type (iii)	Type (iv)
Blogger	15%	16%	57%	12%
Journalist	19%	20%	59%	2%
Photographer	14%	12%	49%	25%
Advertiser	17%	15%	67%	1%
Other	24%	19%	43%	14%

Most of the workers commented about differences on the image content and how it influences emotion. However, we may evidence different kinds of motivations connected to chosen roles. The highest percentage of workers commenting about differences on the image content and emotion were those choosing to be a “journalist”, while workers choosing to be a “photographer” mostly discussed about differences on the global appearance and/or basic attributes of the picture. Workers choosing to be an “advertiser” were those who commented the most about the purposes of the creator or on the effects on the observers.

Considering the question “*describing the people or the types of people who will read or look at what you created*”, we manually grouped the answers into four different types: type (i) general targets, e.g., “*Any people.*”; type (ii) people related to the role of the worker, e.g., “*news readers, journalists*” if the role is “journalist”; type (iii) people who are interested in the content of the picture, e.g., “*People interested in looking for night views of city or buildings at night.*”; and type (iv) people who are acquainted with the workers, e.g., “*my friends who are connected with me on FB, Twitter*”. The results show that the targeted people are also connected to the chosen roles (see the full distribution in Table 1). Workers choosing to be a “journalist” selected very carefully what they were going to create with the modified image (we note that this stands in strong contrast to the professional codes that are actually followed by visual journalists, mentioned above). They also intended to attract the readers and engage people that could be somehow related to that kind of picture. Moreover, we noticed that 59% of the answers were strongly connected with the content of the picture. Workers choosing to be an “advertiser” described their purposes in advertising something connected with the content of the image (67%). Their target were somehow similar to the journalists one, but more precise on age and group of people. Workers choosing to be a “blogger” or “photographer” provided very wide and creative answers on what they are creating with the modified image. However, workers choosing to be a “photographer” focused deeply on technical details and points of view, often discussing about the techniques used to modify the picture. Their targets were mainly Photoshop users, friends, or sometimes tourists. These results indicate that people playing different roles (i.e., blogger, journalist, photographer, advertiser) have different motivations for image editing, based also on a different audience. These results suggest that context might play an important role in perception of deception in image editing.

5. CONCLUSIONS

Crowdsourcing is a promising approach to obtain important insights about how different types of manipulations impact users’ perceptions and how the context in which a modified image is used influences human perception of image deceptiveness. The presented work represent a first approach to deeply analyze this research field and to assess the influence on user perception of deception played by the main motivations that lead to manipulate the content. Our crowd sourcing study suggests that both the kind of modification and the image context are factors influencing how people perceive image manipulation. The underlying idea of this study is to provide a human dimension to the trust verification problem, making possible in future work to analyze the relationship between “quantitative” measures of media trustworthiness and the relevant “qualitative” impact on how they are perceived in user communities. In general, crowdsourcing opens many interesting possibilities for user studies on multimedia forensics since it can reach a large pool of workers with different cultural and social characteristics. Future work will pursue deeper insight on the issues that have emerged. Future work will also address possible limitations of this study, such as designing alternative tasks in order to control for possible bias due to question formulation.

Acknowledgments

The research leading to these results has received funding from the European Commission's 7th Framework Programme under grant agreements No. 287704 (CUBRIK).

REFERENCES

- [1] D. L. M. Sacchi, F. Agnoli, and E. F. Loftus, "Changing history: Doctored photographs affect memory for past public events," *Appl. Cognit. Psychol* (21), pp. 1005–1022, 2007.
- [2] E. Delp and M. Wu, "Digital forensics," *IEEE Signal Processing Magazine* **2**(26), pp. 14–15, 2009.
- [3] J. Howe, *The Rise of Crowdsourcing*, Wired, 2006.
- [4] P. Zontone, M. Carli, G. Boato, and F. De Natale, "Impact of contrast modification on human feeling: an objective and subjective assessment," in *IEEE ICIP*, pp. 1757–1760, 2010.
- [5] H. Farid and M. Bravo, "Perceptual discrimination of computer generated and photographic faces," *Digital Investigation* **8**, pp. 226–235, 2012.
- [6] A. Quinn and B. Bederson, "Human computation: a survey and taxonomy of a growing field," in *SIGCHI Conf. on Human Factors in Computing Systems*, 2011.
- [7] S. Nowak and S. Rüger, "How reliable are annotations via crowdsourcing: a study about inter-annotator agreement for multi-label image annotation," in *Int. Conf. on Multimedia Information Retrieval*, 2010.
- [8] G. Paolacci et al., "Running experiments on Amazon Mechanical Turk," *Judgment and Decision Making* **5**, August 2010.
- [9] W. Mason and S. Suri, "Conducting behavioral research on Amazon's Mechanical Turk," *Behavior Research Methods* **44**, March 2012.
- [10] M. Soleymani and M. Larson, "Crowdsourcing for affective annotation of video: Development of a viewer-reported boredom corpus," in *SIGIR Workshop on Crowdsourcing for Search Evaluation*, 2010.
- [11] R. Vlieghendhart, M. A. Larson, and J. Pouwelse, "Discovering user perceptions of semantic similarity in near-duplicate multimedia files," in *Workshop on Crowdsourcing Web Search*, 2012.
- [12] A. Hanjalic, C. Kofler, and M. Larson, "Intent and its discontents: the user at the wheel of the online video search engine," in *ACM Multimedia*, 2012.
- [13] L. Gross et al., *Image Ethics In The Digital Age*. University Of Minnesota Press, 2003.
- [14] L. Kennedy and S.-F. Chang, "Internet Image Archaeology: Automatically Tracing the Manipulation History of Photographs on the Web," in *ACM International Conference on Multimedia*, pp. 349–358, 2008.
- [15] R. Likert, "A technique for the measurement of attitudes," **4**, pp. 1–55, 1932.
- [16] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing* (13), pp. 600–612, 2004.
- [17] S. Voloshynovskiy, A. Herrigel, N. Baumgaertner, and T. Pun, "A stochastic approach to content adaptive digital image watermarking," *Proc. Int. Workshop on Information Hiding*, pp. 211–236, 1999.