Further Studies on Forensic Features for Source Camera Identification

Ying-Chu Chen, Yongjian Hu, Chang-Tsun Li

Department of Computer Science,
University of Warwick
Coventry CV4 7AL, UK
e-mail: y.c.chen@warwick.ac.uk, Yongjian.Hu@dcs.warwick.ac.uk, c-t.li@warwick.ac.uk

Keywords: Digital image forensics, camera identification, LibSVM , robust camera classifier, feature sets.

Abstract

Most camera identification schemes focus on finding image features that can increase classification accuracy as well as computational efficiency. For forensic investigation purposes, however, these selection criteria are not enough since most real-world photos may have undergone common image processing due to various reasons. Therefore, source camera classifiers must have the capability to resist the influence of common image processing when they tackle these processed photos. In this work, we implement a published camera classifier and investigate the performance of the classifier on images under shearing, histogram equalization, and contrast-stretching operations. Besides, we probe into the impact of camera databases of different sizes on the performance of the classifier.

1 Introduction

With the dramatic development of technology and the decline of manufacturing cost, all kinds of digital devices such as digital camera, video camera, or scanner are easy to obtain and purchase. Posting a video of birthday party or uploading several photos of graduation ceremony on personal Facebook are convenient to everyone. Moreover, with the mature and various kinds of software, those digital data of image or video can be edited unnoticeably and naturally. Movie Titanic reveals impressive grand sights with pre-eminent film digital effects, and attractive model photos on fashion magazines profit from image processing software such as Photoshop and Photoimpact. Processed pictures and videos enrich people's lives through technology growth. However, when a photo or a video clip involves law evidence, can those data still be believed and become a key factor of judgment in court? Owing to the need of reliable and convincing digital evidence, further research and standard techniques in digital forensics domain are in demand.

Image source identification is one of the important branches of forensics research field. Influenced by traditional steganalysis [1, 2], the use of statistical image features turns into one of the most popular ways in camera source identification. Because of the embedded signal or noise in images from a particular camera, it can be chased to identify the camera source from unknown images as the statistics of images captured by different device are believed to be different.

In previous research, different and various statistical features have been proposed and discussed. Farid and Lyu discovered that strong higher-order statistical regularities exist in the wavelet-like decomposition of a natural image, and the embedding of a message significantly alters these statistics and thus becomes detectable [1]. Two statistical features sets have been used in their paper: (i) feature set based on the mean, variance, skewness and kurtosis of the subband coefficients, and (ii) feature set in terms of the errors in an optimal linear predictor of coefficient magnitude.

Avcibas found that steganographic schemes reveal statistical evidence which can be exploited for detection with the aid of image quality features and multivariate regression analysis. Image quality metrics (IQMs) have been used to distinguish between cover and stego images [2].

In [3], Kharrazi et al. proposed 34 features, including average pixel value (3 features), RGB pairs correlation (3 features), neighbour distribution centre of mass (3 features), RGB pairs energy ratio (3 features), wavelet domain statistics (9 features), and IQMs (13 features), to classify 5 cameras with Support vector machines (SVM) and the average identification accuracy reached 88.02%. In [4], Tsai has elaborated the scheme in [3] and re-implemented with 33 features from three categories (colour features, image quality features and wavelet domain features). Tsai's results are consistent with Kharrazi's research on different models and brands cameras.

Scanner identification also has referential value due to the similarity with cameras. In [5], Gou et al. applied 60 statistical noise features from denoising algorithms (30 features), wavelet analysis (18 features), and neighborhood prediction (12 features). The result of their SVM-based classifier is over than 95%. In [6], Khanna et al. presented a scanner source identification method using statistical properties of the sensor pattern noise (SPN). The SPN was first applied to correlation-based camera identification in [7]. The photo response non-uniformity noise (PRNU) is main component of SPN.

Khanna et al. proposed a special method to extract the statistical features from the mean, STD, skewness, and kurtosis of the row correlations and the column correlations on each colour channel. The accuracies from their SVM-based classifier were often better than those in [5]. In addition, Khanna et al. also gave the results when the photographs are manipulated by JPEG compression, contrast stretching and sharpening. From [8]-[10], various statistical image features are available.

Inspired by [14] and [15], we implement a camera classifier and discuss its robustness on images under different image operations. We also examine its performance for camera databases of different sizes.

2 Our Sample Camera Classifier

In this work, we will categorize five feature sets which are *wavelet features, colour features, IQMs, statistical features of difference images*, and *statistical features of prediction errors*. In order to simplify their expression, we index those features as Feature Sets I, II, III, IV, and V, respectively. These features are commonly used in previous work, as mentioned in Section 1. In this section, we will interpret the functions of these feature sets.

2.1 Feature Set I: Wavelet Features

Wavelet Feature describes the correlation between the subband coefficients. The mean, variance, skewness and kurtosis of high-frequency subband coefficients have been used. We use biorthogonal 9/7 wavelet filters and implement one-scale wavelet transform on each colour band. In Figure 1, 3×3×4=36 features are obtained, where the first "3" refers to three colour channels, the second "3" refers to three high-frequency subbands, and "4" refers to four features.

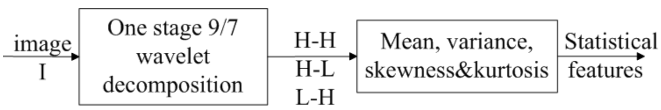


Figure 1. Wavelet features

2.2 Feature Set II: Colour Features

Feature Set II is composed of 3+3+3+3=12 colour features (see Figure 2) including the average value of each colour band, the correlation pair between two different colour bands, the neighbour distribution centre of mass for each colour band and three energy ratios, namely $E1 = |G|^2/|B|^2$, $E2 = |G|^2/|R|^2$, and $E3 = |B|^2/|R|^2$.

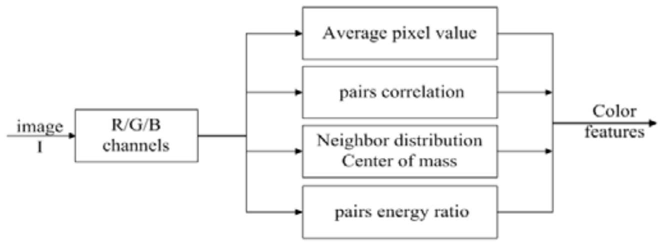


Figure 2. Colour features

2.3 Feature Set III: Image Quality Metrics (IQMs)

In Figure 3, Feature Set III consists of 3+3+6=12 IQMs including three pixel difference-based features, i.e., Minkowski difference, mean absolute error with $\gamma = 1$, and mean square error with $\gamma = 2$; three correlation-based features, i.e., structural content, normalized cross correlation, and Czekonowski correlation; six spectral features, i.e., spectral magnitude error, spectral phase error, spectral phase-magnitude error, block spectral magnitude error, block spectral phase error, and block spectral phase-magnitude error. More detailed information can be found in [15].

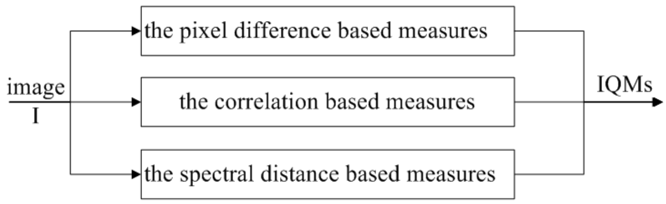


Figure 3. Image quality metrics (IQMs)

2.4 Feature Set IV: Statistical Features of Difference Images

In order to obtain Feature Set IV, the averaging filter, Gaussian filter, median filter, and Wiener adaptive filters with 3×3 and 5×5 neighbourhoods are separately used to acquire the difference images. Similar to [5], we first perform the absolute operation on the difference images, and then take $\log_2$ transformation. Afterwards, we calculate the mean and STD of the $\log_2$-transformed absolute values. 2×5×3=30 features are obtained as shown in Figure 4, where "3" refers to three colour channels, "5" refers to five denoising algorithms, "2" refers to two features.

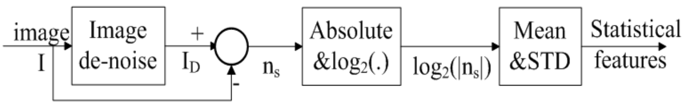


Figure 4. Statistical features of difference images

2.5 Feature Set V: Statistical Features of Prediction Errors

Feature Set V consists of 2×2×3=12 statistical features of prediction errors (see Figure 5), where "3" refers to three colour channels, the first "2" refers to two different smooth regions, and the second "2" refers to two statistical features.

Strong correlation exists across a natural image, in particular, in smooth regions. Thus, pixel values in smooth regions can be predicted from their neighbouring pixels with high accuracy. For images from different cameras, however, linear prediction error is probably different. The mean and STD of linear prediction errors are then used as statistical features of prediction errors. The way in [5] is borrowed to obtain Feature Set V.

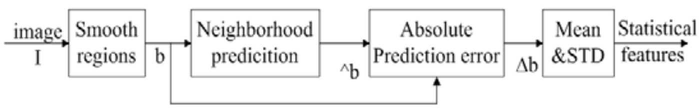


Figure 5. Statistical features of prediction errors

3 Experiments and Discussions

The above five feature sets form our feature vector of 36+12+12+30+12=102 dimensions. We use this vector as the input of a camera classifier. Since the LIBSVM toolbox [12] with a nonlinear RBF kernel is frequently used for camera/scanner identification in literature, we adopt it in our experiments for the sake of comparison.

We take a way as in [15] to investigate the performance of the classifier on test images manipulated by three common image operations: shearing, histogram equalization and contrast-stretching transformation. Photos are captured with ten cameras: Canon IXUS70, Canon IXUS80IS, Canon IXUS500, Canon IXUS800IS, Fujifilm F200EXR, Nikon D70, Nikon D80, Nikon D90, Panasonic DMC-LX2, and Sony DSC-H2. Each camera captures 300 photos, half the photos are used as the training set, and the rest of the photos are used as the testing set. For fair comparison, we take 1024×1024 as a standard size for all images. Each image is cropped from the centre of a photo, as shown in Figure 6.

The classification accuracy of the classifier using all the five Feature Sets on unprocessed images is 71%. When the classifier uses one feature set a time, its accuracies from Set Ⅰ to Set Ⅴ are 70%, 27%, 17%, 36%, and 24% , respectively.

3.1 Experiment Results of Shear Images

In some media, especially in fashion magazine industry, photo-retouching is an increasingly worsening problem. Over slender and exaggerated pictures of models may cause teenagers' imitation and result in harm to their health, even anorexia. Shearing is one of the ways to prettify images. One of the most common used forms of spatial transformation is the *Affine Transform* which can be written in matrix form as (1).

$$\begin{bmatrix} x & y & 1 \end{bmatrix} = \begin{bmatrix} w & z & 1 \end{bmatrix} \mathrm{T} = \begin{bmatrix} w & z & 1 \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & 1 \end{bmatrix} \qquad (1)$$

We use horizontal shearing where $\mathrm{T} = \begin{bmatrix} 1 & 0 & 0 \\ \alpha & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, and set the tilt degree α as 5 degree. From Figure 7, we can observe that the average classification accuracy of our sample classifier on images after shearing is 34%. When the classifier only uses one feature set each time, its classification accuracy from Set Ⅰ to Set Ⅴ are 23%, 32%, 28%, 30%, and 28%, respectively. Apparently, shearing operation degrades the performance of the classifier using all the five Feature Sets. However, for the classifier separately using Feature Set II, III, and V, the classification accuracies become a little better.

3.2 Experiment Results of Equalized Images

Histogram equalization is a common image processing operation. Its main purpose is to increase global contrast to images, especially when the helpful detail of the image is represented by close contrast values. Figure 8 shows the original image (left) and the equalized image (right). We can observe that histogram equalization increases dynamic range and makes image details better visible.

In order to equalize colour images, we transfer the images from RGB (Red, Green, Blue) colour space to HSI (Hue, Saturation, Intensity) colour space. After the intensity has been histogram equalized, we transfer images from HSI to RGB. The average classification accuracy of the classifier on histogram equalized images is 16%. When using one feature Set, the accuracy of the classifier from Set Ⅰ to Set Ⅴ are 20%, 22%, 16%, 25%, and 21%, respectively, as shown in Figure 7. It can be seen that such image processing has great impact on the classifier, and degrades the performance of the classifier using all the five Feature Sets as well as separately using each Feature Set.

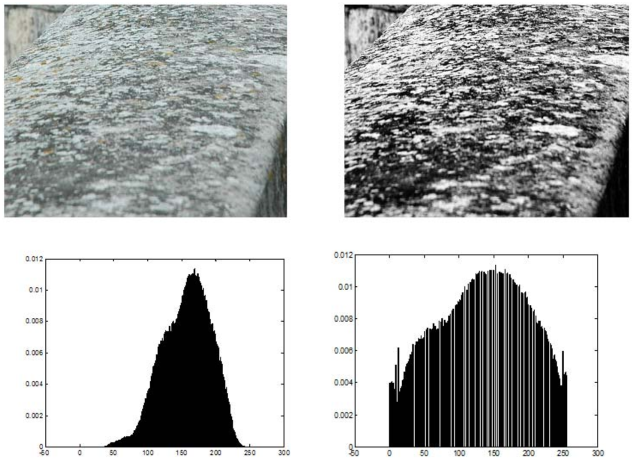


Figure 8. Original image (left) and equalized image (right)

3.3 Experiment Results of Contrast-Stretching Images

Contrast-Stretching transformation increases the contrast between the darks and the lights. Formula (2) and Figure 4 show a typical contrast-stretching transformation, where r represents the intensities of the input image, s the corresponding intensity values in the output image, and the parameter E controls the slope of the function.

$$s = T(r) = \frac{1}{1 + \left(m / r\right)^{E}} \qquad (2)$$

We take contrast-stretching images with E=4 and use the mean of the image intensities as m. The average classification accuracy of the classifier is 13%. The accuracy of the classifier from using Set Ⅰ to Set Ⅴ are 16%, 27%, 15%, 18%, and 17%, respectively. As a result, contrast-stretching has enormous impact on the performance of the classifier. It degrades the performance of the classifier using all the five Feature Sets as well as separately using each Feature Set.

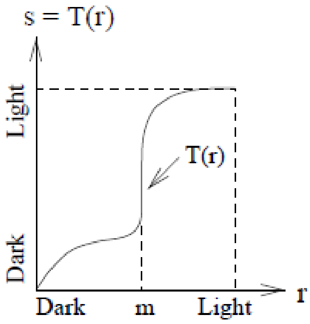


Figure 4. Contrast-Stretching transformation [16]

4 The Performance on Larger Image Database

In order to investigate the impact of the size of camera database on the performance of the classifier, we use the camera database of our previous work in [15] as the basic database and add three cameras into the database each time to change the size of database, and then investigate the change of performance of the classifier. Specifically, we implement the experiments of 13 cameras, 16 cameras, and 19 cameras. The nine more cameras are Canon IXUS70, Canon IXUS80IS, and Canon IXUS500, Sony DSC-H2, Fujifilm F200EXR, Nikon D70, Nikon D80, Nikon D90 and Panasonic DMC-LX2. For distinguishing 10 cameras (*X1* to *X10*) used in [15], we name the newly added cameras as *X11*, *X12, X13, X14, X15, X16, X17, X18,* and *X19*, respectively.

4.1 Experimental Results on Original Unprocessed Images for 10, 13, 16, 19 Cameras

From Figure 8.(a), it can be observed that the classifier using all the five image feature sets has the best performance, and the accuracies are 92%, 88%, 88%, 86% for 10, 13, 16, and 19 cameras, respectively. A noteworthy result is that the classifier using Set I (Wavelet Features) has a good performance close to that of the classifier using all the five feature sets. The accuracies of using Set I are 92%, 88%, 87%, 85% for 10, 13, 16, and 19 cameras, respectively. This means that Sets II to V only have minor contribution to the accuracy increase of the classifier under this circumstance. So we think that the use of Wavelet Features is good for fast classification and time saving when test images are not manipulated.

4.2 Experimental Results on Compressed Images for 10, 13, 16, 19 Cameras

From Figure 8.(b), we can see the accuracies decline as the camera number increases. The classifier using only Feature Set II (Colour Features) has the best average accuracy, 37%. The classifier using all the five Feature Sets, Set I, Set III, Set IV, and Set V have the classification accuracies 27%, 14%, 25%, 19% and 27%, respectively. However, the Standard Deviation (STD) of accuracies using Set II is 8.25, which is much higher than that of using Set IV (Statistical feature of difference images) , i.e., 3.96. This implies that the use of Set II only has high accuracy when the database is small. With the increase of camera number, the use of Set IV will ensure more steady performance of the classifier than other Feature Sets.

4.3 Experimental Results on Cropped Images for 10, 13, 16, 19 Cameras

In Figure 8.(c), the classifier that only uses Set IV (Statistical features of difference images) has the highest average accuracy 70%. The accuracies of this classifier are 84%, 65%, 70%, and 60% for 10, 13, 16, and 19 cameras, respectively. Nevertheless, the accuracies of using Set IV have also the highest STD 10.19, which means that predicted accuracy might fail when the number of cameras increases. On the other hand, when all the five feature sets are used, the classifier has the most stable performance since its accuracies do not change much when the size of camera database changes.

4.4 Experimental Results on Scaled Images for 10, 13, 16, 19 Cameras

In Figure 8.(d), the classifier that uses all the five Feature Sets has the best classification accuracies, which are 53%, 56%, 50%, and 56%, respectively. In addition, the accuracies of the classifier using all the five Feature Sets also have the lowest STD, 3.09. Undoubtedly, using all the five Feature Sets is a good choice for scaled images. Meanwhile, the use of Set IV (Statistical feature of difference images) results in close accuracies of the classifier: 58%, 51%, 55%, and 50% for 10, 13, 16, and 19 cameras, respectively. Their corresponding STD is also 3.09. Therefore, the use of Set IV is also good for handling scaled photos.

5 Conclusions

Based on our experiments, we have two conclusions:

a) Among shearing, histogram equalization and contrast-stretching, the latter two operations have more negative impact on the performance of the classifier than shearing. Probably, histogram equalization and contrast-stretching have more global effects on an image than shearing does.

b) The size of camera databases has apparent impact on the performance of the classifier. Generally, the classification accuracy descreases when the camera database becomes larger. However, for images under different image processing operations, the degradation degree of the performance of the classifier is different.

The problem of camera identification is a complex one with no universally applicable solution so far. According to our experiments, different image operations and different sizes of camera databases affect the accuracy of the camera classifier with different degrees. Finding robust statistical features against variation of image alteration and the increase of camera database size will be our future work.

6 Acknowledgement

This work is partially supported by the EU FP7 Digital Image and Video Forensics project (Grant Agreement No. 251677).

References

[1] H. Farid and S. Lyu. "Detecting hidden messages using higher-order statistics and support vector machines," Proc. of 5th Int. Workshop on Information Hiding, Springer-Verlag, Berlin, Heidelberg, vol. 2578, pp. 340-354, 2002.

[2] I. Avcibas, N. Memon, B. Sankur, "Steganalysis using image quality metrics," IEEE Trans. on Image Process., vol. 12, no. 2, pp. 221-229, Jan. 2003.

[3] M. Kharrazi, H.T. Sencar, and N. Memon, "Blind source camera identification," Proc. of IEEE Int. Conf. on Image Processing, Singapore, pp. 709-712, Oct. 24-27, 2004.

[4] M.-J. Tsai and G.-H. Wu, "Using image features to identify camera sources," Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Processing, vol. 2, 2006.

[5] H. Gou, A. Swaminathan, and M. Wu, "Intrinsic sensor noise features for forensic analysis on scanners and scanned images," IEEE Trans. on Inf. Forensics Security, vol. 4, no. 3, pp. 476-491, Sept. 2009.

[6] N. Khanna, A.K. Mikkilineni, and E.J. Delp, "Scanner identification using feature-based processing and analysis," IEEE Trans. on Inf. Forensics Security, vol. 4, no. 1, pp. 123-139, Mar. 2009.

[7] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," IEEE Trans. on Inf. Forensics Security, vol. 1, no. 2, pp. 205-214, Jun. 2006.

[8] T. Filler, J. Fridrich, and M. Goljan, "Using sensor pattern noise for camera model identification," Proc. of IEEE Int. Conf. on Image Processing, pp. 1296-1299, 2008.

[9] M.-J. Tsai, C.-L. Lai, and J. Liu, "Camera/Mobile phone source identification for digital forensics," Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Processing, pp. 221-224, 2007.

[10] M.-J. Tsai and C.-S Wang, "Adaptive feature selection for digital camera source identification," Proc. of IEEE Int. Symposium on Ciruits and Systems, May 2008.

[11] I. Avcibas. Image quality statistics and their use in steganalysis and compression. Ph.D. Thesis, Bogazici University 2001.

[12] C.-C. Chang and C.-J. Lin, LIBSVM: A Library for Support Vector Machines 2001. http://www.csie.ntu.edu.tw/~cjlin/

[13] C.-T. Li, "Source camera identification using enhanced sensor pattern noise," IEEE Trans.on Inf. Forensics Security, vol. 5, no. 2, pp. 280-287, Jun. 2010.

[14] T. Gloe, M. Kirchner, P. Winkler, and R. Bohme, "Can we trust digital image forensics," Proc. of the 15th International Conference on Multimedia, Sept. 2007.

[15] Yongjian Hu, Chang-Tsun Li, and Changhui Zhou. Selecting forensic features for robust source camera identification. in Computer Symposium (ICS), 2010 International. 2010.

[16] Dmitrij Csetverikov, Basic Algorithms for Digital Image Analysis. http://visual.ipan.sztaki.hu

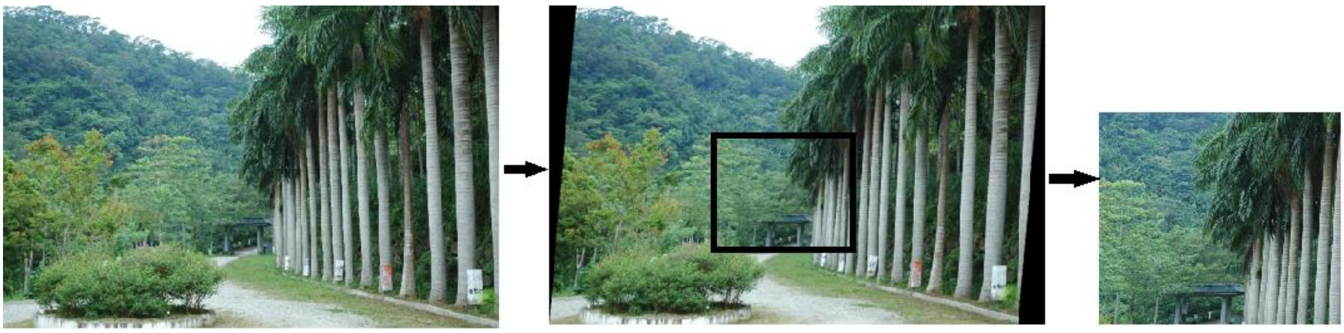


Figure 6. Shearing and cropping an image.

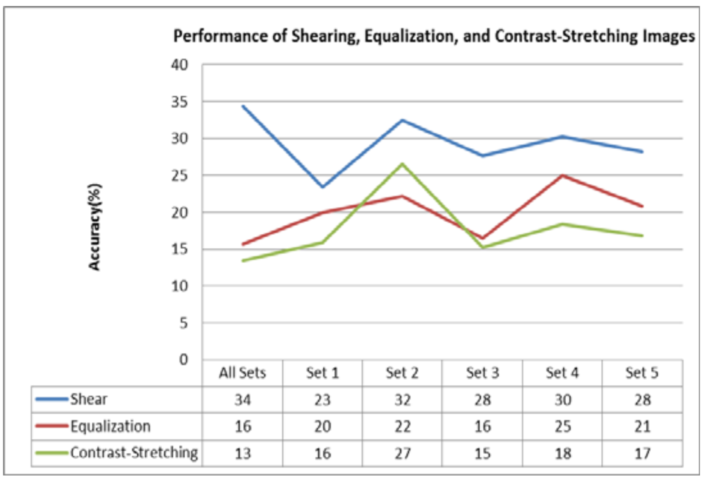
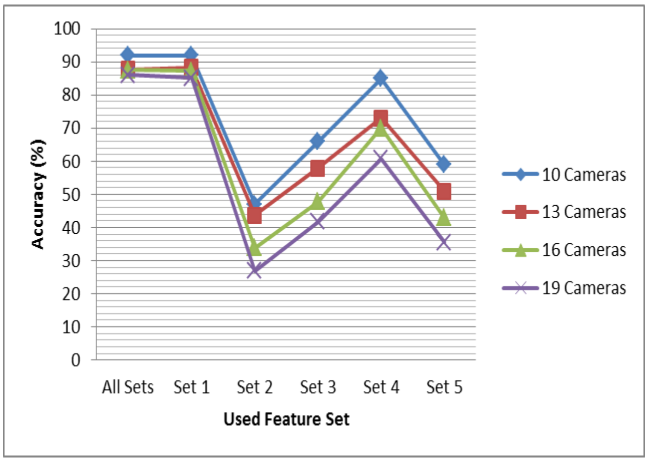
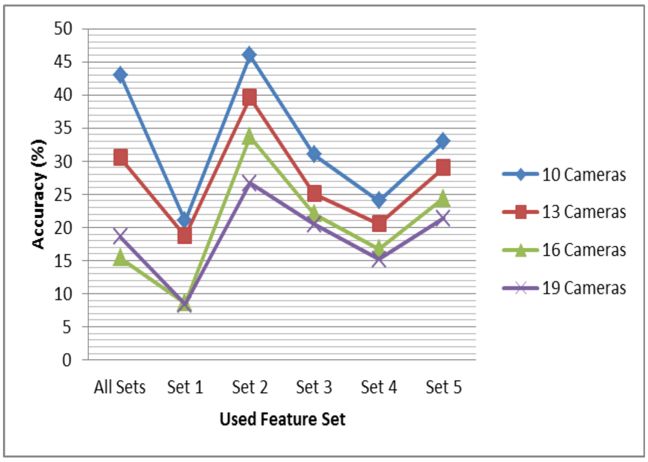


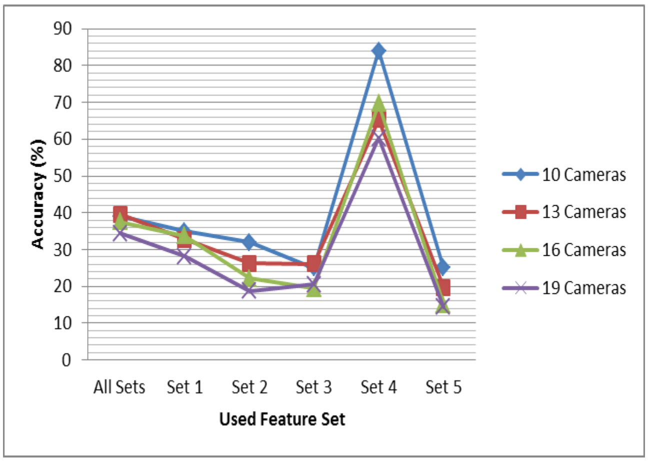
Figure 7. Classification accuracy of the classifier on shear, Equalized, and Contrast-Stretching Images.



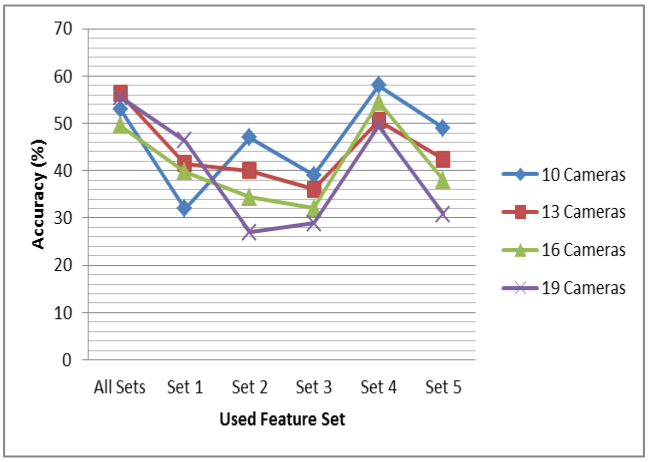
(a)



(b)



(c)



(d)

Figure 8. (a) Performance of the classifier on original unprocessed images. (b) Performance of the classifier on compressed images. (c) Performance of the classifier on cropped images. (d) Performance of the classifier on scaled images.