High-Profile Forensic Analysis of Images

Siome Goldenstein, Anderson Rocha

Institute of Computing, University of Campinas (UNICAMP)
Campinas, SP — Brazil
{siome, anderson.rocha}@ic.unicamp.br

Keywords: Digital Image Forensics, Image Forgery Detection, Image Tampering, Image Manipulation.

Abstract

Amidst many different forms of image manipulation, how to convince a jury of a tampering? Is traditional expert opinions enough? In this paper, using a high-profile Brazilian case as a guideline, we explain how we can take advantage of important statistical methodologies and state-of-the-art techniques to verify evidences of digital tampering, beyond reasonable doubt.

1 Introduction

In our digital age, images and videos reach us at unprecedented frequency and, unfortunately, there are currently no established methodologies to verify their authenticity and integrity in an automatic manner [1]. As a consequence, on a daily basis we are faced with numerous images and videos — and it is likely that at least a few have undergone some level of manipulation.

At the same time our understanding of the technological, ethical, and legal implications associated with image editing falls far behind. When such modifications are no longer innocent image adjustments and start implying legal threats to a society, it becomes paramount to devise and deploy efficient and effective approaches to detect such activities [2, 3].

In this context, on Sunday April 5^{th}, 2009, the brazilian newspaper *Folha de São Paulo*, one amongst the top largest printed circulations with approximately 360,000 copies on Sundays[1,2], published on pages A8-A10 an article on how the current Brazilian Secretary of State Dilma Rousseff (the Brazilian Chief of Staff and a possible runner for the presidential office on the coming 2010 election) has actively participated in the resistance against the military regime, such as the planning and preparations of robberies and kidnappings. As part of the article, the newspaper printed an alleged image of the Repression Police Internal files of Secretary Rousseff, in Figure 1, and stated that it came from the *Public Archive of São Paulo*[3], that houses the collection of documents from that period of time.

Secretary Rousseff contacted the newspaper's Ombudsperson arguing against the veracity of the evidence published, claiming that all her activities during the Brazilian military regime were non-violent, and that the "police records" could not have come from the *Public Archive of São Paulo* Collection.

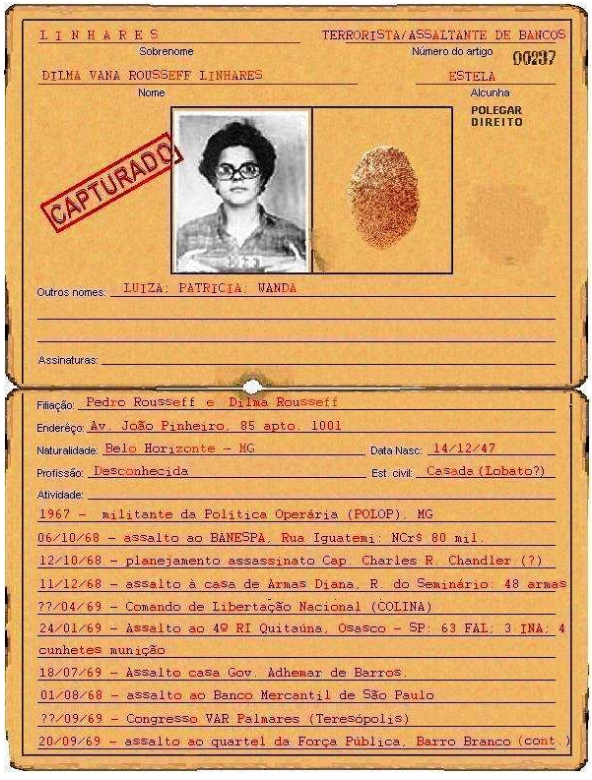

Figure 1. Object in question, the alleged Image of Repression's Secret Police files on the Secretary of State.

On the edition of April 25^{th}, 2009, the newspaper acknowledged two errors in their former article: that the image was received by e-mail from an anonymous source, and that the image was treated as authentic. They also stated that the authenticity

[1] http://www1.folha.uol.com.br/folha/conheca/
[2] http://en.wikipedia.org/wiki/Folha_de_S._Paulo
[3] http://www.arquivoestado.sp.gov.br/proin.php

of the document could not be proved or disproved at that point.

The Secretary contacted our University, which has a standing tradition for Forensic Analysis in Brazil, and requested a forensic authenticity analysis of the digital image that could stand in court, if necessary.

The case of Secretary Rousseff we report here is not alone. Photo and video retouching and manipulation are being more and more used for political purposes. In the 2004 US presidential campaign, a photomontage appeared in several newspapers surprising John Kerry's allies. It showed Kerry and Jane Fonda standing together at a podium in a 1970s anti-war rally. Further investigations showed that Kerry's picture was taken at an anti-war rally in Mineola, NY., on June 13^{th}, 1971 by Ken Light while Fonda's picture was taken during a speech at Miami Beach, FL. in August, 1972 by Owen Franken [1].

In a similar case, shortly after Sarah Palin was announced as the vice presidential nominee for the Republican ticket in 2008, an image was widely distributed across the Internet sporting her wearing an American flag bikini and a holding a rifle. Later on, the photo was revealed to be a composite of Palin's head, and somebody else's body.

Different versions of the image in Figure 1, the object in question that was used by the newspaper in their article, has been circulating on different internet venues since mid 2008. The image has several unusual and suspicious artifacts that indeed hint a forgery. The goal of this paper is to describe the methodology we have used in our forensic analysis to prove the doctoring beyond reasonable doubt.

2 Methodology

People copy and republish images on the internet, resampling, cropping, changing gamma settings and re-compressing them. Each one of these operations changes the properties of the final image, adds new artifacts, and can make forgery analysis harder. This was certainly the case here, where the object in question has been widely redistributed for several months prior to its appearance on the newspaper.

The newspaper's online version of the article, available only for subscribed users, had a very low resolution version of the image (the text was not even readable). The Secretary's staff provided us with an additional four versions of the image, obtained from different web-sites and blogs, for further analysis.

2.1 Choosing the Right Image for Analysis

For digital image forensics techniques, the least post-processing information performed over the analyzed image the better. However, for a "live" image in the web, to point out its first tampered version or the *patient zero* is not an easy task. For our analysis, we needed to identify which image contained the least amount of modification not tagged as tampering (e.g., resizing, resampling, and re-compressing) before any further analysis.

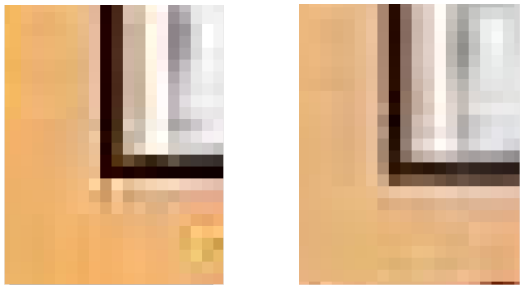

Figure 2. Zoom of image 1 (left) and image 4 (right).

There are many versions of analyzed image over the internet, which slightly differs from one another. Although these differences do not alter the gross appearance (and thus do not change the meaning) of the image, such representational differences can fool even sophisticated detection schemes. We had to decide which of the five provided images (Table 1) was the best one for analysis, and whether it was close enough to the first version of the image.

The *Public Archive of São Paulo* organizes its 1964-1985 political dossiers in 6"$\times$ 4" paper files. One paper file captured at 96 DPIs in any scanner has about 576 pixels wide. From this information, we can eliminate image 2 since its geometry is incompatible with the DPI information. This image was probably re-compressed.

We also can discard images 3 and 5 as a resampling of the possible original source as they have geometry and DPI information incompatible with the expected.

Performing a visual inspection of images 1 and 4, we observe that vertical and horizontal straight lines in image 1 are more precise and well-defined than in image 4. Furthermore, lines in image 4 contain blurred artifacts pointing out possible resampling telltales. Figure 2 illustrates this behavior with a zoom of both images on the same region. This initial procedure has lead us to focus the forensic analysis on image 1.

2.2 Data set Reference Construction

As stated earlier, the *Public Archive of São Paulo* houses, since 1991, a collection with the documents of the *Departamento de Ordem Política e Social* (DEOPS). As part of the collection, there are the internal dossiers of left-wing politic activists during the military repressive regime, organized in 6"$\times$ 4" cartoon files. The Secretary does indeed have a file there, but it does not look like the published image in any way.

Items from the Library's special collections are not allowed to leave the premises. We spent a day collecting data on-site using their three available pieces of scanning equipment, namely: HP 5470 (Scanner 1), HP G4050 (Scanner 2), and WideTek 25 (Scanner 3). Forensic techniques need to compare the test object to some reference pattern with similar properties.

1	http://www.viomundo.com.br/img/Curriculum_Vitae___Dilma_Roussef.JPG
2	http://www.economiaepolitica.com.br/images/presidente.jpg
3	http://www.misturinha.com/2/wp-content/uploads/2009/04/dilma.jpg
4	INFORME RECIFE, Boletim 27-05
5	http://www1.folha.uol.com.br/fsp/brasil/fc0504200908.htm

	md5sum	Size	DPI	Compression
1	afe8ebf8f99b63dd4cbf3133bad597b8	596 × 784	96 × 96	JPEG 75%
2	6020a7d2296f94d3505907fb8d16afe7	596 × 784	72 × 72	JPEG 85%
3	941e3ea5a7579d0e5ead53d70d16da3b	510× 671	72 × 72	JPEG 80%
4	5dce7e352ab41166fdaa41c378fb5acc	567× 746	96 × 96	JPEG 70%
5	07f2b9618002c7025aa8c8e143a42c11	266× 350	100× 100	JPEG 72%

Table 1. The five images in the analysis, their original location, and intrinsic properties.

We selected the files from 30 random individuals that contained pictures, typewritten text, and stamped information. We scanned them independently in each of the three scanners at resolutions of 96, 300, and 600 DPIs, yielding a total of 90-115 images per scanner (some files had information on both sides). At this stage the images were not compressed.

This careful process of data collecting ensures that the analyzed reference data represents the typical paper files of the stored political archives in the period of time of the alleged image of Secretary Rousseff. The collected data contains paper aging and typographic artifacts, photographic process, foldings, and imperfections due to mishandling over time.

2.3 Photography Splicing

One giveaway sign of manipulation is the absence of color information in the region of the photograph in the object in question (Figure 1). This strongly suggested it was the result of a splicing process from a grayscale photography into the object.

In this section, we calculate the probability of this occurrence in a non-manipulated image, and show beyond reasonable doubt that this part is the result of a forgery.

We used the data set from Section 2.2, collected for this analysis, as reference. We resampled all images to 96×96 DPIs, and stored them in the JPEG format with compression quality of 75%. This procedure ensures a comparison data set with similar properties in both resolution and compression artifacts to the object under analysis. There is no need to test against multiple compression factors, since in the JPEG files luminance and chrominance are treated independently [4, 5].

The analysis was focused on the variability of color information, using the saturation channel (S) of the HSV color model. For every image on the data set, we manually identified the photograph regions, and used all these pixels to examine the variability of the S channel. Similarly, in the object of analysis, all the pixels in the interior of the photograph have zero value in the saturation channel (excluding the stamp). This can be visually illustrated in Figure 3.

Figure 3. On the left, the saturation channel (S) from the upper half of the image under analysis. On the right, the saturation channel from one of the images from the reference data set (Section 2.2).

The collected data set is appropriate to model black and white pictures developed between 1967 and 1969 using the standard chemical process at that time, properly aged over the years, and then scanned using different color scanning devices.

In Figure 4, we plot the histogram of the values of the S channel from the reference set of pixels, and used inspection and maximum likelihood estimation and quantitative best-fit comparisons [6] to find and fit a good distribution, in this case a Cauchy with $\widehat{\mu} = 0,10402560$ and $\widehat{\sigma} = 0,03617861$. The red curve in Figure 4 displays the shape of this Cauchy distribution.

The Cauchy CDF value for the Saturation at zero is 0.1065. According to developed model, this is the probability of a photograph pixel, under these conditions, having zero saturation. Finally, for a photograph region of 100×150 pixels, there are $N = 15,000$ pixels, and the probability that all of these pixels have zero saturation is of the order of $10^{-30,000}$. This very low probability show, with no doubt, that the photograph in the object in question is the result of a copy and paste operation.

2.4 Source Identification

In this section, we assess whether or not the analyzed image comes from a scanner from the *Public Archive of São Paulo*. As we stated in Section 2.2, documents from libraries' special collections do not leave their premises. In these situations, photocopier machines, or digital scanning devices, allow researchers to take home copies of the documents of interest.

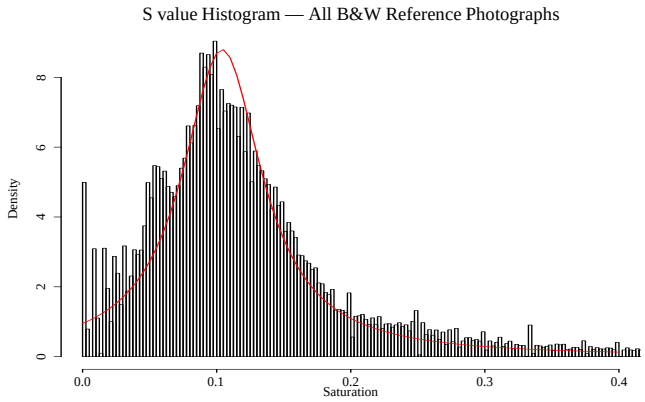

Figure 4. Histogram of the Saturation value (from HSV color representation) from the pixels of all B&W reference photographs in all the three available scanning devices. Additionally, in red we show the best fit parametric distribution (Cauchy) to approximate the data.

How can we check if a digital image was scanned by one of the devices available at the special collection premises? In the case we are describing in this article, the reporter had first alleged the document originated from the *DEOPS* special collection at the *Public Archive of São Paulo,* implying it was scanned there. Upon authorization, we had local access to the Archive's three available scanning equipment: HP 5470 (Scanner 1), HP G4050 (Scanner 2), and WideTek 25 (Scanner 3). We used a random set of images from their collection with the proper artifacts due to aging, storing, and mishandling conditions. Using this data, we were able to point out statistically what is the probability that the object of analysis was scanned in the *Public Archive of São Paulo*.

Recently, Khanna et al. [7] have proposed a technique to determine the source scanner that captured an image. This technique can point out which device, out of a known pool, is the most likely to have captured a test image. However, for this paper, we needed to verify the likelihood of one test image (the analyzed document) was captured, and not doctored later on, by each scanner. For that, we have extended Khanna?s solution.

The making process of any capturing device, be it a scanner or a camera, introduces several defects in the imagery sensors, and hence, it creates noise in the image pixels acquisition process. For forensics purposes, the noise of interest is divided in two types depending on the its impact on the final value of the captured pixels. The first kind of noise comes from localized defects, pixel traps, saturated pixels, or even no saturation at all. The second kind is called pattern noise, it comes from spatial noise that remains from one image to another and it is caused by dark currents and the device's non-uniform photo-responsiveness.

Vendors try to fix the first kind of noise with dedicated hardware and software within the capturing devices. The second type of noise if much harder to fix, and since it causes less variability to the pixel values, it is often ignored by the vendors. In the forensic scenario, we can use this pattern noise to describe an image-noise signature and consequently infer its acquisition source.

We have compared the compatibility of the noise-signature of the object in question to noise-signatures of the images captured as reference data from the Archive. In this analysis, we take into account if the object in question in this paper has undergone tampering operations after its acquisition.

Khanna et al [7] have described several features in order to represent the pattern noise more compactly. They have also proposed how to calculate the pattern noise signature through the statistics (moments) of noise values across the image rows and columns. Scanners have a periodicity between different rows of the fixed component of the sensor noise. To detect the similarity between different rows of the noise, the method uses the correlation between each of the N rows of the sensor noise with the average row.

As described in the paper, we have verified that the feature which measures the correlations among the color channels (F_{15}, in the paper's nomenclature) is not effective for images that have undergone JPEG compression and resampling.

In a difference from the original scenario presented in [7], which used machine learning techniques [8], we needed to verify and give probability estimates for each capturing device. In this scenario, a purely statistical analysis was more appropriate.

In order to calculate the pattern noise-signatures, we have considered a scenario in which each captured paper file image, in each scanner, is replicated in different JPEG compression levels (70%, 75%, 85%). We have used different compression levels due to the several JPEG compressions found in the object in question (Table 1). We have discarded the borders for the analysis and considered only centered regions of 450×300 pixels.

In Section 2.3, we have showed that there was a splicing operation of a background image and a photograph to create the upper half of the object under analysis. For source identification, we have used only the lower half of the object of Figure 1.

For each scanner and for each compression level, we have calculated the Mahalanobis [9] distance of the pattern noise-signature feature vector to the average pattern noise-feature vector of the reference set. We also have calculated the Mahalanobis distance of the average noise feature vector signature of the lower half of the object under analysis to the average signature of each capturing device.

Figure 5 presents the histogram of the Mahalanobis distances between the noise-signature of each image for its average, for several compression levels and scanners. Columns represent Scanner 1, Scanner 2, Scanner 3, and All Scanners, respectively. Rows represent JPEG compression levels of 70%, 75% e 85%. In red, we show the Mahalanobis distance of the object in question to the average signature in consideration.

Scanner 2 has generated images with poorer quality than the

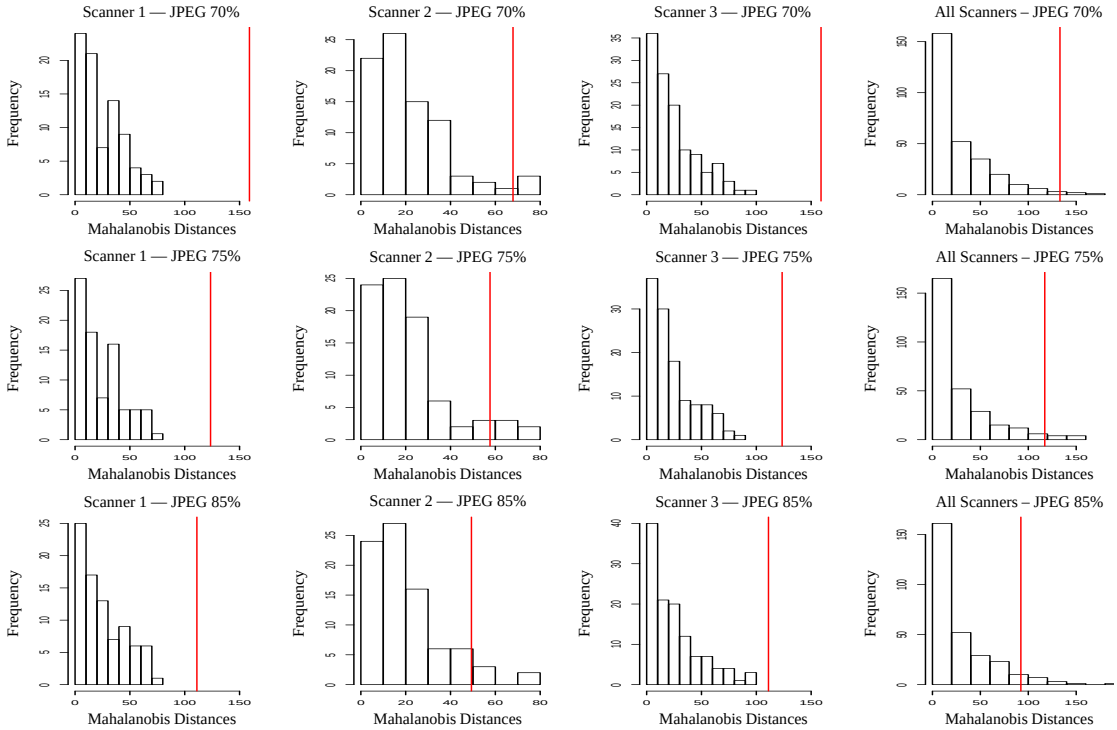

Figure 5. Histogram of Mahalanobis distances of each image's scanner-noise descriptor to the average scanner-noise descriptor. Columns represent scanner 1, scanner 2, scanner 3, and all scanners together respectively. Rows represent JPEG compression rates of 70%, 75% e 85%. In red, we show the distance value of the Mahalanobis of the image in question.

others, and generated more saturated pixels. Scanner 2 has a different statistical behavior than Scanner 1 and Scanner 3, as depicted in Figure 5. Scanners 1 and 3 histograms of Mahalanobis distances can be properly modeled by an exponential distribution (this analysis is not here), but this does not hold for Scanner 2. The image-noise from the object in question is not statistically compatible with the image-noise that comes from documents from Scanners 1 and 3.

Figure 6 allowed us to perform a further analysis. It displays the Mahalanobis distances of the pattern noise feature vectors for images captured on Scanner 2 under different resolutions, resampled at 96 DPIs, and then stored as JPEGs images with compression levels of 70%, 75% and 85% to the average signature of images coming from Scanner 2.

We calculated the CDF from the histogram-approximated PDF, without a parametric model. The probability that an image, captured by Scanner 2, that has not undergone any further manipulation, would have a Mahalanobis distance of the image-noise to the average signature of the group smaller than the distance of the object in question to the average of the group is 94.5%.

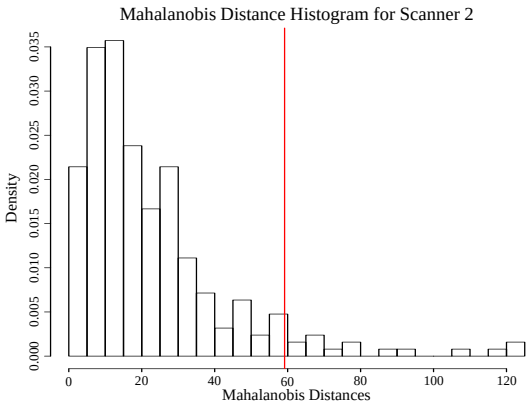

Figure 6. Histogram of Mahalanobis distances of each image's scanner-noise descriptor to the average scanner-noise descriptor including JPEG compressions of 70%, 75%, and 85% on scanner 2. In red, we show the distance value of the Mahalanobis of the image in question.

2.5 Letter Variability

In the object in question, the shape of the letters in the text are too regular. In scanned images of documents with typewritten text there is always some variability due to both the mechanical typing process and to the scanning noise, as can be seen in the lower half of Figure 7.

We focused our analysis only on the letter "a," since it is the most frequent character in the portuguese language, and allowed us obtain a larger set of examples. We manually iden-

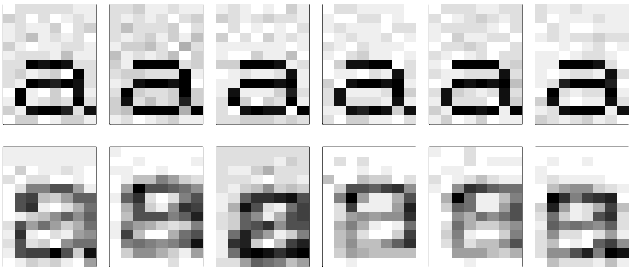

Figure 7. Letter variability analysis. Top: six random examples of character "a" from the object in question. Bottom: six random examples from character "a" from the reference data set.

tified 54 8×13-pixel regions in the lower half of the object in question containing the letter "a". Similarly, we selected 10 samples from the comparison data set (Section 2.2) that are representative of the expected properties of the whole set (96 DPIs and JPEG compression of 75%, the same as the object in question). On these 10 samples, we manually identified all the 190 8×13-pixel regions containing characters "a".

Figure 7 is visually convincing, but we performed a proper statistical analysis to show mathematically that the text is the result of digital manipulation.

We have represented each region as a 104-dimensional vector, and have calculated the covariance matrix of the two sets: the letters from the test object and the letters from the reference set. To compare their variability, we have looked into the sum of the eigenvalues of their covariance matrices. We have not assumed a parametric model, and we have used a 1,000 repetition *Bootstrap* [10] procedure to estimate the confidence intervals of the sum of the covariance matrix's eigenvalues of each of these groups, shown in Table 2. Even using high-confidence values intervals, there is no superposition, and we can state, beyond reasonable doubt, that the text in the object in question did not come from a scanning process of a typewritten document.

Confidence	Object Interval	Reference Interval
99%	(12.826, 18.799)	(102.520, 142.796)
99,9%	(11.998, 19.627)	(96.932, 148.384)
99,99%	(11.302, 20.323)	(92.240, 153.075)
99,9999%	(10.142, 21.483)	(84.414, 160.902)
99,999999%	(9.169, 22.456)	(77.854, 167.462)

Table 2. Confidence Intervals for Letter Variability.

3 Conclusions

In this case study, we have used both parametric and non-parametric statistical techniques. We were able to show that the object in question was indeed a fabrication. The photograph in the composed image is the result of a splicing operation from a different grayscale image, the text is the result of a digitally-manipulated insertion, it has not originated from a scanning procedure of a typewritten document, and there is only approximately 5% chance that parts of the object in question has been scanned in one of the available devices at the *Public Archive of São Paulo*.

This was not a difficult forensic case, but the analysis has to go beyond the initial gut feeling, and its findings have to give quantitative measures to substantiate its conclusions. The statistical methodology we used here remains solid beyond this paper, and allows a practitioner to explore the different possibilities to verify source, variability, and color consistency of a test object against a properly constructed reference data set.

Acknowledgments

We kindly thank FAPESP (Award Number 2008/08681-9 and 07/52015-0), and CNPq (Award Number 309254/2007-8 and 551007/2007-9) for the support of our related work.

References

[1] A. Rocha, *Classifiers and machine learning techniques for image processing and computer vision*. Phd thesis, University of Campinas (Unicamp), Campinas, Brazil, Mar 2009.

[2] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of re-sampling," *IEEE Transactions on Signal Processing (TSP)*, vol. 53, no. 2, pp. 758–767, 2005.

[3] A. Rocha and S. Goldenstein, "Progressive Randomization: Seeing the Unseen," *Computer Vision and Image Understanding (CVIU)*, 2010. To appear.

[4] B. Furht, ed., *Encyclopedia of Multimedia*. Springer, 2006.

[5] D. Salomon, *Data Compression: The Complete Reference*. Springer, 4^{th} ed., 2007.

[6] J. K. Lindsey, *Introductory Statistics: The Modelling Approach*. Oxford University Press, 2^{nd} ed., 2003.

[7] N. Khanna, A. Mikkilineni, and E. Delp, "Scanner identification using feature-based processing and analysis," *IEEE Transaction on Information Forensics and Security*, vol. 4, no. 1, pp. 123–139, 2009.

[8] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 1 ed., 2006.

[9] T. Hastie, R. Tibshirani, and J. Friedman, *Elements of Statistical Learning. Data Mining, Inference and Prediction*. Springer, 2001.

[10] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*. Chapman & Hall/CRC, 1994.