IDENTIFICATION OF RECAPTURED PHOTOGRAPHS ON LCD SCREENS

Hong Cao and Alex C. Kot

School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

ABSTRACT

With advances in image display technology, recapturing good-quality images from the high-fidelity artificial scenery on a LCD screen becomes possible. Such image recapturing posts a security threat, which allows the forgery images to bypass the current forensic systems. In this paper, we first recapture some good-quality photos on different LCD screens by properly setting up the recapturing environment and tuning the controllable settings. In a perceptional study, we find that such finely recaptured images can hardly be identified by human eyes. To prevent the image recapturing attack, we propose a set of statistical features, which capture the common anomalies introduced in the camera recapturing process on LCD screens. With a probabilistic support vector machine classifier, comparison results show that our proposed features work very well, which outperform the conventional image forensic features in identification of the finely recaptured images.

Index Terms— image forensics, local binary pattern, LCD, loss of details, recaptured photos, texture

1. INTRODUCTION

To restore the trustworthiness of digital images, image forgery detection [10] has been intensively studied in recent years through detection of certain intrinsic image regularities or some common tampering anomalies. Frequently, the tell-tale cues useful for image forensics such as lens distortion, sensor noise pattern and statistics, demosaicing regularity and JPEG characteristics are directly associated with the image creation pipeline, where the light signals are converted into a digital image. Though some forensic methods can efficiently expose the direct tampering made on an image, most existing methods are unable to expose the indirect scenery forgery, where the scenery to be captured is artificially created. Though creating a physical scenery in general can be a very difficult and expensive task, with the aid of today's ubiquitous and high-fidelity display technology, generating a virtual scenery of reasonable fidelity is still relatively easy and such technology is potentially exploited to defeat the current image forensics systems.

As illustrated in Fig. 1, an image forger can first display the tampered images with a high-quality artificial display
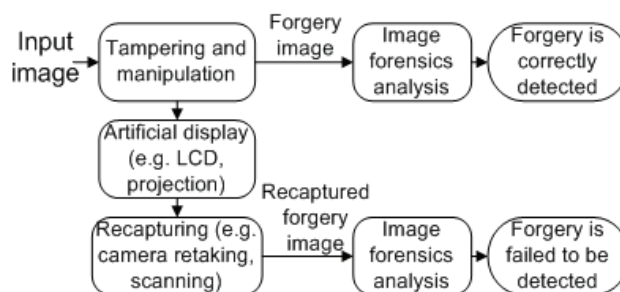

Fig. 1 Comparison of Image Forensics Results on the Direct Forgery Images and the Recaptured Forgery Images

e.g. through high-fidelity printing, liquid crystal display (LCD) or projection. Through proper set-up, the forger can recapture the artificially generated scenery and use the recaptured image to fool the image forensic system. It should be noted that during the image recapturing process, the tampering anomalies, e.g. splicing discontinuity and resampling artifacts, are automatically removed and the intrinsic image regularities which are originally disturbed due to the tampering operations are automatically restored. Moreover, creating such kind of forgery requires no specialty and can be implemented by novices of modern computer and photography technology. In a real-life example [1], a hunter recaptured a fake tiger scenery, which is made of a paper tiger poster, to prove the very presence of a commonly believed extinct tiger species. After being accepted by the local authority and published on Internet, these recaptured photos sparked large-scale controversy. To prevent such a security loophole, we consider identification of the recaptured images is an important task to facilitate the current image forensic system.

In previous research, the work in [2] initially brought up the issue of "rebroadcast attack" in image forensics. By formulating it as a binary classification problem, the author shows the printed-and-scanned photos can be accurately distinguished from the natural photos by using 72 wavelet statistical features and a simple linear discriminant classifier. In [4], a set of image geometry features are proposed to distinguish photos from computer graphics (CG). At the same time, the authors show experimentally that both the geometry features and wavelet features [2-3] can be used to efficiently classify recaptured CGs from natural photos. In [5], by assuming the artificial scenery is planar, e.g. a face image, the authors find the specular component measured on recaptured images can serve as a distinctive feature to

(a) a natural image

(b) a casually recaptured image

Fig. 2 Comparison of a Natural Image and its Casually Recaptured Version on a LCD Screen

Fig. 3 Our Image Recapturing Environment where the Tunable Settings to Achieve Better Quality of Recaptured Images are Highlighted

differentiate a recaptured photo and a natural photo. In this paper, we consider identification of finely recaptured photos on common LCD screens. Since image recapturing is commonly accompanied with some image quality losses, we first study the settings for recapturing good-quality images. A perceptional study is then designed to test the human ability in identification of our finely recaptured images. Based on our observation, we further propose several sets of image features, including texture features, loss-of-detail features and color features to identify the recaptured images from natural images.

This paper is organized as follows: Section 2 investigates on the environmental setup and the settings for recapturing good-quality images. Section 3 describes a survey on human's ability to identify the finely recaptured images. Section 4 proposes several types of image features for automatic identification of recaptured images and experimentally demonstrates their effectiveness. Finally, section 5 concludes this paper and indicates the future extension.

2. RECAPTURING GOOD-QUALITY IMAGES

Casually recapturing a scenery displayed on a LCD screen often lead to poor quality of recaptured images. As shown in Fig. 2, we can easily observe some obvious artifacts such as the texture patterns, loss of fine details and color degradation from the poor-quality example. Generally, such poor-quality recaptured photos are useless and they can be easily identified by human eyes. The question now becomes whether we can recapture reasonably good-quality images from the ubiquitous LCD screens in a common environment so that they can be used to fool human eyes. To perform such test, we have set up an image recapturing environment as illustrated in Fig. 3. It should be noted that such an environment contains a large number of controllable settings including the camera settings, the LCD settings and the environmental settings. These setting can be tuned to recapture good-quality images.
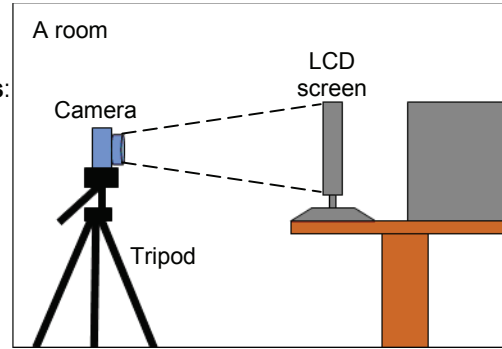
With 3 available digital still cameras (Canon Powershot A620, Olympus Mju 300 and Olympus E500 DSLR) and 3 common LCDs (Philips 19" 190B6CG, NEC 17" AccuSync and Acer 17" AL712), we experimentally determine the best settings for each of 9 camera-LCD combination by tuning all the controllable settings. In the tuning process, we found that by adjusting the camera's color temperature and LCD's RGB color ratios, brightness, etc., color of the recaptured images can be made close to their original images. By adjusting the camera's shutter speed, capturing mode, zooming and the distance between camera and LCD, the visible textured pattern can be largely eliminated. And by setting the LCD's resolution to its maximum, the amount of blurring (loss of details) can be minimized. Though with the best setting, we still can observe some obvious quality degradation if we compare the recaptured images with their corresponding original images, visual quality of these finely recaptured images are significantly better than the casually recaptured images.

After determining the best recapturing setting, we recapture 300 photos from LCD scenery for each camera-LCD combination, where the displayed contents are 300 natural photos. Out of these 300 natural photos, 100 are taken by the 3 available cameras, 100 are downloaded from *Flickr* website and the rest 100 are tampered photos where about 10% of the total image area has been altered. In such a way, for a total of 9 camera-LCD combinations, we form a finely recaptured dataset of 2700 photos.

3. HUMAN IDENTIFICATION OF FINELY RECAPTURED IMAGES

To find out how well human beings can identify the finely recaptured images, we perform the following survey in a trained scenario. Since most our survey participants have not seen a recaptured image before, we first explain about the recapture images on LCDs and provide each participant with two sample pairs of the original and the recaptured images as shown in Fig. 4. On a computer screen, the

Natural images　　　Recaptured images

Fig. 4 Two pairs of Training Samples of Nature and Recaptured Images

Fig. 5 Samples of 50 Images Used in Human Recaptured Image Identification Survey

participant can closely examine the image differences caused by the recapturing. After about 2 to 5 minutes' training and when the participant felt confident, we let the participant to browse through 50 selected photos as shown in Fig. 5, which contain 20 natural photos and 30 recaptured photos mixed in random order. The participants are allowed to take their own time to closely examine each image before they classify the image either into the "natural" category or the "recaptured" category. This survey is conducted on 30 participants, who are mostly university students and staffs. By comparing the participants' answers with the ground-truth answers, we find on average, the type-I error rate is 19.8%, i.e. natural images been misclassified as "recaptured" and the type-II error rate is 51.1%, i.e. the recaptured images been misclassified as "natural". The standard deviations of the type-I error rate and type-II error rate are respectively 12.8% and 18.3%. These large error rates suggest that common people are poor in differentiating the finely recaptured photos from natural photos. Therefore, these recaptured images can potentially fool both human eyes and the current image forensic system.

4. COMPUTER IDENTIFICATION OF FINELY RECAPTURED IMAGES

4.1. Forensic Features

To prevent the security loophole of the image recapturing attack, reliable automatic identification of the finely recaptured images on LCD screens is highly desirable. By formulating the problem as a binary classification task, we propose the following 3 types of features to capture the unique anomalies introduced by the recapturing process:

Local binary pattern (LBP): Texture patterns at fine scale are often easily observable on a poor-quality recaptured image. Formation of these textures is due to the aggregation of the regular tiny structures appearing on surface of a LCD, the unique polarity inverse driving pattern [11] and the periodic recharging of the CMOS capacitors, which affect brightness of the tiny LCD cells. Though these texture patterns are not obvious in our finely recaptured image, we consider complete elimination of the textures is very

difficult. To capture such texture anomalies, we compute multiple-scale LBP features [6]. The LBP features are a normalized occurrence histogram of some "uniform" patterns computed using the operator $LBP_{P,R}^{riu2}$ [6], where P is the dimension of the angular space and R determines the resolution. The computed LBP features are invariant to image rotation and contrast changes, which are highly desirable properties for our identification of recaptured images. Moreover, by varying R and P correspondingly, LBP features can be easily extended to multiple scales. By first converting an input color image into a gray image, we compute a total of 80 LBP features at multiple scales using the 4 operators: $LBP_{8,1}^{riu2}$, $LBP_{16,2}^{riu2}$, $LBP_{24,3}^{riu2}$ and $LBP_{24,4}^{riu2}$.

Multi-Scale Wavelet Statistics (MSWS): Loss of fine details is inevitably coupled with the image recapturing on LCD screens. With today's LCD technology, the display resolution of a common LCD can still hardly match with the sensor resolution of a common DSC. Considering that usually more information is lost at the fine scale than at the coarse scale, the loss of fine details can be characterized by measuring the amount of local image variations at multiple scales. To do this, we first perform N-level wavelet decomposition separately on the R, G and B channels using a standard Haar filter. Let {LL, HL, LH, HH}$_{cn}$ represent the nth–level decomposition of the cth color channel, where $1 \leq n \leq N$ and $c \in$ {R, G, B}. For each high-frequency band in {HL, LH, HH}$_{cn}$, we compute the mean and standard deviation of the absolute wavelet coefficients as our features. By setting N=3 and for 3 color channels, we derive a total of $3 \times 3 \times 3 \times 2 = 54$ such features.

Color Features (CF): Though majority of color artifacts can be eliminated with the best settings of the camera and LCD, the color of our finely recaptured images still looks different from their original images. Typically, the color of a recaptured image appears a bit washed out and more easily saturated. To capture these color anomalies, we compose a set of 21 color features including 3 average pixel values, 3 pairs correlations, 3 neighbor distribution centers of mass, 3 pairs energy ratios [7] from the RGB color space and 9 color moments computed from HSV color space [8].

1792

Table 1 Performance Comparison in Recaptured Image
Identification in Terms of Equal Error Rate (EER)

Features	Dimension	EER (%)
LBP	80	0.9
MSWS	54	1.1
CF	21	17.4
LBP+MSWS+CF	155	0.5
Wavelet Stats [3]	216	3.4

4.2. Identification Experiment

To test the effectiveness of the proposed features, we have set up an image dataset containing 2000 natural photos and 2700 finely recaptured photos described in section 2. Out of the natural photos, 300 photos are taken with the same 3 cameras used in image recapturing in section 2 and the remaining 1700 photos are taken with 9 other cameras from 5 different brands including Canon, Casio, Lumix, Nikon and Sony. For both classes, we randomly select 80% of the photos for training and the remaining for testing. Such random apportion is repeated for 5 times so that we have 5 different combinations of training and testing datasets. The various types of features are then computed from the cropped central block of 1024×1024 from each photo.

After feature extraction, we train a probabilistic support vector machine (PSVM) classifier using the LIBSVM tools by following the guild in [9]. By averaging the test results from the 5 different apportions, the various feature sets are compared in Table 1 in terms of equal error rates (EERs). Based on the results, we can see both the LBP and the MSWS features perform excellently by giving low EERs of about 1.0%. This suggests both types of features are highly effective in capturing the artifacts introduced by the image recapturing process. By combining the LBP, MSWS and CF features, the identification performance can be further improved to a very low EER of 0.5%. We have also compared our proposed features with the wavelet statistics features in [3]. It should be noted that according to [4], these wavelet statistics features have better performance than the geometry features [4] in identification of recaptured computer graphic images. Hence we consider these wavelet statistics are the best statistical features in the literature for identification of the recaptured scenery. The comparison result shows our combined features have clear advantage over these wavelet statistics features, where the EER of our combined features is 85.1% lower than the wavelet statistics features.

With our trained PSVM classifier based on our combined features, we also classify the 50 survey photos used in human survey in Section 3. All the 50 photos are correctly classified.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we study identification of the finely camera recaptured images on LCD screens. Through proper setup of the image recapturing environment and by tuning the controllable settings, we can recapture the artificial sceneries displayed on LCD screens with reasonably good quality. In a human identification study, we notice that such recaptured images can hardly be identified successfully by human eyes. Hence, these finely recaptured photos potentially post a threat for image forgers to both walk around the current forgery detection system and to fool human eyes. To prevent such a security threat, we propose using several types of statistical features for classification of the recaptured images from natural images. Our proposed features capture the textured patterns, the loss-of-fine-details characteristics and the color anomalies introduced in the image recapturing process. Comparison results show the proposed features works excellently in identification of the recaptured photos and they outperform the state-of-art forensics features in identification of the finely recaptured images on LCD screens.

In the future, we will further investigate other common image recapturing attacks such as recapturing on high-quality printed photos, projection and scanning of high-quality printed photos. These different image recapturing processes possibly introduce the similar artifacts we have found in this paper. Our current method will be adapted and extended to address all sorts of common image recapturing attacks.

6. REFERENCES

[1] "Rare-Tiger Photo Flap Makes Fur Fly in China," *Science*, vol. 318, no. 5852, pp. 893, 2007.

[2] S. Lyu, "Natural Image Statistics for Digital Image Forensics," Dartmouth College, 2005.

[3] S. Lyu and H. Farid, "How Realistic is Photorealistic?," *IEEE Trans. on Signal Processing*, vol. 53, pp. 845-850, Feb 2005.

[4] T.-T. Ng, S.-F. Chang, J. Hsu, L. Xie, and M.-P. Tsui, "Physics-Motivated Features for Distinguishing Photographic Images and Computer Graphics," in *Proc. of ACM Int. Conf. on Multimedia*, 2005, pp. 239-248.

[5] H. Yu, T.-T. Ng, and Q. Sun, "Recaptured Photo Detection Using Specularity Distribution," in *Proc. of ICIP*, 2008, pp. 3140-3143.

[6] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution Gray-Scale and Rotation Invariant Textureclassification with Local Binary Patterns," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 971-987, 2002.

[7] M. Kharrazi, H. T. Sencar, and N. Memon, "Blind Source Camera Identification," in *Proc. of ICIP*, vol. 1, 2004, pp. 709-712.

[8] Y. Chen, Z. Li, M. Li, and W.-Y. Ma, "Automatic Classification of Photographs and Graphics," in *Proc. of ICME*, vol. 9-12, 2006, pp. 973-976.

[9] C.-W. Hsu, C.-C. Chang, C.-J. Lin, "A Practical Guide to Support Vector Classification," 2008

[10] H. Farid, "A Survey of Image Forgery Detection," *IEEE Signal Processing Magazine*, vol. 26, pp. 16-25, 2009.

[11] F.-T. Pai, "Method for Driving a Liquid Crystal Display in a Dynamic Inversion Manner," US Patent 7109964, 2006.