

Block Size Forensic Analysis in Digital Images

Steven Tjoa, W. Sabrina Lin, H. Vicky Zhao, K. J. Ray Liu

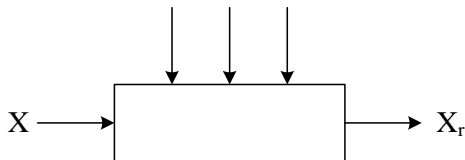
Signals and Information Group
Dept. of Electrical and Computer Engineering
University of Maryland

April 20, 2007



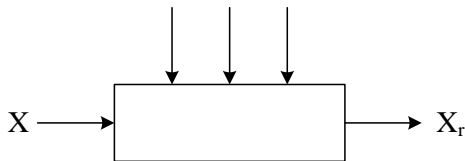
- Multimedia security is typically extrinsic in nature.

Keys, Watermarks, etc. (Extrinsic Information)



- Multimedia security is typically extrinsic in nature.

Keys, Watermarks, etc. (Extrinsic Information)



- Alternative: Examine the intrinsic information of a signal.



(Use the intrinsic information in X_r .)

- Use *non-intrusive forensic analysis*.

Existing work: forgery detection; component forensics.

Existing work: forgery detection; component forensics.

New problem: Source Coder Forensics for Digital Images.

- Given an uncompressed image, perform forensic analysis upon the image.

Existing work: forgery detection; component forensics.

New problem: Source Coder Forensics for Digital Images.

- Given an uncompressed image, perform forensic analysis upon the image.
- What compression has been performed, if any?
 - Was any preprocessing performed, such as blocking?
 - What transform method, if any, was used?
 - What is the quantization scheme?

Existing work: forgery detection; component forensics.

New problem: Source Coder Forensics for Digital Images.

- Given an uncompressed image, perform forensic analysis upon the image.
- What compression has been performed, if any?
 - Was any preprocessing performed, such as blocking?
 - What transform method, if any, was used?
 - What is the quantization scheme?
- What are the parameters of the coding scheme?

Existing work: forgery detection; component forensics.

New problem: Source Coder Forensics for Digital Images.

- Given an uncompressed image, perform forensic analysis upon the image.
- What compression has been performed, if any?
 - Was any preprocessing performed, such as blocking?
 - What transform method, if any, was used?
 - What is the quantization scheme?
- What are the parameters of the coding scheme?
- Establish a confidence measure.

- **Digital Image Integrity.** Picture content is of crucial importance in many scenarios. Verify the similarity between the digital photograph and the actual scene.

Motivation

- **Digital Image Integrity.** Picture content is of crucial importance in many scenarios. Verify the similarity between the digital photograph and the actual scene.



Q: "Was the image compressed?"

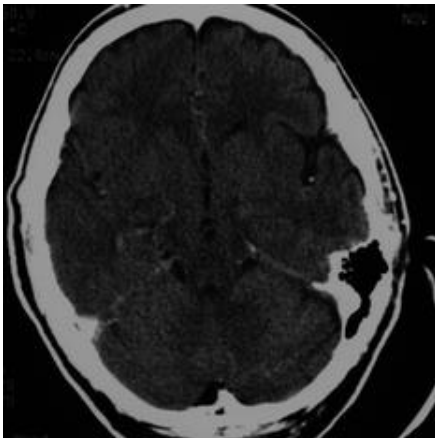
Motivation

- **Digital Image Integrity.** Picture content is of crucial importance in many scenarios. Verify the similarity between the digital photograph and the actual scene.

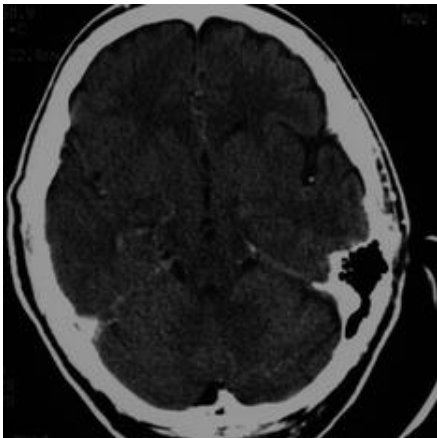


Q: "Was the image compressed?"

A: "Yes. As a result, the artifacts obscure the true image content."



Q: "Was the image compressed?"



Q: "Was the image compressed?"

A1: "No. Therefore, I am 95% confident in my diagnosis."

A2: "Yes. Therefore, I am only 80% confident in my diagnosis."

- **Patent Infringement.** Determine a specific encoding mechanism.

- **Patent Infringement.** Determine a specific encoding mechanism.

UPDATE 1-Philips charges Kodak infringed on imaging patent

Wed Apr 19, 2006 11:10am ET

[Email This Article](#) | [Print This Article](#) | [Rep](#)

COMPANY NEWS:

- ▶ [EU mergers and takeovers \(Aug 14\)](#)
- ▶ [Garmin doubles Europe Q2 market share-Canalys](#)
- ▶ [EU mergers and takeovers \(Aug 11\)](#)

(adds details, company comment)

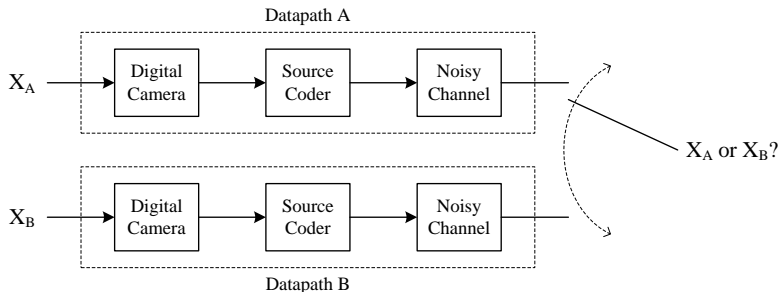
[\[-\] Text](#)

NEW YORK, April 19 (Reuters) - Dutch company Philips Electronics (PHG.AS: [Quote](#), [Profile](#), [Research](#)) on Wednesday said it sued Eastman Kodak Co. (EK.N: [Quote](#), [Profile](#), [Research](#)), claiming several of Kodak's cameras infringe on a patent related to the compression of digital images.

(“Philips charges Kodak infringed on imaging patent,” Reuters [Online], Apr. 2006.)

- **Datapath Integrity.** Validate the steps taken to process data, and assess the trustworthiness and reliability of these operations.

- **Datapath Integrity.** Validate the steps taken to process data, and assess the trustworthiness and reliability of these operations.



- **Quality Assurance.** Monitor and control image quality.

- **Quality Assurance.** Monitor and control image quality.
- **Rate Control.** Minimize the distortion given a rate constraint. Avoid unnecessary losses during recompression.

- **Quality Assurance.** Monitor and control image quality.
- **Rate Control.** Minimize the distortion given a rate constraint. Avoid unnecessary losses during recompression.
- **Restoration.** Obtain information about previous processing to customize restoration algorithms by reducing assumptions.

Detection of Pre-processing

Motivation:

- Detection of block processing is an obvious first step. Incorrect estimation could invalidate subsequent tests.

Detection of Pre-processing

Motivation:

- Detection of block processing is an obvious first step. Incorrect estimation could invalidate subsequent tests.
- Block artifact measurement is well-established. However, existing work imposes strong assumptions upon the input data.

Detection of Pre-processing

Motivation:

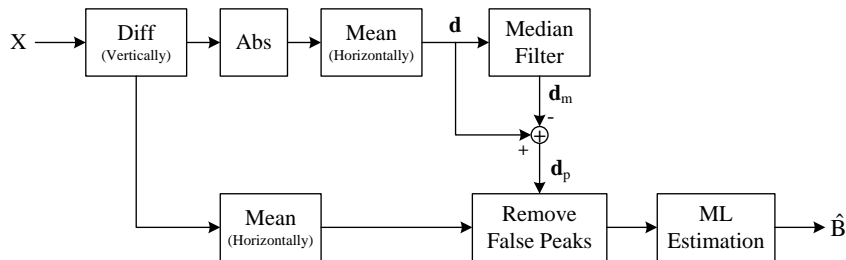
- Detection of block processing is an obvious first step. Incorrect estimation could invalidate subsequent tests.
- Block artifact measurement is well-established. However, existing work imposes strong assumptions upon the input data.
- We propose a novel scheme to detect the presence of block processing in an image, to estimate the block size, and to measure confidence.

Detection of Pre-processing

Motivation:

- Detection of block processing is an obvious first step. Incorrect estimation could invalidate subsequent tests.
- Block artifact measurement is well-established. However, existing work imposes strong assumptions upon the input data.
- We propose a novel scheme to detect the presence of block processing in an image, to estimate the block size, and to measure confidence.
- Idea: Extract a reduced feature – the *one-dimensional block artifact signature* – and analyze its periodicity.

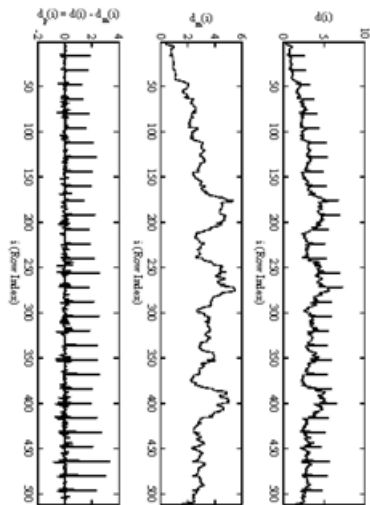
Algorithm



Repeat for the other direction to obtain both vertical and horizontal block size estimates.

- X : Image
- d_p : One-dimensional block artifact signature
- \hat{B} : Block size estimate

Median Filtering



Left: The block artifact signature, $d_p(i)$.

Removal of False Peaks

Will produce a
spurious peak.



ML Estimation of Block Size

Suppose $d_p(i)$ consists of a known periodic signal plus zero-mean i.i.d. Gaussian noise.

$$d_p(i) = s(i) + n(i)$$

Maximize $p(\mathbf{d}_p | \mathbf{s}, \sigma^2, B)$ with respect to the signal parameter \mathbf{s} , the noise variance σ^2 , and the period B .

ML Estimation of Block Size

Suppose $d_p(i)$ consists of a known periodic signal plus zero-mean i.i.d. Gaussian noise.

$$d_p(i) = s(i) + n(i)$$

Maximize $p(\mathbf{d}_p | \mathbf{s}, \sigma^2, B)$ with respect to the signal parameter \mathbf{s} , the noise variance σ^2 , and the period B .

- 1 Obtain the signal estimate $\hat{s}(i; B)$.

ML Estimation of Block Size

Suppose $d_p(i)$ consists of a known periodic signal plus zero-mean i.i.d. Gaussian noise.

$$d_p(i) = s(i) + n(i)$$

Maximize $p(\mathbf{d}_p | \mathbf{s}, \sigma^2, B)$ with respect to the signal parameter \mathbf{s} , the noise variance σ^2 , and the period B .

- 1 Obtain the signal estimate $\hat{s}(i; B)$.
- 2 Estimate the noise variance, $\hat{\sigma}^2(B)$

ML Estimation of Block Size

Suppose $d_p(i)$ consists of a known periodic signal plus zero-mean i.i.d. Gaussian noise.

$$d_p(i) = s(i) + n(i)$$

Maximize $p(\mathbf{d}_p | \mathbf{s}, \sigma^2, B)$ with respect to the signal parameter \mathbf{s} , the noise variance σ^2 , and the period B .

- 1 Obtain the signal estimate $\hat{s}(i; B)$.
- 2 Estimate the noise variance, $\hat{\sigma}^2(B)$
- 3 Estimate the block size:

$$\hat{B} = \underset{B}{\operatorname{argmin}} \hat{\sigma}^2(B)$$

ML Estimation of Block Size

Suppose $d_p(i)$ consists of a known periodic signal plus zero-mean i.i.d. Gaussian noise.

$$d_p(i) = s(i) + n(i)$$

Maximize $p(\mathbf{d}_p | \mathbf{s}, \sigma^2, B)$ with respect to the signal parameter \mathbf{s} , the noise variance σ^2 , and the period B .

- 1 Obtain the signal estimate $\hat{s}(i; B)$.
- 2 Estimate the noise variance, $\hat{\sigma}^2(B)$
- 3 Estimate the block size:

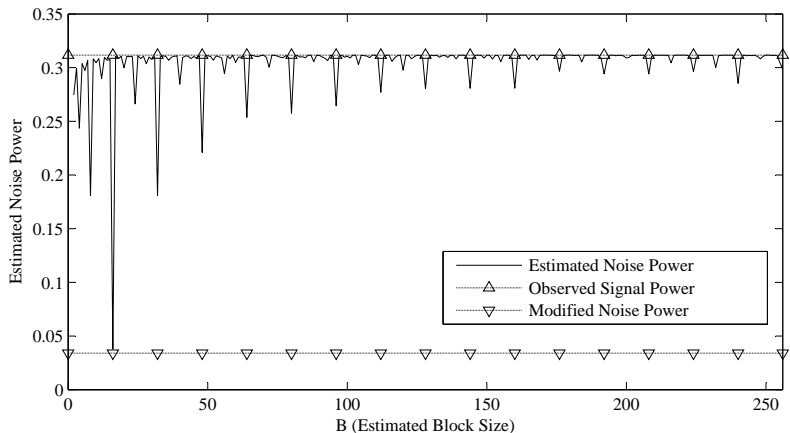
$$\hat{B} = \underset{B}{\operatorname{argmin}} \hat{\sigma}^2(B)$$

- 4 Calculate the observed signal-to-noise ratio:

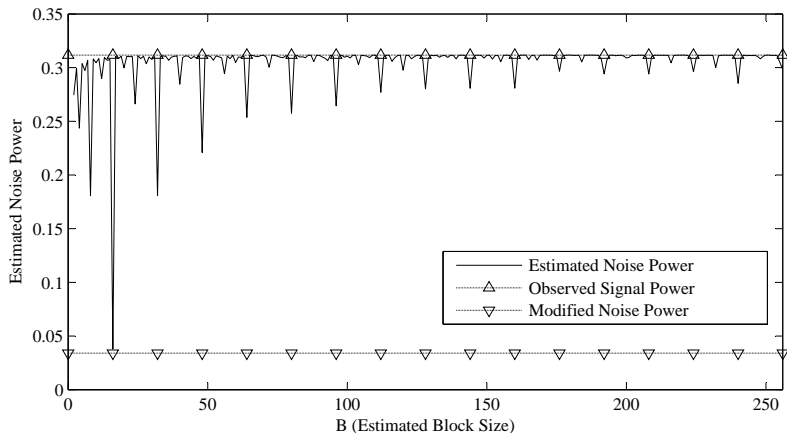
$$\text{OSNR} = \frac{P_{d_p}}{\sigma^2}$$

This is our measure of detection accuracy.

ML Estimation of JPEG-compressed *Lena* with a block size of 16:

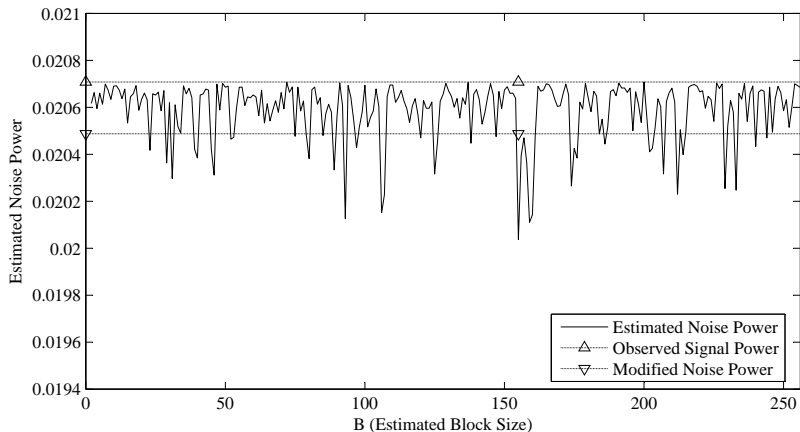


ML Estimation of JPEG-compressed *Lena* with a block size of 16:

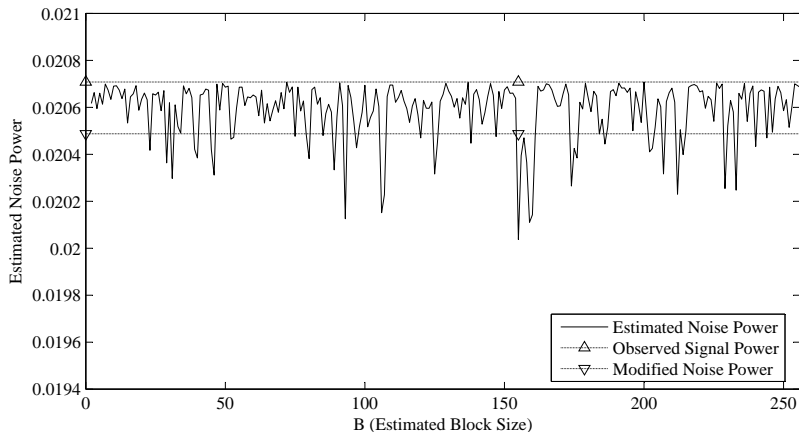


- The block size estimate is $\hat{B} = B = 16$.
- OSNR is high. Block processing is present.

ML Estimation of uncompressed *Lena*:



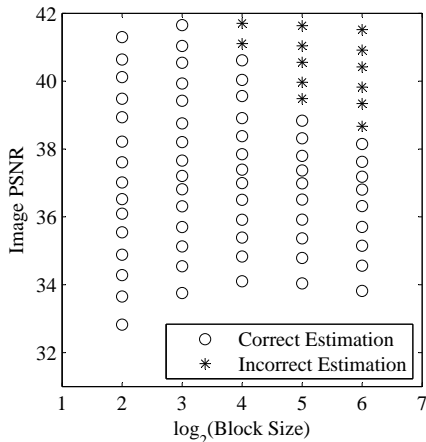
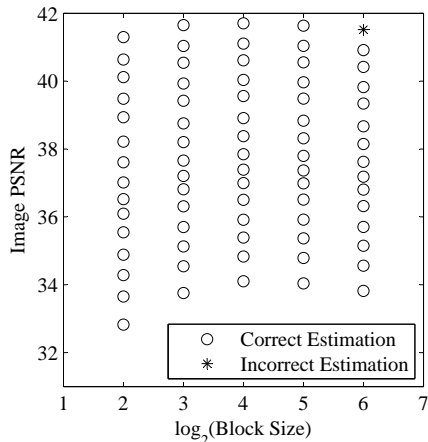
ML Estimation of uncompressed *Lena*:



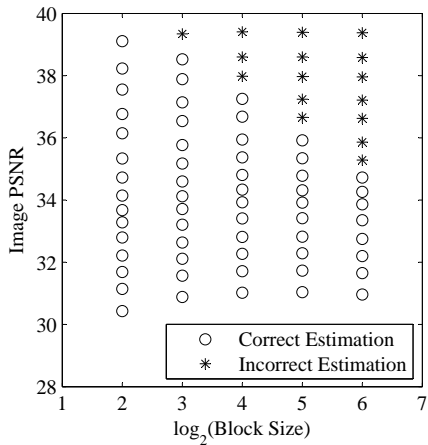
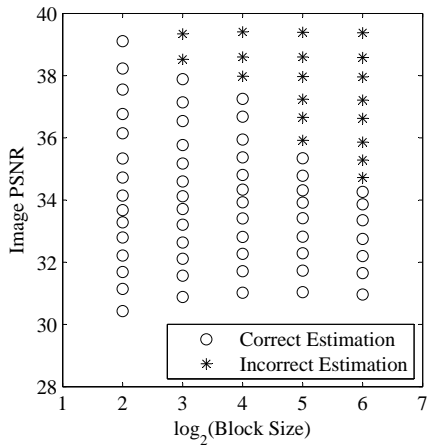
- OSNR is low. Block processing is absent.

Results

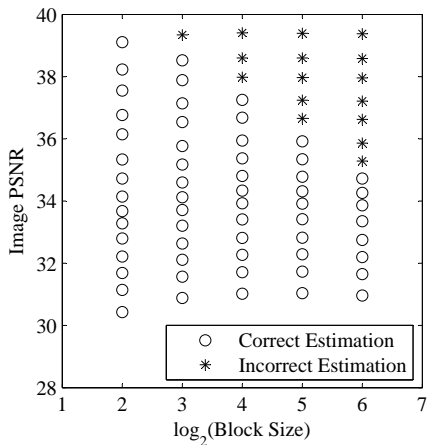
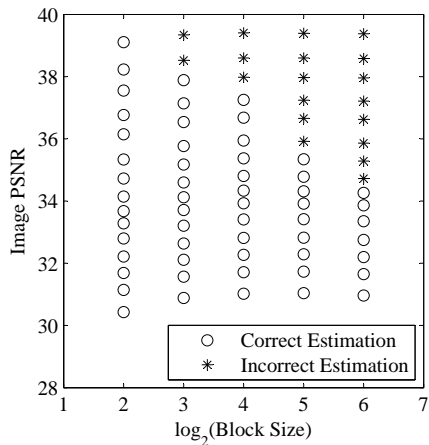
Lena, vertical dimension (left), horizontal dimension (right):



Goldhill, vertical dimension (left), horizontal dimension (right):

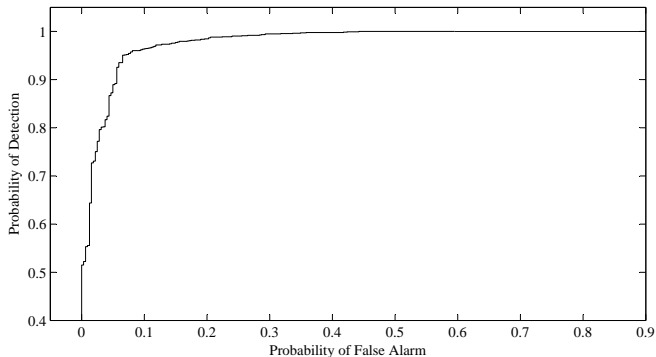


Goldhill, vertical dimension (left), horizontal dimension (right):

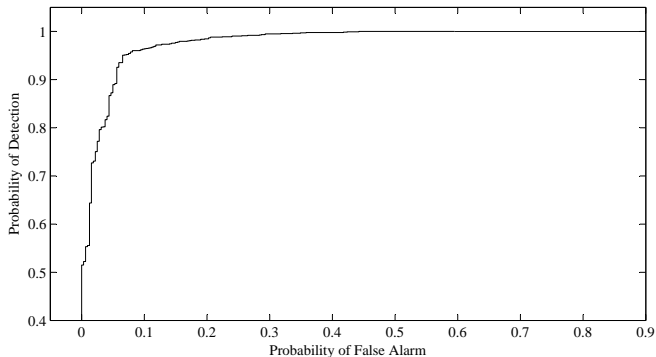


Estimation is successful even when artifacts are imperceptible.

ROC as a function of OSNR, computed from 24 digital photographs across both dimensions, with various block sizes and quality factors:



ROC as a function of OSNR, computed from 24 digital photographs across both dimensions, with various block sizes and quality factors:



Achievable points:

- $P_F = 7.4\%$, $P_D = 95.0\%$
- $P_F = 16.5\%$, $P_D = 98.0\%$

Conclusions

- Proposed a novel non-intrusive forensic method for block size analysis.
- Through median filtering, removal of false peaks, and maximum-likelihood estimation, we can accurately estimate the block size used by an image source coder *without assumptions* upon the nature of the coding procedure.

- Proposed a novel non-intrusive forensic method for block size analysis.
- Through median filtering, removal of false peaks, and maximum-likelihood estimation, we can accurately estimate the block size used by an image source coder *without assumptions* upon the nature of the coding procedure.
- Block sizes can be estimated when artifacts are completely imperceptible.
- Although estimation success may be image dependent, the OSNR remains a reliable measure of detection accuracy.

Future Work: Complete a non-intrusive forensic methodology for digital image source coder identification.

- Which transform is used, if any?
- What specific quantization scheme is used?
- What are the parameters of the coding scheme?
- How do we classify images into families of source coders?

Future Work: Complete a non-intrusive forensic methodology for digital image source coder identification.

- Which transform is used, if any?
- What specific quantization scheme is used?
- What are the parameters of the coding scheme?
- How do we classify images into families of source coders?

Thank you.