

Cross-Class and Inter-class Alignment Based Camera Source Identification for Re-compression Images

Guowen Zhang, Bo Wang^(✉), and Yabin Li

Dalian University of Technology, Dalian City 116024,
Liaoning Province, People's Republic of China
bowang@dlut.edu.cn

Abstract. With the sophisticated machine learning technology developing the state of art of model based camera source identification has achieved a high level of accuracy in the case of matching identification, which means the feature vectors of training and test sets follow the same statistical distribution. For a more practical scenario, identifying the camera source of an image transmitted via social media applications and internet is a much more interesting and challenging work. Undergoing serials of manipulations, re-compression for instance, the feature vectors of training and test sets mismatch, thus decreasing the identification accuracy. In this paper, cross-class and inter-class alignment based algorithms, inspired by transfer learning, are proposed to minimize the distribution difference between the training and the test sets. Experiments on four cameras with five image quality factors indicate that the proposed cross-class, inter-class alignment based algorithms and their combination outperform the existing LBP method, and presents high identification accuracies in re-compression images.

Keywords: Cross-class alignment · Inter-class alignment
Camera source identification · Re-compression

1 Introduction

As the development of the social platforms and online media brings great convenience to people, digital images, as one kind of multimedia, are becoming much more important in our daily lives. At the same time, various image processing software and applications have been very popular, and people do not need professional skills for image tampering any longer. Image tampering and forging become easy and convenient. Recently, several events of tampering and forgery of digital images have had negative effect on various fields, justice and science for instance. People's confidence in the news and social integrity has declined. More seriously, when the tampered image is used as an evidence in the court, there would be lots of troubles in the justice system.

Because of the increasing requirements of digital media integrity and authenticity analysis, the digital image forensics technology has been rapidly developed in the last decade [1]. As one of the most important branch of digital images forensics, source camera identification focuses on identifying the device type, device model or device individual which captures the image. Correspondingly, the source camera identification

is divided into three categories [2]. Device-based identification, which aims to explore the type of device capturing the image [3], usually focuses on distinguishing the computer generated image, camera image, cell phone image and scanned image. Model-based identification pays more attention on identifying the device model [4], for instance camera model, which generates the digital image. Camera-based approaches look forward to specifying the camera individual. In this paper, only the camera model identification approaches could be investigated and discussed.

In recent years, several algorithms for model-based camera source identification have been proposed. Most of these methods follow a framework, which considers the problem of camera model identification as the classification issue, and as a result, it could be solved by machine learning approaches. As a matter of fact, numerous existing methods of source camera identification in the framework have achieved an approximately high detection accuracy more than 90%. A typical approach proposed by Xu *et al.*, extracts 354-dimensional features based on local binary patterns to distinguish camera models. Considering 13 brand and 18 camera models, the average accuracy of the LBP method reaches up to 98% [5]. Another outstanding work by Swaminathan *et al.* constructs an efficient liner model to estimate the interpolation coefficients of color filter array [6]. Using the CFA coefficients as the feature vector, a distinguished classifier is established for camera source identification, and a high average discrimination accuracy of 90% is achieved on a dataset of 19 camera models [4]. Although the decent high detection accuracies, these algorithms set a default scenario that the training and test image samples are considered raw and unprocessed images, in other words, in the laboratory environment.

Considering a more practical scenario, the digital image, captured by a camera or cell phone, is uploaded to the internet via the fashionable social media applications, for instance Twitter and Facebook in the U.S., QQ, WeChat and Microblog in China. Sequentially, the image could be spread and forwarded. When the digital image is adopted as an evidence in the court, the digital image is usually manipulated and processed in the internet. Because of the popularity of the social media applications, this scenario is more practical and significant for the real world. In this paper, the more practical and meaningful scenario is explored and discussed.

In the more practical social media application based scenario, the digital images may experience retouching, geometric transformation such as resizing, re-compression and even D/A, A/D transform and so on, because of the limitations of the social media application platforms and communication channels [2]. It is the enhanced operations and manipulations that make the test images, which are used for camera source identification, no longer the raw ones from the devices. Furthermore, the statistical characteristics and feature distributions deviated from that of the training raw digital images. The deviation of the training raw image samples and the test manipulated image samples makes the practical camera source identification a challenging task.

Typically, the manipulations in the social media applications and the internet is complicated and manifold. To simplify the model, we focus on the re-compression operation in the image pipeline. In spite of the restrictions of terminal screen resolution, the re-compression image via the social media applications and internet show few visual differences from its raw copy. Thus, JPEG compression is an almost essential operation

of image processing in social networks, which makes the simplification reasonable and feasible. But as mentioned before, the statistical characteristics and feature distributions vary largely, which means the identification accuracy in the laboratory environment is no longer reliable.

Wang *et al.* analyzes the influence of JPEG compression on the typical camera model identification approaches [7]. The classification accuracies of the existing algorithms decrease rapidly with the decreasing of JPEG compression quality. An intuitionistic solution could construct an online training system, which generates the training image samples and trains the classification model online, according to the re-compression quality factor obtained from the test samples. In this scheme, the computation costs is huge and online training is time-consuming. Obviously, the scheme of online training system could work while it is unrealistic.

In this paper, we focus on the camera source identification of re-compression images. Instead of the online training system, algorithms based on cross-class alignment (CCA) and inter-class alignment (ICA), inspired by transfer learning, are proposed.

The paper is organized as follows. In Sect. 2, following by the instruction of motivation of transfer learning, CCA and ICA based methods are proposed, conjunction with a description of the LBP features used in the method. The experiments are demonstrated and discussed in Sect. 3. Finally, the paper is concluded in Sect. 4.

2 Proposed Algorithms

In this section, we first give a brief introduction of the transfer learning, which inspires us to design the algorithms. Subsequently, the cross-class and inter-class alignment based algorithms are introduced in detail.

2.1 Transfer Learning

In the practical scenario of camera source identification for a digital image, the most important reason of fast reduction of identification accuracy is the variation of statistical characteristics and feature distributions between training set and the test image samples, which is caused by the re-compression manipulation used in the social media applications and internet [7]. The variation makes the existing methods deviate from the assumption that training set and test set are subject to the same distribution. Without the foundation of classification model, the source camera identification methods based on machine learning spontaneously degrade in the performance field.

To solve the mismatch of the distribution, an important approach named transfer learning is proposed. It relaxes the two restrictions of the basic assumptions in traditional machine learning: first one is that the statistical characteristics and feature distributions of training data and test data follow the identically distribution, and the other is that there are enough labeled samples to train a good classification model. Transfer learning aims to use the learned knowledge of training data to solve the problems that there are few or even not any labeled data in test data [8]. Transfer learning suggests that if the training data and test samples are inherently correlated, the training progress in the classifier

definitely contributes to the classification of the test samples, even though the distribution of training data varies from that of the test samples.

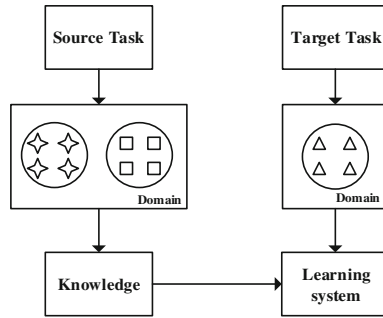


Fig. 1. Transfer learning.

The ‘domain’ contains the feature space χ and the marginal probability distribution $P(x)$, where the feature space χ denotes eigenvector space. The ‘task’ can be considered as the tag space y and the target prediction function $f(\bullet)$, where $f(\bullet)$ is used to predict the label of the test sample equivalent to $P(x|y)$. In the other word, the ‘task’, including the source task T_s and target task T_t in the camera source identification, means classification, as Fig. 1 illustrated. In this scenario, transfer learning try to solve the problem described as following.

$$\begin{cases} D_s \neq D_t \\ T_s = T_t \end{cases} \quad (1)$$

The key of transfer learning means constructing a projection, or a transformation to minimize the distribution difference between the training set and the test set.

2.2 Cross-Class Alignment Based and Inter-class Alignment Based Approaches

Inspired by the transfer learning, we try to design a transformation to minimize the distribution deviation. In this case, two independent approaches are proposed. In our work, the re-compression manipulation is the only factor that makes the distributions of source domain and the target domain deviated. Considering the characteristics of JPEG compression, a Gaussian model is used to evaluate the deviation of these two domains.

Cross-Class alignment based approach:

Based on the assumption that the test image samples are all re-compressed, a cross-class alignment base approach, which means a global alignment between the source domain (training set) and the target domain (test set), is proposed. It supposes that we have a set of n_s samples $S = \{p_1, p_2, \dots, p_{n_s}\} \in R^d$ in the source domain and a set of n_t samples $T = \{q_1, q_2, \dots, q_{n_t}\} \in R^d$ in the target domain, where d is the dimension of the feature vector. To evaluate the difference between these two Gaussian model domains, the

minimization of expectation and the standard deviation are considered, as Eqs. (2) and (3) shows.

$$E(\varphi(s^j)) = E(t^j) \quad (2)$$

$$\sigma(\varphi(s^j)) = \sigma(t^j) \quad (3)$$

where s^j denotes of the j -th feature of training samples, and t^j disnotes the j -th feature of the test samples, $j = 1, 2, \dots, d$. $E(s^j)$, $\sigma(s^j)$ is defined to represent the expectation and the standard deviation of the j -th feature of samples in source domain. The transformation $\varphi(\cdot)$ for each feature is defined as following [9].

$$\varphi(s_i^j) = (s_i^j - E(s^j)) \frac{\sigma(t^j)}{\sigma(s^j)} + E(t^j) \quad (4)$$

where $j = 1, 2, \dots, d$, and $i = 1, 2, \dots, n_s$.

Inter-class alignment based approach:

The global alignment, cross-class alignment, focuses on the re-compression effect between the training set and the test set. Considering the alignment between the labels, an inter-class alignment base approach is presented. Correspondingly, we aim to minimize the intra-class expectation and standard deviation, as Eqs. (5) and (6) shows.

$$E(\varphi(s^j), y) = E(t^j, y) \quad (5)$$

$$\sigma(\varphi(s^j), y) = \sigma(t^j, y) \quad (6)$$

where $j = 1, 2, \dots, d$.

Similar with the cross-class alignment, the transformation $\varphi(\cdot)$ could be described as following, with a specified label for each class [10].

$$\varphi(s_i^j) = (s_i^j - E(s^j, y_i)) \frac{\sigma(t^j, y_i)}{\sigma(s^j, y_i)} + E(t^j, y) \quad (7)$$

where $j = 1, 2, \dots, d$, $i = 1, 2, \dots, n_s$.

The problem of the Eq. (7) is that the label of the target domain is unavailable, which means the $E(t^j, y)$ and $\sigma(t^j, y)$ are not available. As a result, we have to use the joint estimate $p(y|t_i)$ instead of the label y . To obtain the $p(y|t_i)$, we directly train a classifier model using the training samples and predict the test samples. Therefore, the approximate $E(t^j, y)$ and $\sigma(t^j, y)$ could be computed as following:

$$E(t^j, y) \approx \frac{1}{\sum_{i=1}^{n_i} p(y|t_i)} \sum_{i=1}^{n_i} t_i^j p(y|t_i) \quad (8)$$

$$\sigma(t^j, y) \approx \sqrt{\frac{1}{\sum_{i=1}^{n_i} p(y|t_i)} \sum_{i=1}^{n_i} (t_i^j - E(t^j, y))^2 p(y|t_i)} \tag{9}$$

where $j = 1, 2, \dots, d$.

Furthermore, the cross-class alignment and inter-class alignment can be combined, to minimize the influence of re-compression.

2.3 LBP Features

To verify the proposed alignment based algorithms, efficient features are required in our approaches. LBP [5], CFA [6] and IQM [11] are admitted outstanding feature vectors used for model-based camera source identification. In our work, the LBP features are adopted and also in our future work, the other feature vectors are certified resultful.

The LBP features, propose by Xu *et al.*, are designed based on uniform gray-scale invariant local binary patterns [3], which can be described as:

$$LBP_{p,R}^{u2} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p \tag{10}$$

where R is the radius of a circularly symmetric neighborhood used for local binary patterns, and P is the number of samples around the circle. We set $R = 1, P = 8$. g_c and g_p represent gray levels of the center pixel and its neighbor pixels, respectively. Which is showed in Fig. 2.

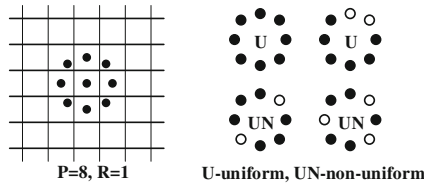


Fig. 2. (Left) Constellation of neighborhood. (Right) Example of ‘uniform’ and ‘non-uniform’ local binary patterns.

Function s is defined as:

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{11}$$

The differences between the central pixels and the neighborhood pixels are firstly calculated. Subsequently, the differences are binary quantized and coded according to the function s . to form a 8-dimensional histogram with a total of $2^8 = 256$ bins named local binary patterns. Inspired by [12], both of the ‘uniform’ and ‘non-uniform’ local binary pattern are included in [5]. Considering the majority of ‘uniform’ local binary

pattern in the total patterns, only 58 ‘uniform’ patterns are merged with the ‘non-uniform’ patterns to generate 59 effective patterns. For each color channel, the LBP features are extracted from (i) original image, (ii) its prediction-error counterpart, and (iii) its 1st-level diagonal wavelet subband, resulting in a total of $59 \times 3 = 177$ features, as Fig. 3 illustrated. Respect to the same processing strategy for red and blue channel because of Bayer CFA, we only extract LBP features from red and green channels to reduce the dimension of the feature vector. Finally, a total of $177 \times 2 = 354$ features are achieved.

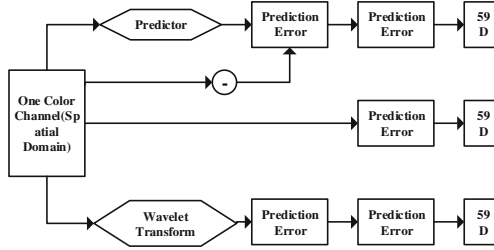


Fig. 3. Feature extraction framework for one color channel

3 Experimental Results

3.1 Experimental Setup and Parameters

To testify the performance of the proposed algorithms, four camera models from ‘Dresden Image Database’ are used in our experiments. For each camera model, 350 image samples are selected randomly and each image is cropped into 6 non-overlap sub-images. Therefore, 2100 images samples are obtained in each camera model, as Table 1 shows.

Table 1. Image dataset.

Camera model	Resolution	Images number	Samples number	Abbr
Agfa_DC-830i	3264×2448	350	2100	AGF
Kodak_M1063	3664×2748	350	2100	KOD
Pentax_OptioA40	4000×3000	350	2100	PET
Sony_DSC-W170	3648×2736	350	2100	SON

In all of our experiments, LibSVM [13] is used as the classifier to train the classification model and classify the test image samples. 1500 images of each camera are randomly selected as the training set, and the remaining 600 images of each camera are treated as the test samples. To simulate the re-compression manipulations, five typical quality of JPEG images are investigated in our experiments, including the original JPEG, which means the initial JPEG quality factors used in the camera, and the standard quality factors of 100, 90, 80 and 70.

3.2 Experimental Result and Analysis

The LBP method is used as the baseline in our experiments. Table 3 shows the average identification accuracies for the different training models with various quality images, and the different quality test images. It is easy to draw a conclusion that the classification reaches the highest identification accuracy when the training model matches the test image in the matter of image quality, as the diagonal elements indicate. For instance, with the model trained by the original JPEG images, a high accuracy of 94.04% is achieved for the original test samples. In the mismatching case, such as the re-compression quality factor of 100, the average accuracy decrease to 90.63%. While for the JPEG quality factor of 80 and 70, the classifier is considered to be out of work, as the accuracies drop to 47.29% and 33.92% (Table 2).

Table 2. Average accuracies of different quality images for the baseline of LBP.

Training test	Original	100	90	80	70
Original	94.04	90.63	82.04	47.29	33.92
100	89.38	93.54	69.54	41.88	32.46
90	66.04	61.67	93.96	43.46	30.83
80	52.54	45.96	58.38	86.13	59.54
70	44.58	37.58	46.33	63.67	80.50

In our following experiments, we use the raw images as the training set, which is widely considered as the best strategy in the practical scenario. With the CCA based algorithm, the confusion matrixes for re-compression qualities of 100, 90, 80 and 70 are shown in Tables 3, 4, 5 and 6.

Table 3. Confusion matrix of CCA method for quality factor of 100.

	AGF	KOD	PET	SON
AGF	93.00	3.33	1.17	2.33
KOD	2.67	93.33	1.17	3.00
PET	1.50	0.17	97.50	0.83
SON	6.83	5.33	0.83	87.00

Table 4. Confusion matrix of CCA method for quality factor of 90.

	AGF	KOD	PET	SON
AGF	75.83	13.17	2.83	8.17
KOD	8.17	86.50	0.83	4.67
PET	2.00	1.17	96.83	0.33
SON	5.33	5.00	3.17	87.00

Table 5. Confusion matrix of CCA method for quality factor of 80.

	AGF	KOD	PET	SON
AGF	47.00	25.83	19.17	8.00
KOD	13.50	58.50	6.00	22.00
PET	9.33	5.33	76.83	8.50
SON	23.50	20.67	3.00	52.83

Table 6. Confusion matrix of CCA method for quality factor of 70.

	AGF	KOD	PET	SON
AGF	34.50	21.50	33.00	11.00
KOD	10.50	59.50	7.17	22.83
PET	13.50	11.67	67.33	7.50
SON	21.00	39.33	8.50	31.17

Similarly, Tables 7, 8, 9, 10 and 11 show the details of the experimental results of ICA based algorithm for various JPEG quality factors.

Table 7. Confusion matrix of ICA method for quality factor of 100.

	AGF	KOD	PET	SON
AGF	94.17	3.00	0.67	2.17
KOD	2.50	95.00	1.00	1.50
PET	1.17	1.17	97.50	1.17
SON	6.33	10.50	0.67	82.50

Table 8. Confusion matrix of ICA method for quality factor of 90.

	AGF	KOD	PET	SON
AGF	69.83	4.83	5.33	20.00
KOD	2.00	85.00	1.00	12.00
PET	1.17	1.33	97.17	0.33
SON	1.33	1.17	5.83	91.67

Table 9. Confusion matrix of ICA method for quality factor of 80.

	AGF	KOD	PET	SON
AGF	19.50	16.00	38.67	25.83
KOD	22.67	38.50	15.33	23.50
PET	2.83	2.83	88.17	6.17
SON	31.33	23.00	6.83	38.33

Table 10. Confusion matrix of ICA method for quality factor of 70.

	AGF	KOD	PET	SON
AGF	19.67	2.67	71.83	5.83
KOD	34.50	26.83	25.17	13.50
PET	16.00	4.17	78.00	1.83
SON	35.83	12.00	28.33	28.33

Table 11. Confusion matrix of CCA + ICA method for quality factor of 100.

	AGF	KOD	PET	SON
AGF	93.83	2.83	1.00	2.33
KOD	2.33	94.17	1.17	2.33
PET	1.17	0	97.67	1.17
SON	6.83	4.50	0.67	88.00

For comparison, a combination of CCA and ICA base algorithm is also evaluated with the same image data set and experimental parameters. The confusion matrixes are shown in Tables 11, 12, 13 and 14.

Table 12. Confusion matrix of CCA + ICA method for quality factor of 90.

	AGF	KOD	PET	SON
AGF	67.67	14.00	4.83	13.50
KOD	3.00	92.67	–	4.33
PET	1.50	1.33	97.00	0.17
SON	1.67	3.17	2.33	92.83

Table 13. Confusion matrix of CCA + ICA method for quality factor of 80.

	AGF	KOD	PET	SON
AGF	27.83	33.83	22.50	15.83
KOD	6.83	39.17	7.67	46.33
PET	5.33	2.67	80.67	11.33
SON	6.00	8.50	2.83	82.67

Table 14. Confusion matrix of CCA + ICA method for quality factor of 70.

	AGF	KOD	PET	SON
AGF	18.17	11.33	50.50	20.00
KOD	7.00	41.00	11.00	41.00
PET	20.50	12.83	51.50	15.17
SON	16.50	29.00	10.50	44.00

By investigating all of the confusion matrixes of CCA, ICA and the combination of CCA and ICA, we can find that the classification performance can be promoted in different degree. For instance, the CCA, ICA and combination reaches 92.83%, 92.29% and 94.64% respectively for the image quality factor of 100, compared with the accuracy of 90.63% of the baseline, as Table 15 shows. Meanwhile for the quality factors of 100 and 90, the combination of CCA and ICA obtains the best results of 94.64% and 87.54%. But for the low quality factors of 80 and 70, the CCA based algorithm shows the highest accuracies of 58.79% and 48.13%. A further analysis indicates that the increasing inaccurate tags labeled in the ICA based algorithm have negative effects on the combination of CCA and ICA.

Table 15. Comparison of the proposed algorithms and the baseline.

		LBP	CCA	ICA	CCA + ICA
Quality	100	90.63	92.83	92.29	94.64
	90	80.24	86.54	85.91	87.54
	80	47.29	58.79	46.25	57.58
	70	33.92	48.13	37.08	38.67

4 Conclusion

This paper focused on identifying the camera source of image with different JPEG quality re-compression manipulations. Inspired by transfer learning, cross-class alignment and inter-class alignment based algorithms are presented. Experiments indicate that the proposed CCA, ICA and the combination outperform the baseline. In the case of re-compression quality factors of 100 and 90, the average accuracies of 94.64% and 87.54% are achieved by the combination algorithm. Meanwhile for the quality factors of 80 and 70, decent accuracies of 58.79% and 48.13% are shown respectively. Although for the quality factors of 100 and 90, the algorithms have good performance, the accuracies achieved in the case of re-compression quality factors of 80 and 70 illustrate the algorithms should be improved.

Acknowledgement. This work is supported by the National Science Foundation of China (No. 61502076) and the Scientific Research Project of Liaoning Provincial Education Department (No. L2015114).

References

1. Piva, A.: An overview on image forensics. *ISRN Sig. Process.* **2013**, 1–22 (2013)
2. Wang, B., Yang, F.: An overview and trends on digital image source forensics. *J. Inf. Secur. Res.* **2**(5), 501–511 (2016)
3. Mckay, C., Swaminathan, A., Gou, H., et al.: Image acquisition forensics: forensic analysis to identify source. In: *IEEE International Conference on Acoustics, Speech and Processing*, pp. 1657–1660. IEEE (2008)

4. Deng, Z., Gijsenij, A., Zhang, J.: Source camera identification using auto-white balance approximation. In: 2011 IEEE International Conference on Computer Vision (ICCV). IEEE, pp. 57–64 (2011)
5. Xu, G., Shi, Y.Q.: Camera model identification using local binary patterns. In: Proceedings IEEE International Conference on Multimedia and Expo (ICME), pp. 392–397. IEEE (2012)
6. Popescu, A.C., Farid, H.: Exposing digital forgeries in color filter array interpolated images. *IEEE Trans. Signal Process.* **53**(10), 3948–3959 (2005)
7. Bo, W., Yin, J.F., Li, Y.B.: Analysis of the effect form the JPEG compression on source camera model forensics. *J. Netw. Inf. Secur.* (2016)
8. Zhang, F.: Survey on transfer learning research. *J. Softw.* **26**(1), 26–39 (2015)
9. Arnold, A., Nallapati, R., Cohen, W.W.: A comparative study of methods for transductive transfer learning. In: International Conference on Data Mining Workshops, pp. 77–82 (2007)
10. Li, X., et al.: Generalized transfer component analysis for mismatched JPEG steganalysis. In: IEEE International Conference on Image Processing, pp. 4432–4436 (2014)
11. Avcibas, I.: Steganalysis using image quality metrics. *IEEE Trans. Image Process.* **12**(2), 221–229 (2003)
12. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)
13. Chang, C.-C.: LIBSVM: a library for support vector machine. *ACM Trans. Intell. Syst. Technol.* **2**(3), 1–17 (2011)