

How to identify fake images? : Multiscale methods vs. Sherlock Holmes

Minsu Park^a, Minjeong Park^b, Donghoh Kim^{1,c}, Hajeong Lee^d, Hee-Seok Oh^e

^aDepartment of Information and Statistics, Chungnam National University, Korea;

^bStatistical Research Institute, Statistics Korea, Korea;

^cDepartment of Mathematics and Statistics, Sejong University, Korea;

^dDepartment of Internal Medicine, Seoul National University Hospital, Korea;

^eDepartment of Statistics, Seoul National University, Korea

Abstract

In this paper, we propose wavelet-based procedures to identify the difference between images, including portraits and handwriting. The proposed methods are based on a novel combination of multiscale methods with a regularization technique. The multiscale method extracts the local characteristics of an image, and the distinct features are obtained through the regularized regression of the local characteristics. The regularized regression approach copes with the high-dimensional problem to build the relation between the local characteristics. Lytle and Yang (2006) introduced the detection method of forged handwriting via wavelets and summary statistics. We expand the scope of their method to the general image and significantly improve the results. We demonstrate the promising empirical evidence of the proposed method through various experiments.

Keywords: image analysis, intraclass correlation, ridge regression, wavelet packet analysis

1. Introduction

Nowadays, imitation for various products such as handwriting documents and famous artworks is prevalent. It is hard to identify fakes from the original one, and acrimonious disputes even happen to argue which one is genuine. Artistic professionals, experts or detectives may identify the authenticity of counterfeit products. The famous detective Holmes may solve the case using his renowned skills of astute observation and deductive reasoning with a magnifying glass. However, this procedure is subjective and is based on the accumulated experience of human experts. Human mainly uses the five senses of sight, hearing, smell, taste and touch to perceive the environment. Among the five senses, sight is the most powerful one. It is known that the majority of brain activity is involved with image treatment. A famous Chinese philosopher, Confucius, said, “A picture is worth a thousand words,” which means an image is very effective in capturing the features of products. The new technology makes it possible to record the visual observation of an image. The skills of observation, such as cameras, telescopes and microscopes, can compensate for the limitations of the human eyes (Dougherty, 2009). The development of computers also contributes to the expansion of image analysis. Computers can be utilized for the complex analysis of image data. These tasks include image segmentation,

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MIST) (No. 2021R1C1C1009976, 2018R1D1A1B07042933, and 2021R1F1A1047506).

¹ Corresponding author: Department of Mathematics and Statistics, Sejong University, 209, Neungdong-ro, Gwangjin-gu, Seoul 05006, Korea. E-mail: donghoh.kim@gmail.com

Published 30 November 2021 / journal homepage: <http://csam.or.kr>

© 2021 The Korean Statistical Society, and Korean International Statistical Society. All rights reserved.

object recognition, 3-D feature analysis and many other tasks associated with the broad area of visual observation (Hesamian *et al.*, 2019; Kasturi and Trivedi, 1990). The various methods have been developed to analyze an image objectively.

For handwriting analysis, besides the information on the writing tools, materials, self-expression, spacing and shape of stroke, various features such as the character form, slant and size of writing have been used to identify the forgeries. For the objective method, the distinct features can be extracted from an image through the statistical methods, and those features can be utilized. Lyu *et al.* (2004) proposed a method based on wavelet coefficients and basic statistics to distinguish the genuineness of artworks from fakes. Their approach provides a desirable direction to develop an objective statistical procedure. However, the approach of Lyu *et al.* (2004) is limited in that only the partial part of an image is used for their analysis. Lytle and Yang (2006) proposed a wavelet-based method considering the relation between wavelet coefficients by linear regression. Their approach has a restriction in that the number of predictors is not enough to find proper relation between the wavelet coefficients.

In this paper, we propose novel methods based on multiscale methods such as wavelet and wavelet packet transform coupled with a regularized regression approach. The essential features that distinguish the proposed method from existing ones are three-fold: (a) A multiscale framework of wavelets and wavelet packets is adapted to deal with images with various types of spatial structure. The flexible multiscale framework effectively captures the spatial characteristics of an image and extends the applicable scope of an image. (b) A regularized regression technique is applied to find proper relation between wavelet coefficients with a large number of predictors, and at the same time, to avoid overfitting problems of high-dimensional data. (c) We extract the distinct feature of an image through a regularized regression.

The rest of this paper is organized as follows. In Section 2, we briefly review wavelets and wavelet packets as multiscale methods and ridge regression as a regularized regression technique. The proposed methods are described in Section 3. In Section 4, we conduct numerical studies to investigate the empirical performance of the proposed method. Finally, conclusions are addressed in Section 5.

We remark that in literature various approaches have been proposed to detect similarity between images. Aljanabi *et al.* (2018) proposed the information-theory approach based on the entropies. The similarity and recognition measures between images were defined as the distance between two images, and were combined with a joint histogram to compare similarity between images. Chechik *et al.* (2010) utilized the approaches based on learning algorithms. The learning algorithm minimizes the hinge loss function with fast processing speed and accuracy. Recently, deep learning-based methods have been studied. Appalaraju and Chaoji (2017) used a multi-scale convolutional neural network (CNN) in a deep Siamese network, and trained the difference between the images by considering the distance of dichotomous image pairs.

2. Backgrounds

To make the present paper to be a self-contained material, we briefly discuss wavelets and wavelet packets, and a regularized regression technique. In addition, the intraclass correlation coefficient is introduced to measure the similarity between two groups of observations.

2.1. Wavelets and wavelet packets

Wavelets and wavelet packets have the advantage of obtaining a variety of additional information according to the spatial position, orientation and scale through domain transform. The various information obtained from the transform is utilized for comparative analysis of image data.

Wavelets are a family of orthonormal basis functions with several useful properties that have been popularly used in various fields such as mathematics, statistics and engineering. For analyzing an image, we focus on two-dimensional wavelets which are constructed by taking tensor products of one-dimensional wavelets. By using wavelets, an image is decomposed according to spatial position, orientation and scale. These local properties reflecting the spatial characteristics will be adapted to analyze an image.

To perform a multiscale analysis for images, we generally use a “pyramidal” algorithm consisting of a series of filters. There are various types of filters, and for a simple example, Haar bivariate filters are used. Let L , H_v , H_h and H_d be low-pass, vertical high-pass, horizontal high-pass and diagonal high-pass filters of Haar wavelets, respectively (Addison, 2002; Nason, 2008). The four filters are given by

$$L = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad H_v = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}, \\ H_h = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}, \quad H_d = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

For a 2×2 matrix M of data

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

the wavelet coefficients for the data matrix M are obtained as

$$\begin{pmatrix} H_v(M) & H_d(M) \\ L(M) & H_h(M) \end{pmatrix},$$

where $L(M) = (a + b + c + d)/2$, $H_v(M) = (a - b + c - d)/2$, $H_h(M) = (a + b - c - d)/2$ and $H_d(M) = (a - b - c + d)/2$. The low-pass filter L produces smoothed wavelet coefficients, and the high-pass filters H_v , H_h and H_d represent detailed wavelet coefficients according to the vertical, horizontal and diagonal directions. Suppose that we have a 4×4 data matrix M . The wavelet-based filtering is sequentially performed. So, the data matrix M is partitioned as

$$M = \begin{pmatrix} M_1 & M_2 \\ M_3 & M_4 \end{pmatrix},$$

where the size of M_i ($i = 1, 2, 3, 4$) is 2×2 . The first filtering produces the wavelet coefficient matrix m as

$$m = \begin{pmatrix} H_v(M) & H_d(M) \\ L(M) & H_h(M) \end{pmatrix}, \quad (2.1)$$

where $L(M)$ is the smoothed wavelet coefficient matrix, and $H_v(M)$, $H_h(M)$ and $H_d(M)$ are the vertical (detailed) wavelet coefficient matrix, the horizontal (detailed) wavelet coefficient matrix and the diagonal (detailed) wavelet coefficient matrix, respectively. Each coefficient matrix is obtained as

$$L(M) = \begin{pmatrix} L(M_1) & L(M_2) \\ L(M_3) & L(M_4) \end{pmatrix}, \quad H_v(M) = \begin{pmatrix} H_v(M_1) & H_v(M_2) \\ H_v(M_3) & H_v(M_4) \end{pmatrix}, \\ H_h(M) = \begin{pmatrix} H_h(M_1) & H_h(M_2) \\ H_h(M_3) & H_h(M_4) \end{pmatrix}, \quad H_d(M) = \begin{pmatrix} H_d(M_1) & H_d(M_2) \\ H_d(M_3) & H_d(M_4) \end{pmatrix}.$$

The second filtering is applied to only the smoothed wavelet coefficient matrix $L(M)$. Then we obtain the wavelet coefficient matrix as

$$\begin{pmatrix} H_v(M) & H_d(M) \\ m_1 & H_h(M) \end{pmatrix},$$

where

$$m_1 = \begin{pmatrix} H_v[L(M)] & H_d[L(M)] \\ L[L(M)] & H_h[L(M)] \end{pmatrix}.$$

In general, for a $2^J \times 2^J$ matrix of data where J is a positive integer, the above filtering process is performed subsequently to the smoothed wavelet coefficient matrices.

Furthermore, we consider the wavelet packet transform, which performs filtering to all sub-bands, while the wavelet transform does not take any further filtering for detailed coefficients. For the coefficients of equation (2.1), the wavelet packet transform takes the four filtering procedures to all sub-bands as follows.

$$\begin{aligned} m_2 &= \begin{pmatrix} H_v[H_v(M)] & H_d[H_v(M)] \\ L[H_v(M)] & H_h[H_v(M)] \end{pmatrix}, \\ m_3 &= \begin{pmatrix} H_v[H_h(M)] & H_d[H_h(M)] \\ L[H_h(M)] & H_h[H_h(M)] \end{pmatrix}, \\ m_4 &= \begin{pmatrix} H_v[H_d(M)] & H_d[H_d(M)] \\ L[H_d(M)] & H_h[H_d(M)] \end{pmatrix}. \end{aligned}$$

Then the wavelet packet coefficients are given by

$$\begin{pmatrix} m_2 & m_4 \\ m_1 & m_3 \end{pmatrix}.$$

For a particular example of two-dimensional wavelet transform and wavelet packet transform, we consider box image in the top panel of Figure 1. By wavelet transform, the box image is decomposed according to eight resolution levels. The middle panel of Figure 1 shows the decomposition results of two most detailed levels in wavelet transform. The first filtering of the box image generates lowpass, vertical, horizontal and diagonal coefficient matrices that represent smooth part, vertical, horizontal and diagonal details of the image, respectively. The next scale of the decomposed image is created by the recursive filtering of the lowpass sub-band. On the other hand, to obtain additional information, the wavelet packet transform in the bottom panel of Figure 1 takes the filtering of four sub-bands (lowpass, vertical, horizontal, and diagonal) recursively for each coefficient matrix of the most detailed level.

2.2. Regularized regression

The linear regression model for $n \times 1$ observation vector \mathbf{y} is expressed as,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad (2.2)$$

where \mathbf{X} is $n \times (p + 1)$ design matrix with p predictors, $\boldsymbol{\beta}$ denotes a $p \times 1$ regression coefficient vector, and $\boldsymbol{\epsilon}$ represents an error term. A regularized regression minimizes the following loss function Q with a penalty term $R(\boldsymbol{\beta})$,

$$Q(\boldsymbol{\beta}) = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + \lambda R(\boldsymbol{\beta}). \quad (2.3)$$

Original Image

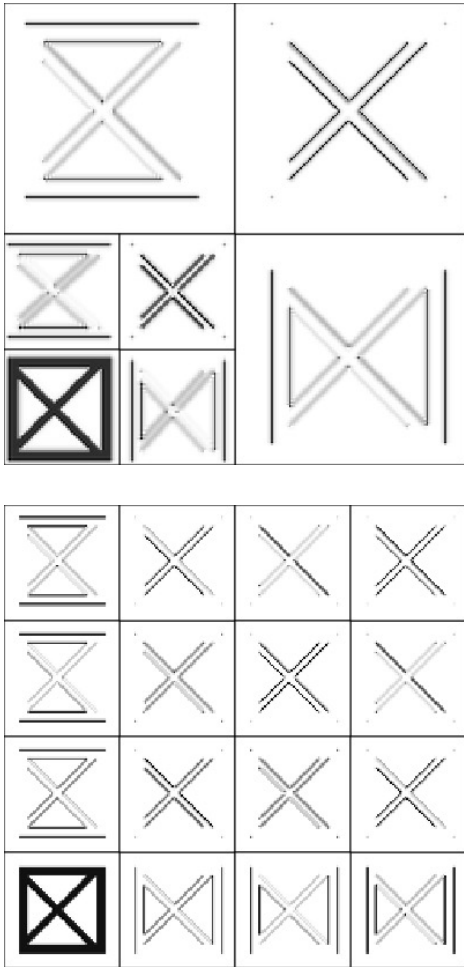


Figure 1: Box image (top) and its two-dimensional wavelet transform (middle) and its wavelet packet transform (bottom).

The parameter λ controls the amount of the regularization term $R(\beta)$, which imposes a penalty on the complexity of coefficient vector β . Ridge regression takes specific form $\sum_{j=1}^p \beta_j^2$ for the regularization term $R(\beta)$. As one can see, a model with $\lambda = 0$ is equivalent to a conventional linear regression model. Ridge regression is popularly used as a regularized regression technique since the regularization term makes the number of the model parameters small. Thus, ridge regression is suitable to handle a high-dimensional problem and to resolve over-fitting for complex models.

2.3. Intraclass correlation coefficient

The intraclass correlation coefficient (ICC) is a measurement of reliability index between individual ratings or measurements (Shrout and Fleiss, 1979). ICC is a modified statistic for the interclass correlation (Pearson correlation), and reflects not only a degree of relatedness but also an agreement between raters. ICC regards the observations for each rater as groups while Pearson correlation measures an association for the paired observations. It is commonly used to quantify the degree to which raters with assigning scores to the observation are similar to each other in terms of a quantitative trait. ICC can be estimated in terms of the random effects models

$$z_{ij} = \mu + \alpha_j + \epsilon_{ij},$$

where z_{ij} is the i^{th} observation in the j^{th} cluster, μ is the population mean as a constant, α_j is the cluster-specific random effect, and ϵ_{ij} is an error term. Then, followed by Shrout and Fleiss (1979), ICC is estimated by $\hat{\rho}$ as,

$$\hat{\rho} = \frac{\hat{\sigma}_\alpha^2}{\hat{\sigma}_\alpha^2 + \hat{\sigma}_\epsilon^2}, \quad (2.4)$$

where $\hat{\sigma}_\alpha^2$ and $\hat{\sigma}_\epsilon^2$ are estimated variances for random effects and errors, respectively. ICC has a range from 0 to 1 and indicates that the raters in each cluster are more similar to each other as it is closer to the value 1.

3. The proposed method

It is necessary to extract a distinct feature representing the characteristics of a given image. The wavelet and wavelet packet transform decompose an image according to the scale, and the detailed scale reflects the subtle, hidden structure of an image. Thus, we focus on the wavelet coefficients of the two most detailed scales and a distinct feature will be extracted based on the relation between these wavelet coefficients. In the proposed algorithm, two-dimensional discrete wavelet transform is implemented according to the Mallat's pyramidal algorithm with Daubechies' least-asymmetric wavelets (Mallat, 1989a; Mallat, 1989b). For a given image, we derive certain relation among wavelet coefficients of smooth part and detailed parts of an image. If this relation changes, we can deduce that an image is modified. The regression is used to model the relation among wavelet coefficients. The wavelet coefficients of a block in the most detailed level will be modeled by the wavelet coefficients of surrounding blocks at the same level and the wavelet coefficients in the next coarser level. To find a proper relation between the coefficients, the ridge regression is adapted to cope with the high-dimension problem at this step. The regression coefficient estimates and residuals are used to assess the quality of the regression model, and the skewness based on the estimates and residuals will represent the characteristics of the given image. If the skewness is similar to each other for given two images, then it is identified that two images are identical. The similarity is measured by ICC of

skewness. Suppose that we have two images of interest. The proposed method is implemented by the following steps when wavelet packet transform is used.

1. Feature extraction : For each image, a distinct feature is extracted.

- (1) (Transformation) Take wavelet packet transform (wavelet transform) of an image and then obtain the coefficients corresponding to each sub-band of two most detailed levels.
- (2) (Regression) The coefficients corresponding to each sub-band of the most detailed level are partitioned into certain blocks. Let \mathbf{y} be wavelet packet coefficients of a specific block in the most detailed level. Fit ridge regression minimizing the loss function (2.3), where the predictors of the design matrix \mathbf{X} of model (2.2) are formed by coefficients of neighboring blocks in the same level and coefficients in the next coarser level. Repeat the regression for all the partitioned blocks.
- (3) (Measurement) Obtain the estimated regression coefficients and residuals. Let $\hat{\beta}_s$ and \mathbf{e}_s be the vector of the estimated regression coefficients and residuals from all blocks of the smooth part. For detailed part, denote the corresponding vectors as $\hat{\beta}_v$ and \mathbf{e}_v for vertical direction, $\hat{\beta}_h$ and \mathbf{e}_h for horizontal direction, and $\hat{\beta}_d$ and \mathbf{e}_d for diagonal direction. Then, compute skewness δ for each vector and obtain eight features as

$$\delta\hat{\beta}_s, \delta\hat{\beta}_v, \delta\hat{\beta}_h, \delta\hat{\beta}_d, \delta\mathbf{e}_s, \delta\mathbf{e}_v, \delta\mathbf{e}_h, \delta\mathbf{e}_d. \quad (3.1)$$

2. Comparison : The similarity between two images is judged by eight features.

- (1) (Similarity) For example, among eight features, consider the skewness $\delta\hat{\beta}_v^1$ and $\delta\hat{\beta}_v^2$ based on vertical detailed coefficients for given two images. Regard two feature vectors $\delta\hat{\beta}_v^1$ and $\delta\hat{\beta}_v^2$ as the observations from raters in two clusters, and estimate ICC ρ by equation (2.4). For each feature in equation (3.1), ICC is estimated.
- (2) (Testing) Perform eight tests under each null hypothesis $H_0 : \rho = 0$, and adjust p -values by the Bonferroni correction. As the number of rejected tests increases, the similarity of two images is bigger.

Since the feature extraction step plays a crucial role, we explain the procedure in detail when wavelet packet transform is used. Suppose that we have a $2^J \times 2^J$ image, where J is a positive integer. The two-dimensional wavelet packet transform produces four kinds of wavelet packet coefficient matrices including smooth part (S) and vertical (V), horizontal (H) and diagonal (D) detailed part as in Figure 2. Note that the resolution level is from the finest level $J - 1$ to the coarse level 0. Then, we divide each $2^{J-1} \times 2^{J-1}$ wavelet packet coefficient matrix at the finest level $J - 1$ into several blocks. For example, in the case that the number of blocks is $16 = 2^2 \times 2^2$, the size of each block is $2^{J-3} \times 2^{J-3}$. At the resolution level $J - 2$, four blocks are constructed for each smooth and detailed parts, as shown in Figure 2. For identifying the relation between wavelet packet coefficients, the regression model (2.2) is used. We consider the wavelet packet coefficients of a particular block V_C as the response \mathbf{y} . The wavelet packet coefficients of neighboring blocks at levels $J - 1$ and the wavelet packet coefficients of all the blocks at level $J - 2$ are covariates that consist of the design matrix \mathbf{X} of model (2.2). By the ridge regression, the predicted values $\hat{\mathbf{y}}$ and regression coefficient estimates $\hat{\beta}_v$ are obtained. In addition, the residuals are stabilized as $\mathbf{e}_v = \log_2(|\mathbf{y}|) - \log_2(|\hat{\mathbf{y}}|)$. The regression coefficients may represent the relation between wavelet packet coefficients for a specific block and its neighboring blocks,

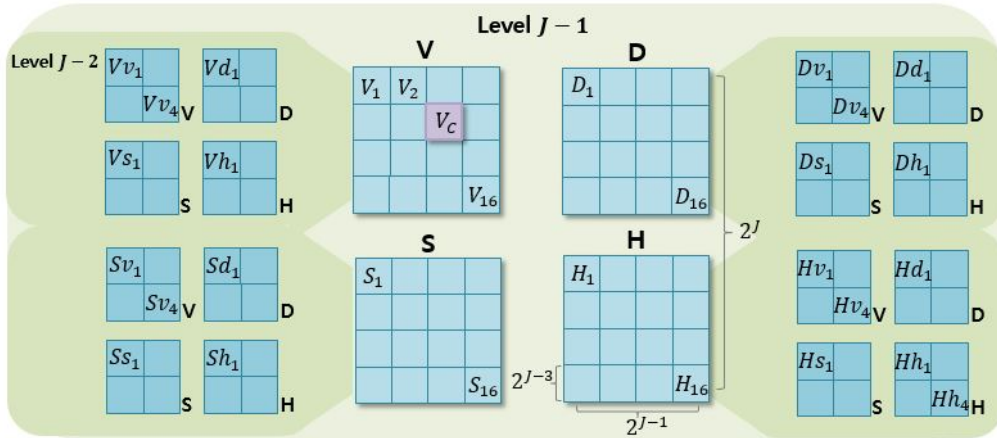


Figure 2: Partitions of wavelet packet coefficient matrices.

and the residuals capture the unusual properties, such as outliers or influential points. Subsequently, the skewness is calculated for the regression coefficient estimates and the residuals. This procedure is repeated for all the blocks. Then, the resulting skewness $\delta_{\hat{\beta}_v}$ and δ_{e_v} is a vector with size 16 when we partition the vertical wavelet packet coefficient matrix as 16 blocks, as shown in Figure 2. Finally, by applying the model (2.2) to other detailed wavelet packet coefficient matrices and the smooth wavelet packet coefficient matrix, we obtain the eight features of skewness in equation (3.1). Lytle and Yang (2006) remarked that the skewness rather than variance or kurtosis is more useful to represent the feature of an image. The skewness rather than variance or kurtosis adequately represents the distinct features such as the extreme values and asymmetric distribution.

4. Numerical study

This section reports the case study results to assess the empirical performance of the proposed method. The proposed algorithm is supposed to identify the difference between two images. For the analysis, we consider three numerical examples, including a simulated image, portrait image and handwriting image.

For comparison, the following five methods are studied:

1. WTR: the proposed method with wavelet transform and ridge regression,
2. WTL: the proposed method with wavelet transform and conventional regression,
3. WPR: the proposed method with wavelet packet transform and ridge regression,
4. WPL: the proposed method with wavelet packet transform and conventional regression and
5. WT3: the algorithm of Lytle and Yang (2006).

The performance of the methods is measured by the number of rejected tests with the null hypothesis that ICC is zero under a significance level of 0.05. Note that the proposed method uses wavelet coefficients of both smooth part and detailed parts while the algorithm suggested by Lytle and Yang (2006) is based on the skewness only for three directions of detailed wavelet coefficients. Thus, WT3

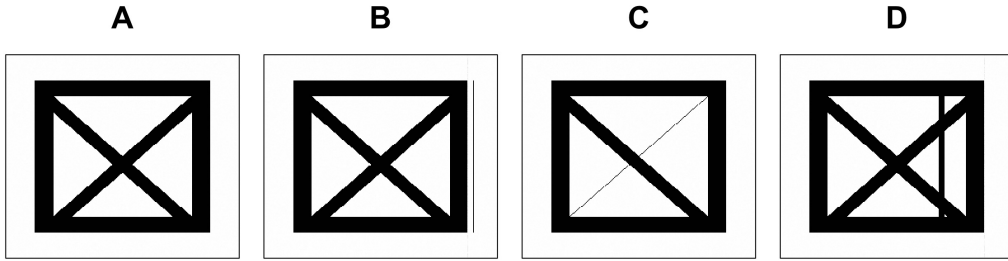


Figure 3: Four simulated box images.

Table 1: The number of rejected tests and the mean of ICCs inside parenthesis for box images

	WTR	WTL	WPR	WPL	WT3
A vs. B	8 (0.823)	8 (0.845)	8 (0.945)	8 (0.927)	6
A vs. C	4 (0.492)	2 (0.470)	0 (0.199)	0 (0.199)	5
B vs. C	3 (0.402)	1 (0.379)	0 (0.159)	0 (0.195)	5
C vs. D	1 (0.326)	1 (0.258)	0 (0.114)	0 (0.053)	2

performs six tests to judge the similarity, while the proposed method runs eight tests. For this study, images of size 256×256 are used, and 16 blocks for each sub-band are employed.

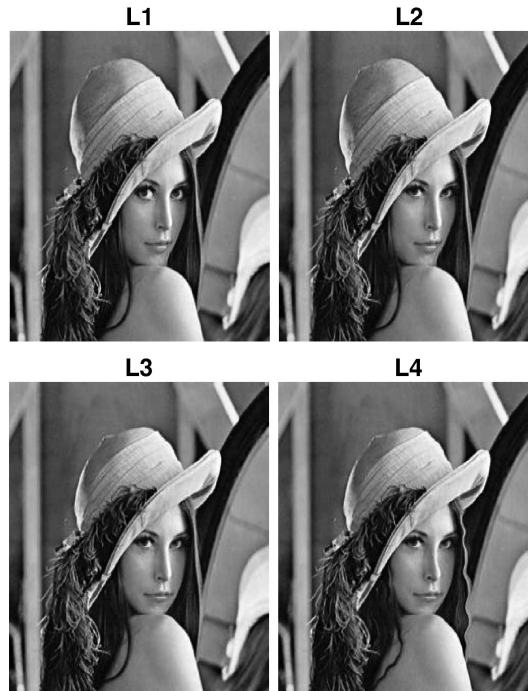
4.1. Simulated box image

Figure 3 shows four different box images in which the thickness of the diagonal lines is different, or there exists additional line. Box image **A** has the thick diagonal lines, and image **B** is created by adding an extra thin line to the right of image **A**. Image **C** is imitated by thinning the diagonal line of image **A**, and image **D** is generated by adding a thick vertical line inside box **A**. The simulated images are generated in a way that the difference between images **A** and **B** is negligible, and the differences between image **C** and other images are apparent. For each pair of two images, similarity tests are implemented by the five methods and the number of rejected tests is given in Table 1. For reference, the mean of ICC values is reported inside parenthesis. Note that WT3 procedure is not based on ICC.

The proposed methods correctly identify that the difference between images **A** and **B** is negligible. For pairs of image **C** and other images, the ICC mean values of WPR and WPL are smaller compared to the ICC mean values of WTR and WTL methods and test results of WPR and WPL imply that the two images are different, while it seems that WT3 is not effective to identify these differences since image **C** is not clearly distinguishable from other images. From the results, the proposed algorithm based on the wavelet packet transform yields the most reliable result when the image is changed by line thickness or additional line.

4.2. Portrait image

Image **L1** of Figure 4 is Lena image. Images **L2**, **L3** and **L4** are constructed by gradually changing face, hair and hat of image **L1**. More specifically, the eyes of image **L2** are smaller than the original image. For image **L3**, the hat size is decreased, and the chin and lip are increased with the changes made in image **L2**. Finally, along with the previous cumulative changes, the ridge of the nose holds highly and hair is twisted in image **L4**. The same procedure is performed for each pair of image **L1** and other images. The results are listed in Table 2. WT3 fails to detect the difference of Lena images regardless of the severity of modification. The number of rejected tests is the same as six, which implies

Figure 4: *Lena* image and its modified images.Table 2: The number of rejected tests and the mean of ICCs inside parenthesis for *Lena* images

	WTR	WTL	WPR	WPL	WT3
L1 vs. L2	7 (0.734)	7 (0.737)	7 (0.782)	7 (0.765)	6
L1 vs. L3	6 (0.612)	4 (0.578)	4 (0.528)	5 (0.531)	6
L1 vs. L4	6 (0.668)	3 (0.550)	3 (0.573)	3 (0.546)	6

that the two images are identical. On the other hand, the proposed methods show the improved result that the number of rejected tests is decreased as the severity of modification is increased. The proposed method by coupling wavelet packet transform and ridge regression model provides reasonable results concerning the degree of image alteration. For images **L1**, **L2** and **L3**, the ICC mean values of the proposed methods show the decreasing tendency as the severity of modification is increased.

4.3. Handwriting analysis

Figure 5 displays Bengali alphabet samples of handwritings from the database (<https://www.isical.ac.in/~ujjwal/download/Banglabasiccharacter.html>). This database provides 37,858 handwritten image samples for Bengali basic characters. We consider three characters of 'DA,' 'DDA' and 'DDHA'. Four handwriting images are chosen for comparison with two handwriting images **DA** by different people, an image **DDA** and an image **DDHA**. For all pairs of images, similarity tests are implemented. It is desirable to conclude that the difference of images **DA1** and **DA2** is negligible, and other pairs of images are different. Overall, the proposed method outperforms the existing method WT3 except for pairs of images **DA1** and **DA2**. The result of the method WT3 implies that two-character

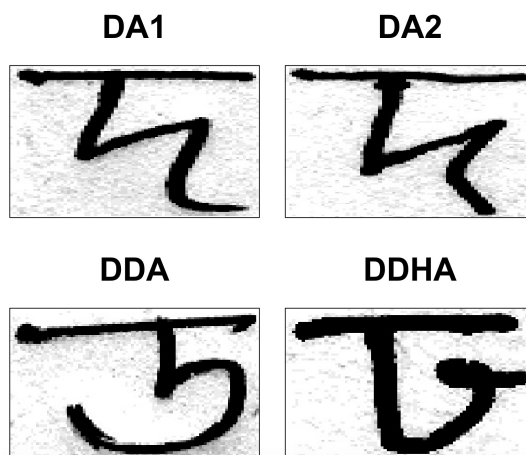


Figure 5: Four samples of handwriting

Table 3: The number of rejected tests and the mean of ICCs inside parenthesis for handwriting analysis

	WTR	WTL	WPR	WPL	WT3
DA1 vs. DA2	2 (0.274)	2 (0.268)	4 (0.323)	4 (0.472)	6
DA1 vs. DDA	0 (0.040)	0 (0.003)	0 (0.074)	0 (0.021)	0
DA1 vs. DDHA	0 (0.088)	0 (0.084)	0 (0.095)	1 (0.051)	0
DA2 vs. DDA	0 (0.081)	0 (0.030)	0 (0.130)	0 (0.135)	4
DA2 vs. DDHA	0 (0.039)	0 (0.077)	0 (0.005)	0 (0.038)	0
DDA vs. DDHA	0 (0.054)	0 (0.047)	0 (0.024)	0 (0.002)	3

images **DA1** and **DA2** written by different people are identical, while the proposed method does not completely judge the similarity, as shown in Table 3. For images **DA1** and **DA2**, the ICC mean values and the number of rejected tests of **WPR** and **WPL** are larger compared to those of **WTR** and **WTL** methods, which implies that **WPR** and **WPL** are more efficient to identify the similarity of images **DA1** and **DA2**.

5. Concluding Remarks

In this paper, we proposed the multiscale-based method for identifying the discrepancies between different images. Its performance is evaluated through a simulated box image, portrait image and handwriting image. The proposed method is implemented by coupling wavelet (packet) transform with regression. The proposed method is efficient in the simulation study when the regression model is coupled with wavelet packet transform rather than wavelet transform. The results from numerical experiments suggest that the proposed method possesses promising empirical properties.

We remark that the regularized regression approaches cope with the high-dimensional problem at the feature extraction step to resolve over-fitting for complex models, and the proposed procedure can be extended utilizing the various regularized regression approaches. The block size for the regression must be pre-determined for the proposed algorithm. The optimal block size for an image may improve its performance. Furthermore, it is possible to expand the algorithm to compare the multiple images by adapting multiple comparison procedure for the test of ICC. We leave these issues for future research since it is beyond the scope of this paper.

References

- Addison PS (2002). *The Illustrated Wavelet Transform Handbook*, Napier University, Edinburgh.
- Aljanabi MA, Hussain ZM, and Lu SF (2018). An entropy-histogram approach for image similarity and face recognition, *Mathematical Problems in Engineering*, **2018**, 1–18.
- Appalaraju S and Chaoji V (2017). *Image Similarity Using Deep CNN and Curriculum Learning*, arXiv preprint arXiv:1709.08761, 1–9.
- Chechik G, Sharma V, Shalit U, and Bengio S (2010). Large scale online learning of image similarity through ranking, *Journal of Machine Learning Research*, **11**, 1109–1135.
- Dougherty G (2009). *Digital Image Processing for Medical Applications*, California State University, Channel Islands.
- Hesamian MH, Jia W, He X, and Kennedy P (2019). Deep learning techniques for medical image segmentation: achievements and challenges, *Journal of Digital Imaging*, **32**, 582–596.
- Kasturi R and Trivedi MM (1990). *Image Analysis Applications*, Marcel Dekker, New York.
- Lytle B and Yang C (2006). Detecting forged handwriting with wavelets and statistics, *Rose Hulman Undergraduate Mathematics Journal*, **7**, 1–10.
- Lyu S, Rockmore D, and Farid H (2004). A digital technique for art authentication. In *Proceedings of the National Academy of Sciences*, **101**, 17006–17010.
- Mallat SG (1989a). Multifrequency channel decompositions of images and wavelet models. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **37**, 2091–2110.
- Mallat SG (1989b). A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, 674–693.
- Nason GP (2008). *Wavelet Methods in Statistics with R*, Springer-Verlag, New York.
- Shrout PE and Fleiss JL (1979). Intraclass correlations: uses in assessing rater reliability, *Psychological Bulletin*, **86**, 420–428.

Received March 21, 2021; Revised September 12, 2021; Accepted September 23, 2021