Detecting Digital Forgeries Using Bispectral Analysis

Hany Farid

Perceptual Science Group, MIT, Cambridge, MA 02139

With the rapid increase in low-cost and sophisticated digital technology the need for techniques to authenticate digital material will become more urgent. In this paper we address the problem of authenticating digital signals assuming no explicit prior knowledge of the original. The basic approach that we take is to assume that in the frequency domain a "natural" signal has weak higher-order statistical correlations. We then show that "un-natural" correlations are introduced if this signal is passed through a non-linearity (which would almost surely occur in the creation of a forgery). Techniques from polyspectral analysis are then used to detect the presence of these correlations. We review the basics of polyspectral analysis, show how and why these tools can be used in detecting forgeries and show their effectiveness in analyzing human speech.

1 Introduction

The Federal Rules of Evidence outlines the requirements for the introduction of material in a United States court of law [1]. These Rules cover various aspects relating to the introduction of audio recordings and photographs as evidence. The courts have stated that to prove the content of a recording or photograph, the original is required (Article X, Rule 1002). Where Rule 1001 of the same article defines original as:

> An "original" of a writing or recording is the writing or recording itself or any counterpart intended to have the same effect by a person executing or issuing it. An "original" of a photograph includes the negative or any print therefrom. *If data are stored in a computer or similar device, any printout or other output readable by sight, shown to reflect the data accurately, is an "original".*

Given the ease with which digital material can be manipulated it is surprising that the courts have adopted such a broad definition of original. To contend with the technical and legal issues involved in the use of digital material as evidence the courts have turned to the Department of Justice Computer Crime and Intellectual Property Section (CCIPS). Section VIII of the CCIPS Federal Guidelines for Searching and Seizing Computers points out many of the concerns involved in the use of digital media as evidence, but in the end is left to simply conclude:

> For the time being, however, most computer evidence can still be altered electronically - in dramatic ways or in imperceptible detail - without any sign of erasure. *But this does not mean that electronic evidence, having become less distinctive, has become any less admissible. It simply may require us to authenticate it in other ways.*

With the rapid increase in low-cost and sophisticated digital technology the need for techniques to authenticate digital material will become more urgent. In this paper we address the problem of determining whether a digital signal has been tampered with from the time of its recording. As illustrated in Figure 1 our work differs from recent work in digital watermarking where the goal is to embed in a signal an imperceptible and distinctive pattern that is impervious to a broad class of manipulations; the watermark is stored and later used for authentication (see [2] for a review). Such techniques have primarily been developed for protecting copyrights. Our work is also distinct from the vast authentication literature which encodes and encrypts certain aspects of what is assumed to be the authentic signal; this signature is later used for authentication (see [3] for a review). These techniques have been popular in trying to ensure that signals are not tampered with during transmission over insecure lines.

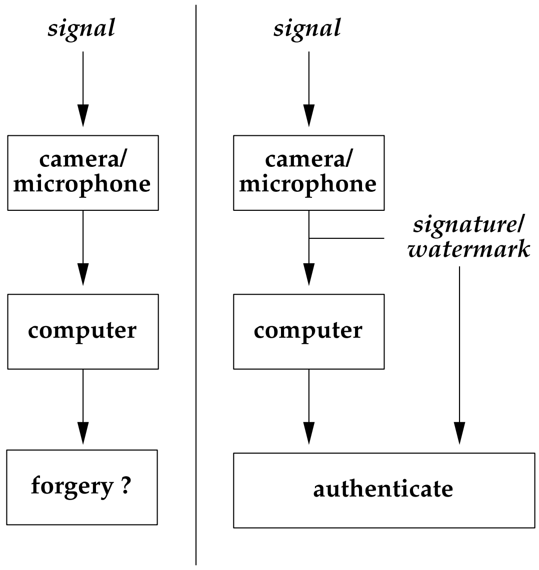

Figure 1: Unlike other authentication techniques (right) we are interested in verifying the authenticity of digital signals (left) with no explicit prior knowledge of the original.

Unlike both of these approaches we assume no explicit prior knowledge of the original signal. The basic approach that we take to detecting digital forgeries is to assume that in the frequency domain a "natural" signal has weak higher-order statistical correlations. We then show that if this signal is passed through a non-linearity (which would almost surely occur in the creation of a forgery), that "un-natural" higher-order correlations are introduced (in magnitude and phase). Techniques from polyspectral analysis [4] are then used to detect the presence of these correlations, and used as an indication of digital tampering. This paper focuses primarily on analyzing human speech: we show that in the frequency domain such signals have weak higher-order correlations, but that when tampered with such correlations are introduced. To begin, we introduce the necessary tools from polyspectral analysis and then show how and why these tools can be useful in detecting forgeries. Several examples showing the effectivness of this approach are given.

signal	frequency	phase
$x(n)$	ω_1, ω_2	ϕ_1, ϕ_2
$x^2(n) + x(n)$	ω_1, ω_2 $2\omega_1, 2\omega_2$ $\omega_1 + \omega_2, \omega_1 - \omega_2$	ϕ_1, ϕ_2 $2\phi_1 - \frac{\pi}{2}, 2\phi_2 - \frac{\pi}{2}$ $\phi_1 + \phi_2, \phi_1 - \phi_2$

Figure 2: A simple non-linearity introduces new harmonics with higher-order correlations.

2 Bispectral Analysis

For purposes of illustration consider a signal composed of a sum of two sinusoids:

$$x(n) \;=\; \sin(\omega_1 n + \phi_1) + \sin(\omega_2 n + \phi_2). \qquad (1)$$

Consider now passing this signal through a simple non-linearity:

$$y(n) \;=\; x^2(n) + x(n). \qquad (2)$$

Expanding the polynomial and rewriting in terms of the harmonics using basic trigonometric identities gives:

$$\begin{aligned} y(n) \;=\; & \tfrac{1}{2}(1 + \sin(2\omega_1 n + 2\phi_1 - \tfrac{\pi}{2})) \\ & + \tfrac{1}{2}(1 + \sin(2\omega_2 n + 2\phi_2 - \tfrac{\pi}{2})) \\ & + 2\sin((\omega_1 + \omega_2) + (\phi_1 + \phi_2)) \\ & + 2\sin((\omega_1 - \omega_2) + (\phi_1 - \phi_2)) \\ & + \sin(\omega_1 n + \phi_1) + \sin(\omega_2 n + \phi_2). \qquad (3) \end{aligned}$$

Notice that this simple non-linearity introduces new harmonics with frequencies and phases that are correlated (Figure 2). [1] In addition, these correlations simply could not have been introduced through a linear transform. Our goal now is to try to detect these "un-natural" higher-order correlations as a means of detecting the presence of a non-linearity.

The signal is first decomposed according to the Fourier transform:

$$Y(\omega) \;=\; \sum_{k=-\infty}^{\infty} y(k) e^{-ik\omega}, \qquad (4)$$

with $\omega \in [-\pi, \pi]$. It is common practice to use the power spectrum to detect the presence of second-order correlations:

$$P(\omega) \;=\; Y(\omega) Y^*(\omega), \qquad (5)$$

where $Y^*(\omega)$ is the complex conjugate. However the power spectrum is blind to higher-order correlations of the sort of interest to us. However these correlations can be detected by turning to higher-order spectral analysis. For example the bispectrum is used to detect the presence of third-order correlations:

$$B(\omega_1, \omega_2) \;=\; Y(\omega_1) Y(\omega_2) Y^*(\omega_1 + \omega_2). \qquad (6)$$

Comparing the bispectrum with Equation (3) and Figure 2 we can see intuitively that the bispectral response reveals the sorts of "un-natural" higher-order correlations introduced by a non-linearity. That is, correlations between the triples $[\omega_1, \omega_1, \omega_1 + \omega_1]$, $[\omega_2, \omega_2, \omega_2 + \omega_2]$, $[\omega_1, \omega_2, \omega_1 + \omega_2]$, and $[\omega_1, -\omega_2, \omega_1 - \omega_2]$.

From an interpretive stance it will be convenient to express the complex bispectrum with respect to its magnitude:

$$|B(\omega_1, \omega_2)| \;=\; |Y(\omega_1)| \cdot |Y(\omega_2)| \cdot |Y(\omega_1 + \omega_2)|, \qquad (7)$$

and phase:

$$\angle B(\omega_1, \omega_2) \;=\; \angle Y(\omega_1) + \angle Y(\omega_2) - \angle Y(\omega_1 + \omega_2). \qquad (8)$$

Also from an interpretive stance it is helpful to work with the normalized bispectrum, the bicoherence:

$$B_c(\omega_1, \omega_2) \;=\; \frac{Y(\omega_1) Y(\omega_2) Y^*(\omega_1 + \omega_2)}{\sqrt{|Y(\omega_1) Y(\omega_2)|^2 |Y(\omega_1 + \omega_2)|^2}}. \qquad (9)$$

There is much debate in the literature on the "correct" normalization factor; this particular form [6] is adopted because it guarantees that the bicoherence values are in the range $[0, 1]$. In the absence of noise the bicoherence can be estimated from a single realization as in Equation (9). However in the presence of noise some form of averaging is required to ensure stable estimates. A common and convenient form of averaging is to divide the signal into multiple segments. For example the signal $y(n)$ with $n \in [1, N]$ can be divided into K segments of length $M = N/K$, or K overlapping segments with $M > N/K$. The bicoherence is then estimated from the average of each segment's bicoherence spectrum:

$$\hat{B}_c(\omega_1, \omega_2) = \frac{\frac{1}{K}\sum_k Y_k(\omega_1) Y_k(\omega_2) Y_k^*(\omega_1 + \omega_2)}{\sqrt{\frac{1}{K}\sum_k |Y_k(\omega_1) Y_k(\omega_2)|^2 \frac{1}{K}\sum_k |Y_k(\omega_1 + \omega_2)|^2}}. \qquad (10)$$

Correlations beyond second- and third-order can also be detected. For example, the trispectrum is sensitive to fourth-order correlations:

$$T(\omega_1, \omega_2, \omega_3) = Y(\omega_1) Y(\omega_2) Y(\omega_3) Y^*(\omega_1 + \omega_2 + \omega_3). \qquad (11)$$

[1] An observation that Feynman mentioned would "have many practical applications". [5]

And in general the N^{th}-order polyspectrum is sensitive to $N+1^{st}$-order correlations. These polyspectra can also be computed for higher-dimensional signals. For example, a 2-D signal with Fourier transform $Y(\omega_x, \omega_y)$ has a 4-D bispectrum:

$$\begin{aligned} B(\omega_{x1}, \omega_{y1}, \omega_{x2}, \omega_{y2}) \quad &= \\ &Y(\omega_{x1}, \omega_{y1}) Y(\omega_{x2}, \omega_{y2}) Y^*(\omega_{x1}+\omega_{x2}, \omega_{y1}+\omega_{y2}). \end{aligned} \tag{12}$$

For now we will concentrate on the use of the bispectrum/bicoherence for detecting the presence of "un-natural" correlations brought about through some form of non-linear tampering.

3 Detecting Forgeries

In the previous section we saw that a simple squaring non-linearity applied to the sum of two sinusoids introduces new harmonics with distinct higher-order correlations between the individual harmonics, Equation (3). We also saw how the bispectrum/bicoherence can be used to detect these correlations, Equation (6). If such higher-order correlations are weak in "natural" signals then their presence can be used as an indication of tampering thus casting the authenticity of the signal into a suspicious light. In the next section we provide some empirical evidence that for human speech, these higher-order correlations are in fact relatively weak.

But first we need to show why we might expect increased activity in the bicoherence when an arbitrary non-linearity is applied to an arbitrary signal. To this end consider first an arbitrary function $f(\cdot)$:

$$y(n) \;=\; f(x(n)). \tag{13}$$

For notational convenience this system is expressed in an equivalent vector form as $\vec{y} = f(\vec{x})$. The vector valued function $f(\cdot)$ can be expressed in terms of scalar valued functions:

$$f(\vec{x}) \;=\; \left(f_1(\vec{x}) \quad f_2(\vec{x}) \quad \ldots \quad f_N(\vec{x}) \right). \tag{14}$$

Each of these scalar valued functions $f_i(\cdot)$ can subsequently be expressed in terms of its Taylor series expansion about a point $\vec{p}$:

$$f_i(\vec{x}) = f_i(\vec{p}) + \sum_j \frac{\partial f_i}{\partial x_j} x_j + \frac{1}{2} \sum_{j,k} \frac{\partial^2 f_i}{\partial x_j \partial x_k} x_j x_k + \ldots \tag{15}$$

For simplicity this expansion is truncated after the second-order term and rewritten in vector/matrix form as:

$$f_i(\vec{x}) \;\approx\; a + \vec{b}\vec{x} + \frac{1}{2}\vec{x} C \vec{x}, \tag{16}$$

with the scalar $a = f_i(\vec{p})$, the vector $\vec{b}_j = \frac{\partial f_i}{\partial x_j} \mid_{\vec{p}}$, and the matrix $[C]_{jk} = \frac{\partial^2 f_i}{\partial x_j \partial x_k} \mid_{\vec{p}}$, with $i, j, k \in [1, N]$. The point of all of this is simply that the application of an arbitrary non-linearity $f(\vec{x})$ results in an output that contains a sum of a linear term $\vec{b}\vec{x}$ and a quadratic term $\vec{x}C\vec{x}$. Note the similarity with the original example in the previous section, Equation (2).

Consider now the input signal $\vec{x}$ expressed in terms of its Fourier series:

$$x(n) \;=\; \sum_{k=0}^{\infty} a_k \sin(kn + \phi_k). \tag{17}$$

The situation is now very similar to that of the previous section, Equations (1) and (2). That is, the application of an arbitrary non-linearity $f(\cdot)$ to an arbitrary signal $x(n)$ will result in higher-order correlations between various harmonics, and hence an increased bicoherence activity.

In fact, even more specific predictions can be made by looking closely at how a non-linearity affects the bicoherence magnitude and phase. Recall that a non-linearity introduces new harmonics that are correlated with the original harmonics, Figure 2. For example, given a pair of harmonics ω_1 and ω_2, a non-linearity produces a new harmonic $\omega_1 + \omega_2$ whose amplitude is correlated to ω_1 and ω_2. To the extent that such correlations do not occur naturally, these correlations will result in a larger response in the bicoherence magnitude, Equation(7). In addition, if the initial harmonics have phases ϕ_1 and ϕ_2, then the phase of the newly introduced harmonic is $\phi_1 + \phi_2$. It is then easy to see that the bicoherence phase for this pair of harmonics will be 0, Equation (8). Again, to the extent that these higher-order correlations do not occur naturally, the correlations will result in a bias of the bicoherence phase towards 0. A phase bias towards $\pi/2$ may also occur due to correlations of any harmonic ω_1 with itself. That is, the bicoherence phase $\angle B(\omega_1, \omega_1) = \phi_1 + \phi_1 - (2\phi_1 - \pi/2) = \pi/2$.

We predict that a non-linearity will manifest itself in the bicoherence in two ways: for certain harmonics an increase in the magnitude, and for those harmonics a phase bias towards 0 and/or $\pi/2$. In the following section we will show how these principles can be used in the detection of digital forgeries.

4 Results

We first show how the bicoherence can detect the presence of global and local non-linearities of the form described in Section 2. We then show how these ideas generalize to a broader class of non-linearities that might be involved in the creation of digital forgeries.

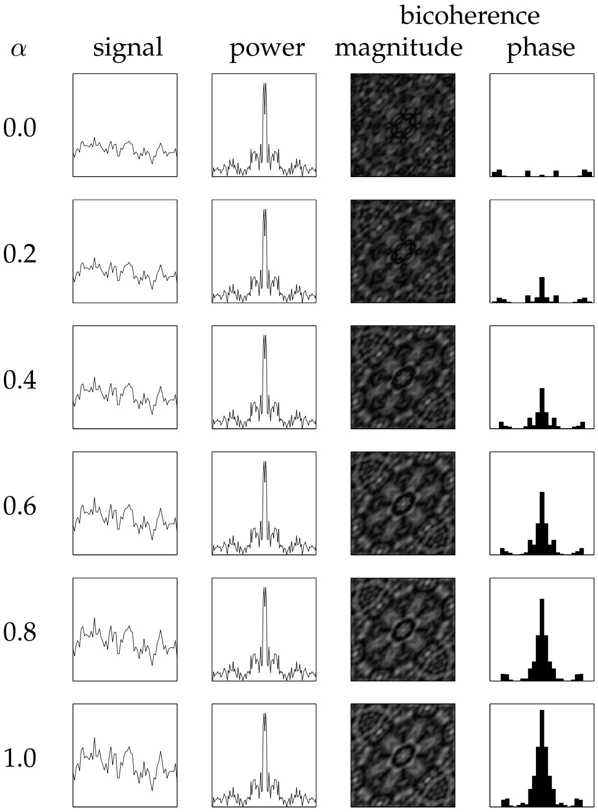

Figure 3: Detecting non-linearities. Shown from left to right is a portion of the input signal $y(n) = \alpha x^2(n) + x(n)$, its normalized power spectrum $P(\omega)$ with $\omega \in [-\pi, \pi]$, its bicoherence $\hat{B}_c(\omega_1, \omega_2)$ with $\omega_1, \omega_2 \in [-\pi, \pi]$, and the bicoherence phase histogram plotted from $[-\pi, \pi]$. As the non-linearity increases there is no change in the power spectrum, but a significant increase in the overall magnitude and phase bias (see also Figure 4).

4.1 Global Non-Linearities

Throughout these examples the input is a 1-D fractal signal $x(n)$ with 2048 samples and a power spectral density of the form $|X(\omega)| = 1/\omega$. In this first set of experiments a signal is subjected to a global non-linearity of the form:

$$y(n) = \alpha x^2(n) + x(n), \tag{18}$$

with $\alpha \in [0, 1]$. The bicoherence is computed by dividing the signal into 127 overlapping segments of length 32 with an overlap of 16. A 128-point windowed DFT, $Y_k(\omega)$, is estimated for each segment from which the bicoherence is estimated, $\hat{B}_c(\omega_1, \omega_2)$ (Equation (10)). [2]

[2]The signal is windowed with a symmetric Hanning window prior to estimating the DFT.

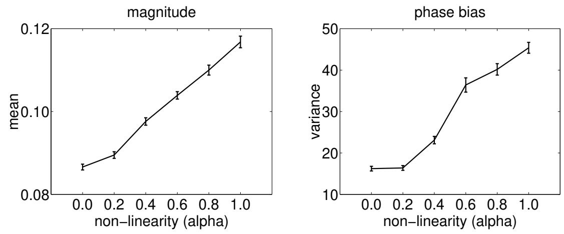

Figure 4: Bicoherence averaged over one-hundred independent signals of the form $\alpha x^2(n) + x(n)$ with $\alpha \in [0.0, 1.0]$. The error bars correspond to the standard error. As the degree of non-linearity increases both the magnitude and phase bias increase (see also Figure 3).

Shown in the first column of Figure 3 are several signals with varying amounts of non-linearity, i.e., increasing values of α in Equation (18). The signals are all plotted on the same scale of $[0, 1]$. Shown within each row is a small portion of the input signal $y(n)$, the normalized power spectrum, the bicoherence magnitude $|\hat{B}_c(\omega_1, \omega_2)|$, and a histogram of the bicoherence phase $\angle \hat{B}_c(\omega_1, \omega_2)$ for those frequencies with $|\hat{B}_c(\omega_1, \omega_2)| > 0.2$. With increasing amounts of non-linearity there is an increase in the magnitude and phase bias towards zero. At the same time the normalized power spectrum remains virtually unchanged - it is blind to the higher-order correlations introduced by the non-linearity.

The increase in bicoherence activity can be quantified by measuring the mean magnitude and the phase bias. To measure the phase bias away from uniform we compute the variance across the bin counts from a discrete histogram of the phases (i.e., deviations from a uniform distribution will have a large variance). To show the general robustness of this measure the bicoherence is measured for one hundred independent fractal signals with varying levels of non-linearity. Shown in Figure 4 are the mean magnitude and phase bias plotted as a function of increasing amounts of non-linearity. Each data point corresponds to the average response over the one hundred trials, the error bars correspond to the standard error. With increasing amounts of non-linearity the mean magnitude and phase bias increase by 35% and 165%, respectively.

4.2 Local Non-Linearities

The bicoherence is able to reliably detect the presence of a simple global non-linearity (Figures 3 and 4). For the purposes of forgery detection it would also be desirable to be able to detect the presence of local non-linearities. To test whether this is possible a small portion of a sig-

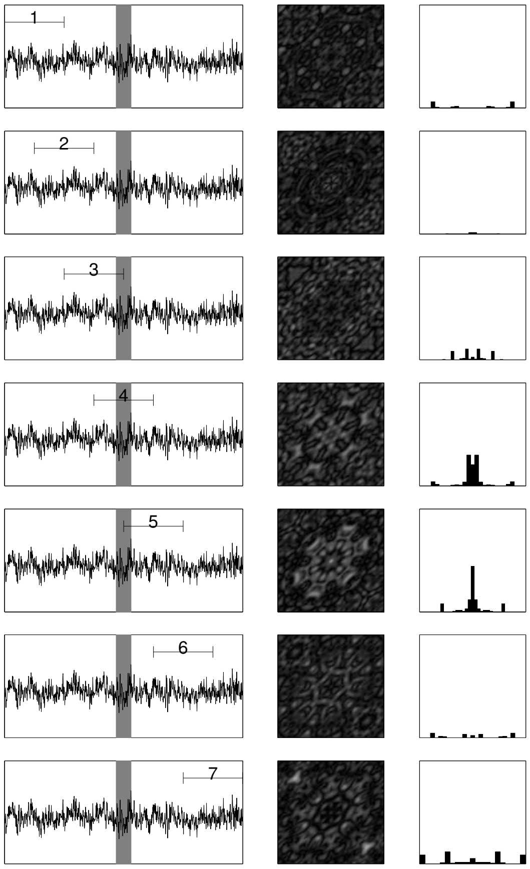

Figure 5: Detecting non-linearities. Shown in the first column is the same input signal, where the middle gray region was subjected to a squaring non-linearity, Equation (2). Shown to the right of each signal is the bicoherence magnitude and phase histogram for seven overlapping windows. Notice the increased activity in the central segment that includes the non-linearity.

nal was exposed to the same squaring non-linearity as above ($\alpha = 1$ in Equation (18)). The full signal contains 8192 samples, with the central 512 segments subjected to the non-linearity. Using the same methodology as above the bicoherence is computed at seven different windows along the signal, where each window contains 2048 samples with an overlap of 1024. Shown in Figure 5 is the bicoherence magnitude and phase. For the central (and one neighboring) window containing the non-linear segment there is an increased activity in both the magnitude and phase bias; for window four the mean magnitude is 0.10 and the phase bias is 42.7. Relative to the global squaring in Figure 4 this amounts to a non-linearity on the order of $\alpha = 0.6$.

4.3 Synthetic Forgeries

We have seen how the bicoherence is able to detect "unnatural" higher-order correlations caused by a simple squaring non-linearity (Figures 3 and 5). We would now like to see how well this generalizes to arbitrary non-linearities. A common technique in digital forging is to splice together signals in such a way that the seam is not perceptually salient. A pair of fractal signals were seamlessly spliced together using a Laplacian pyramid [7]. Each signal, $x_1(n)$ and $x_2(n)$, containing 8192 samples is first decomposed into a seven level Laplacian pyramid. A new pyramid is constructed by combining, at each pyramid level, the left half of $x_1(n)$ with the right half of $x_2(n)$. This new pyramid is then collapsed yielding the "forgery". The rationale for this scheme is that the splicing looks seamless because the low frequencies are blended over a large spatial extent, while the details (high frequencies) are preserved by blending over a smaller extent. Using the same methodology as described above the bicoherence is computed at several windows along the forged signal. Shown in Figure 6 is the bicoherence magnitude and phase for each window. Notice the increased activity in the central region where the splice occurred; the mean magnitude is 0.11 and the phase bias is 54.7. Relative to the global squaring non-linearity this amounts to a non-linearity on the order of $\alpha = 0.8$ (Figure 4).

The robustness of this detection was tested over one-hundred independent signals spliced together in the manner described above. For each signal the bicoherence is measured across the splice point (window four in Figure 6). Shown in Figure 7 is the mean magnitude and phase bias averaged over all the trials. Also shown are the average values for signals that were not subjected to any tampering. On average there is a 26% increase in the mean magnitude and 72% increase in phase bias.

4.4 Audio Forgeries

We analyzed audio recordings from the four volume collection of Great Speeches of the 20^{th} Century (Rhino Records Inc, 1991). The first ten seconds of twenty speeches were digitized. As described in previous sections the bicoherence is estimated for each signal at approximately 200 non-overlapping windows of length 2048. Shown in Figure 8 are the list of speakers and the average magnitude and phase bias. With the exception of one speaker (J. Jackson), the average magnitude and phase bias are 0.109 and 38.2 with a standard error of 0.002 and 3.1, respectively. Again with the exception of one speaker, the average number of windows that contained a magnitude and phase bias above "normal" is just under 6%. Where

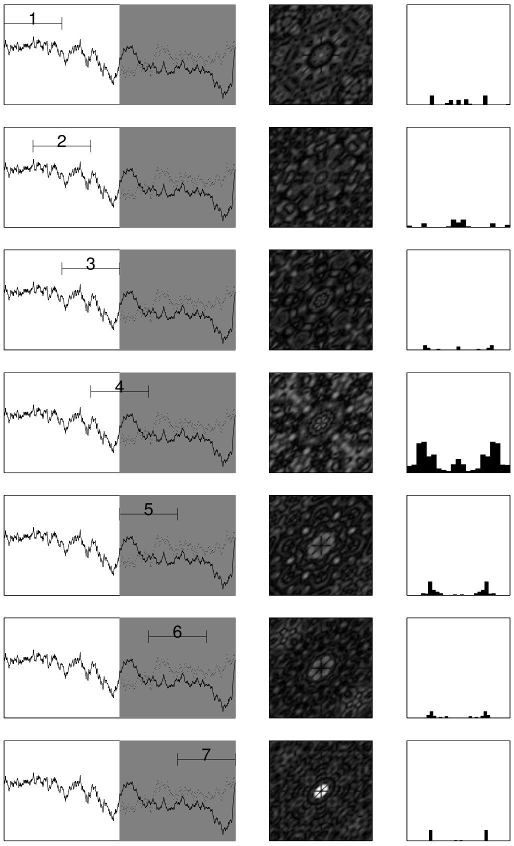

Figure 6: Shown in the first column is the same input signal consisting of a pair of signals spliced together. The dotted line represents the true continuation of the left-half of the signal. Shown to the right of each signal is the bicoherence magnitude and phase histogram for seven overlapping windows. Notice the increased activity in the central segment where the splicing occurred (see also Figure 7).

a magnitude above 0.15 and phase bias above 50 is considered to be outside of the normal range, Figure 9. With respect to the simulations in the previous sections (Figure 4) these natural signals register as containing mid-range non-linearities. This is not surprising given the inherent non-linearities in the recording process and the non-stationarity of the signals. What is encouraging is that even over such a broad range of signals the responses are quite similar thus providing a baseline for future comparisons.

Given the numerous ways in which a forgery can be made there is of course no way of systematically testing the overall efficacy of the bicoherence. Nevertheless, we provide one example starting with the following portion of a rambling speech by Dan Quayle:

signal	magnitude		phase bias	
	mean	std. err.	mean	std. err.
natural	0.087	0.0001	16.1	1.0
forgery	0.11	0.001	27.8	1.5

Figure 7: Bicoherence averaged over one-hundred "natural" and forged signals. Notice that the forged signals have an increased magnitude and phase bias (see also Figure 6).

> Our government unlike many governments and particularly the governments of where the people that founded this country came from is a government that is derived from the people that consent to govern the freedom that is based in the people that then elect their representatives to represent them in a free representative democracy that we have today.
>
> *Dan Quayle*

Several splices were made to this speech and rejoined with local smoothing to yield the more coherent message:

> Our government is a government derived from the people that then elect their representatives in a democracy that we have today.
>
> *Forgery*

The average bicoherence magnitude and phase bias jumped from within normal range, 0.169/41.0, to well outside of the normal range, 0.234/69.2, Figure 9. Even though several people listening to both audio clips were unable to identify the forgery there is a significant increase in the bicoherence activity that would have tagged this recording as suspicious.

In addition to analyzing forgeries made by tampering with natural speech we were curious if the very unnatural sounding computer generated speech would also be tagged as such. We analyzed the short answering machine message "Hello, you have no messages" spoken in a computer generated voice. The bicoherence magnitude and phase bias is 0.225/76.0 and fell well outside of the normal range, 'x' in Figure 9. Although not a forgery in the traditional sense, the computer generated voice clearly does not sound like natural human speech.

Finally, the speech of the Reverend Jesse Jackson is incorrectly tagged as suspicious (the filled circle in top right quadrant of Figure 9). We hypothesize that this is

speaker	average mag	average phase	%
Hank Aaron	0.121	41.5	9.8
Spiro Agnew	0.112	34.9	6.5
Winston Churchill	0.104	31.1	4.1
Mario Cuomo	0.114	39.8	10.9
Mayor Richard Daley	0.118	41.6	9.5
Ameila Earhart	0.101	22.9	0.5
Dwight D. Eisenhower	0.109	68.0	8.5
Barry Goldwater	0.113	51.8	9.5
Jesse Jackson	0.182	88.0	38.5
Lyndon B. Johnson	0.104	25.2	2.0
Edward Kennedy	0.111	41.2	6.1
John F. Kennedy	0.105	32.3	4.1
Martin Luther King	0.097	23.0	0.5
Charles Lindbergh	0.118	46.8	9.0
Douglas MacArthur	0.118	46.5	6.6
Joseph McCarthy	0.105	53.1	8.9
Richard M. Nixon	0.096	26.3	0.5
Gloria Steinem	0.118	48.7	9.2
Harry S. Truman	0.096	12.8	1.0
Malcolm X	0.107	38.5	6.2

Figure 8: Speaker data set. Shown in the second and third columns are the average bicoherence magnitude and phase bias. Shown in the last column is the percentage of windows where both the magnitude and phase bias were above "normal" (see also Figure 9).

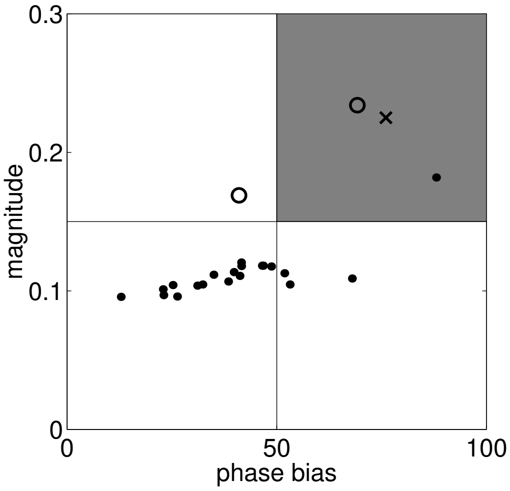

Figure 9: Shown is the average bicoherence magnitude and phase bias for the 20 speakers of Figure 8 (filled circles). The open circles correspond to a speech by Dan Quayle before (top left quadrant) and after tampering (top right quadrant). The 'X' corresponds to a computer generated voice. Points in the top right quadrant are tagged as suspicious.

due to significant amounts of feedback in the recording that was not present in any of the other recordings. This false positive illustrates an important limitation of looking to the bicoherence as a means of identifying digital forgeries. Namely, that we are unable to distinguish between innocuous non-linearities and ones that are meant to deceive the listener. In addition, a sufficiently sophisticated forger could mask any malicious tampering by simply adding a harmless global non-linearity to the signal. While it may not yet be possible to differentiate between various types of non-linearities, their presence should at a minimum cast the authenticity of the signal into a suspicious light.

5 Discussion

This paper addresses the problem of determining whether a digital signal has been tampered with from the time of its recording. Unlike techniques in digital watermarking and authentication, explicit knowledge of the original signal is *not* assumed. It is however assumed that in the frequency domain "natural" signals have weak higher-order correlations. To the extent that this assumption holds, we showed that "un-natural" higher-order correlations are introduced when a signal is passed through a non-linearity (which would almost surely occur in the creation of a forgery). Tools from polyspectral analysis (bispectrum/bicoherence) are employed to detect these correlations and used as an indicator of digital tampering. More specifically we showed that a non-linearity manifests itself with an increase in the bicoherence magnitude and a bias in the bicoherence phase towards 0 or $\pi/2$. We demonstrated the applicability of this technique by first showing that for human speech signals higher-order correlations are in fact weak, and that tampering with such signals is revealed in the bicoherence.

Undoubtedly there has been significant efforts by the law enforcement community in developing techniques for digital forgery detection. However to maximize their effectiveness much of this work remains unpublished. Although there are disadvantages to revealing this new technique, an advantage of analyzing higher-order statistical correlations is that their presence is most likely imperceptible to the visual system, and it is not immediately obvious how they can be removed by manipulating the signal in the spatial domain. So while a forger may be aware of this detection scheme, they may not be able to remove the correlations introduced by their tampering.

Another advantage of the bicoherence is that it is not effected by additive white noise that is often added to an image to disguise evidence of tampering.

There are of course limitations to what the bicoherence is capable of detecting. For example, highly localized and minor tampering will most likely go unnoticed, then again, such manipulations are unlikely to dramatically alter the meaning of a signal. Additionally, rather innocuous factors can lead to an increased activity in the bicoherence, for example, significant non-linearities in the recording process. As a result, the detection of a non-linearity does not immediately imply the discovery of a forgery. However it should give reason to question the authenticity of the signal, suggesting that further inspection is warranted.

Finally, polyspectral analysis is not limited to only third-order statistics of one-dimensional signals. We are extending our analysis to higher-order statistics and applying these techniques to the analysis of digital images and video sequences.

Acknowledgments

The author thanks Ted Adelson for his generosity and support, Mary Bravo for her insightful comments, and the generous support of NIH Grant EY11005-04 and MURI Grant N00014-95-1-0699.

References

[1] United States. *Federal rules of evidence*. West Publishing Company, 1992.

[2] J. Ford, F. Makedon, and C.B. Owen. *Handbook of Multimedia Computing*, chapter Classification and characterization of digital watermarks for multimedia data, pages 299–316. CRC Press, 1999.

[3] D.R. Stinson. *Cryptography : theory and practice*. CRC Press, 1995.

[4] J.M. Mendel. Tutorial on higher order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proceedings of the IEEE*, 79:278–305, 1996.

[5] R. P. Feynman, R. B. Leighton, and M. Sands. *The Feynman Lectures on Physics*. Addison-Wesley Publishing Company, 1977.

[6] J.W.A. Fackrell and S. McLaughlin. Detecting nonlinearities in speech sounds using the bicoherence. *Proceedings of the Institute of Acoustics*, 18(9):123–130, 1996.

[7] E.H. Adelson and P.J. Burt. Image data compression with the Laplacian pyramid. In *Proceedings of the conference on pattern recognition and image processing*, pages 218–223, Dallas, TX, 1981.