Digital Image Ballistics from JPEG Quantization

Hany Farid
Department of Computer Science
Dartmouth College
Hanover NH 03755

Abstract

Most digital cameras export images in the JPEG file format. This lossy compression scheme employs a quantization table that controls the amount of compression achieved. Different cameras typically employ different tables. A comparison of an image's quantization scheme to a database of known cameras affords a simple technique for confirming or denying an image's source. Similarly, comparison to a database of photo-editing software can be used in a forensic setting to determine if an image was edited after its original recording.

1 Introduction

When performing a forensic analysis on a digital image, it is often useful to be able to determine its source (e.g., a digital camera, a digital video camera, or a photo-editing software). Since the JPEG image format has emerged as a virtual standard, most devices and software encode images in this format. This lossy compression scheme allows for some flexibility in how much compression is achieved. Manufacturers typically configure their devices differently to balance compression and quality to their own needs and tastes. This difference, embodied in the JPEG quantization table, can be used to identify the source of an image.

In related work [3], the authors describe a technique to identify a specific camera based on sensor noise. Here, a much cruder identification will be made – at best, that of identifying the camera make and model. One benefit of this approach is that it can be applied to any device or software that employs JPEG compression.

2	1	1	1	1	1	2	1
1	1	2	2	2	2	2	4
3	2	2	2	2	5	4	4
3	4	6	5	6	6	6	5
6	6	6	7	9	8	6	7
9	7	6	6	8	11	8	9
10	10	10	10	10	6	8	11
12	11	10	12	9	10	10	10
2	2	2	2	2	2	5	3
3	5	10	7	6	7	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
2	2	2	2	2	2	5	3
3	5	10	7	6	7	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10
10	10	10	10	10	10	10	10

Figure 1: A sample JPEG quantization table employed by a Nikon Coolpix 2500 digital camera.

2 JPEG Compression

JPEG [1] is a standard lossy compression scheme [1, 6]. Given a three channel color image (RGB), compression proceeds as follows. The RGB image is first converted into luminance/chrominance space (YCbCr). The two chrominance channels (CbCr) are typically subsampled by a factor of two relative to the luminance channel (Y). Each channel is then partitioned into 8×8 pixel blocks. These values are converted from unsigned to signed integers (e.g., from $[0, 255]$ to $[-128, 127]$). Each block is converted to frequency space using a 2-D discrete cosine transform (DCT). Depending on the specific frequency and channel, each DCT coefficient, c, is then quantized by an amount q:

$$\hat{c} \quad = \quad \mathrm{round}(c/q). \tag{1}$$

This stage is the primary source of information loss. A final entropy encoding is then employed.

With some variations, the above sequence of steps are employed by JPEG encoders in digital cameras and photo-editing software. The primary source of variation in these encoders is the choice of quantization that subsequently controls the compression rates and artifacts. The quantization is specified as a table of 192 values – a set of 8×8 values associated with each frequency, for each of three channels (YCbCr). For low compression rates, these values tend towards a value of 1, and increase for higher compression rates. Shown in Figure 1, for example, is the quantization table employed by a Nikon Coolpix 2500 digital camera. Shown, from top to bottom, are the luminance and chrominance channel tables. As is typical, the quantization for the luminance channel is less than for the two chrominance channels, and the quantization level is less for the lower frequency components.

With 192 entries, and values ranging typically from 1 to 25 there are 10^{268} possible quantization tables – in practice, this number is much smaller because the entries are not chosen independent of one another.

3 Ballistics

A single image from each of 204 digital cameras was collected, Table 5 [2]. While some cameras have only a single JPEG quality setting, many others have multiple quality settings. In this current analysis, only a single image at the highest quality was examined.

The JPEG quantization table was extracted from each image and compared for uniqueness. On average, each camera matched 1.43 other cameras. 62 of the 204 cameras had a unique table. The remaining cameras fall into equivalence classes of sizes between 2 and 28, Table 5. In the majority of cases, it is cameras from the same manufacturer that share the same quantization tables. In many cases, only two or three cameras are indistinguishable – the Canon PowerShot family is the

[1] The Joint Photographic Experts Group (JPEG) developed an image file format known as the JPEG File Interchange Format (JFIF), which is most commonly abbreviated as JPEG.

[2] Most of the images were downloaded from the digital photography review site `http://www.dpreview.com`. The metadata of each downloaded image was checked to make sure that the image was not subsequently edited in a photo-editing software, which would have compromised the original JPEG quantization table.

exception, where 28 cameras have identical quantization tables. And in several cases, different makes and models share the same quantization table, with the largest diversity finding Kodak, Nikon, Olympus, Pentax, and Sony cameras in the same equivalence class.

While the JPEG quantization is clearly not perfectly unique, it is reasonably effective at narrowing the source of an image to a single camera make and model or to a small set of possible cameras (on average, this set size is 1.43).

3.1 Software Ballistics

Photo-editing software packages (e.g., Adobe Photoshop, Corel Paint Shop, Microsoft Digital Image Suite, etc.) employ different JPEG quantization tables. Photoshop, unarguably the most popular such package, provides 13 levels of compression (12: lowest compression & highest quality – 0: highest compression & lowest quality). An image was saved at each of these levels for Photoshop CS2 (the most recent version), Photoshop CS, Photoshop 7, Photoshop 4 and Photoshop 3. The quantization tables at each compression level were different from one another. Moreover, at each compression level the tables were the same for all five versions of Photoshop [3]. More importantly, the tables differed from each the 204 digital cameras listed in Table 5. An image can therefore be linked to a specific software (or possibly a set of software packages).

Similar in spirit to forensic techniques for detecting double JPEG compression [2, 4], the presence of quantization tables unique to a photo-editing software can be used to determine that an image was previously viewed and saved from within a photo-editing software. The reliability of such a claim will, of course, require the uniqueness of the software tables to hold over a larger set of cameras.

3.2 Video Ballistics

The MPEG video standard (MPEG-1 and MPEG-2) employs two basic schemes for compression to reduce both spatial redundancy within individual video frames and temporal redundancy across video frames [5]. In a MPEG encoded video sequence, there are three types of frames: intra (I), predictive (P) and bi-directionally predictive (B), each offering varying degrees of compression. These frames typically occur in a periodic sequence. A common sequence, for example, is:

$$I_1\ B_2\ B_3\ P_4\ B_5\ B_6\ P_7\ B_8\ B_9\ P_{10}\ B_{11}\ B_{12}\ I_{13}\ B_{14}\ \cdots,$$

where the subscripts are used to denote time.

I-frames are encoded using the JPEG compression scheme described in Section 2. P-frames are encoded using motion estimation. The motion between a P-frame and its preceding I- or P-frame is estimated. A motion estimated version of the next frame is generated by warping the P-frame according to the estimated motion. The error between this predicted frame, and the actual frame is then computed. Both the motion vectors and the motion errors are encoded and transmitted – the motion errors are encoded using JPEG compression. The encoding of a B-frame is similar to that of a P-frame except that motion estimation for these frames employ past, future, or both of its neighboring I- or P-frames.

In an MPEG video sequence, the I-frames and the motion errors for the P- and B-frames are each JPEG encoded. As such, the quantization tables for these frames may be used to identify the source of a video recording.

4 Discussion

By choosing a JPEG quantization table camera manufacturers and software developers are able to balance the amount of compression and the quality of their images. In so doing, a signature of sorts is embedded within each JPEG image. This simple observation allows for a rather crude form of digital image ballistics, whereby the source of an image can be confirmed or denied.

In order for this ballistic technique to be more effective and reliable, it is important to collect data from an ever-increasing number of cameras and software. In addition, it is important to consider the quantization tables for each JPEG quality setting afforded by a specific device or software.

In collaboration with other ballistic and forensic tools, this technique should prove useful for the forensic examination of digital images and video.

5 Acknowledgments

Thanks to Paul McJones and Sean Parent of Adobe Systems for their assistance extracting the JPEG quantization tables from older versions of Adobe Photoshop. Thanks to Jeff Woodward for writing the batch processing scripts for extracting JPEG quantization tables.

This work was supported by a Guggenheim Fellowship, a gift from Adobe Systems, Inc., a gift from Microsoft, Inc., a grant (FA8750-06-C-0011) from the United States Air Force, and under a grant (2000-DT-CX-K001) from the U.S. Department of Homeland Security, Science and Technology Directorate (points of view in this document are those of the authors and do not necessarily represent the official position of the U.S. Department of Homeland Security or the Science and Technology Directorate).

[3] More recent versions of Photoshop afford more JPEG compression levels. The quantization tables found in older versions, however, are the same as in the most recent versions of Photoshop.

References

[1] Digital compression and coding of continuous-tone still images, Part 1: Requirements and guidelines. ISO/IEC JTC1 Draft International Standard 10918-1, 1991.

[2] J. Lukas and J. Fridrich. Estimation of primary quantization matrix in double compressed JPEG images. In *Digital Forensic Research Workshop*, Cleveland, Ohio, August 2003.

[3] J. Lukas, J. Fridrich, and M. Goljan. Digital camera identification from sensor noise. *IEEE Transactions on Information Security and Forensics*, 1(2):205–214, 2006.

[4] A.C. Popescu and H. Farid. Statistical tools for digital forensics. In *6th International Workshop on Information Hiding*, Toronto, Cananda, 2004.

[5] T. Sikora. *Digital Consumer Electronics Handbook*, chapter MPEG-1 and MPEG-2 Digital Video Coding Standards. McGraw-Hill, 1997.

[6] G.K. Wallace. The JPEG still picture compression standard. *IEEE Transactions on Consumer Electronics*, 34(4):30–44, 1991.

Canon	PowerShot A510	Canon	PowerShot A520	Canon	PowerShot A620
Canon	IXUS 50	Canon	PowerShot S2 IS	Canon	PowerShot S70
Canon	IXUS 55	Canon	IXUS 700	Canon	IXUS 750
Canon	IXUS 800 IS	Canon	PowerShot A20	Canon	PowerShot A70
Canon	PowerShot A700	Canon	PowerShot A95	Canon	EOS 10D
Canon	EOS-1D	Canon	EOS-1D Mark II	Canon	EOS-1DS
Canon	Canon	Canon	EOS 20D	Canon	EOS 300D
Canon	EOS 30D	Canon	EOS 350D	Canon	EOS 5D
Canon	EOS D30	Canon	EOS D60	Canon	PowerShot G1
Canon	PowerShot G2	Canon	PowerShot G3	Canon	PowerShot G5
Canon	PowerShot G6	Canon	PowerShot Pro1	Canon	PowerShot Pro90 IS
Canon	PowerShot S10	Canon	IXUS	Canon	IXUS v
Canon	PowerShot S1 IS	Canon	PowerShot S20	Canon	PowerShot S30
Canon	IXUS 300	Canon	IXUS 330	Canon	PowerShot S40
Canon	IXUS 400	Canon	PowerShot S45	Canon	PowerShot S50
Canon	PowerShot S60	Canon	IXUS II	Canon	IXUS 40
Casio	EX-P700	Casio	EX-Z3	Casio	EX-Z750
Casio	QV-3000EX	Casio	QV-4000	Epson	PhotoPC 3000Z
Fuji	FinePix40i	Fuji	FinePix4700	Fuji	FinePix4900
Fuji	FinePix6800	Fuji	FinePix6900	Fuji	FinePix F810
Fuji	FinePix E550	Fuji	FinePix F10	Fuji	FinePix F30
Fuji	FinePix F601	Fuji	FinePix F700	Fuji	FinePix S1Pro
Fuji	FinePix S2Pro	Fuji	FinePix S3Pro	Fuji	FinePix S500
Fuji	FinePix S5500	Fuji	FinePix S602	Fuji	FinePix S7000
Fuji	FinePix S9500	Fuji	FinePix2700	Fuji	MX-2900ZOOM
Hewlett-Packard	PhotoSmart C812	Hewlett-Packard	PhotoSmart C850	Hewlett-Packard	PhotoSmart C935
Hewlett-Packard	PhotoSmart R707	Kodak	DC290	Kodak	DC3200
Kodak	DC4800	Kodak	DCS Pro 14N	Kodak	DX4900
Kodak	DX7590	Kodak	P850	Kodak	P880
Kodak	V610	Kodak	Z650	Kodak	Z740
Konica Minolta	MAXXUM 7D	Konica Minolta	DiMAGE A200	Konica Minolta	DiMAGE A2
Konica Minolta	DiMAGE Z2	Konica Minolta	DiMAGE Z5	Kyocera	FC-S3
Lieca	DIGILUX 2	Minolta	DiMAGE 5	Minolta	DiMAGE 7
Minolta	DiMAGE 7Hi	Minolta	DiMAGE 7i	Minolta	DiMAGE A1
Minolta	DiMAGE F100	Minolta	DiMAGE S304	Minolta	DiMAGE S404
Minolta	DiMAGE X	Nikon	E2500	Nikon	E3100
Nikon	E4500	Nikon	E4800	Nikon	E5000
Nikon	E5200	Nikon	E5400	Nikon	E5700
Nikon	E700	Nikon	E775	Nikon	E7900
Nikon	E800	Nikon	E8400	Nikon	E8700
Nikon	E880	Nikon	E8800	Nikon	E885
Nikon	E950	Nikon	E990	Nikon	E995
Nikon	COOLPIX P3	Nikon	D1	Nikon	D100
Nikon	D1H	Nikon	D200	Nikon	D2H
Nikon	D2X	Nikon	D50	Nikon	D70
Olympus	C2020Z	Olympus	C2100UZ	Olympus	C3030Z
Olympus	C3040Z	Olympus	C40Z,D40Z	Olympus	C5050Z
Olympus	X-2,C-50Z	Olympus	C70Z,C7000Z	OlympusD	C700UZ
Olympus	C8080WZ	Olympus	E-1	Olympus	E-10
Olympus	E-20	Olympus	E-300	Olympus	E-330
Olympus	E-500	Olympus	SP310	Olympus	SP500UZ
Olympus	uD800,S800	Olympus	u-miniD	Panasonic	DMC-FX01
Panasonic	DMC-FZ30	Panasonic	DMC-FX7	Panasonic	DMC-FX9
Panasonic	DMC-FZ20	Panasonic	DMC-FZ3	Panasonic	DMC-FZ5
Panasonic	DMC-LX1	Panasonic	DMC-LZ2	Panasonic	DMC-TZ1
Pentax	Optio A10	Pentax	ist D	Pentax	ist DS
Pentax	Optio 330	Pentax	Optio 430	Pentax	Optio 550
Pentax	Optio 750Z	Pentax	Optio S	Pentax	Optio S5i
Ricoh	GR	Samsung	Digimax V700	Sigma	SD10
Sony	MVC-CD200	Sony	MVC-CD300	Sony	DSC-D700
Sony	DSC-F505V	Sony	DSC-F707	Sony	DSC-F717
Sony	DSC-F828	Sony	DSC-F88	Sony	DSC-H1
Sony	DSC-H2	Sony	DSC-H5	Sony	DSC-L1
Sony	DSC-P1	Sony	DSC-P150	Sony	DSC-P200
Sony	DSC-P5	Sony	DSC-P71	Sony	DSC-P9
Sony	DSC-R1	Sony	DSC-S70	Sony	DSC-S75
Sony	DSC-S85	Sony	DSC-S90	Sony	DSC-V1
Sony	DSC-V3	Sony	DSC-W7	Sony	DSLR-A100

Table 1: Make and model of 204 cameras.

28	Canon PowerShot A510 – Canon PowerShot A520 – Canon PowerShot A620 – Canon IXUS 50 – Canon PowerShot S2 IS – Canon PowerShot S70 – Canon IXUS 55 – Canon IXUS 700 – Canon IXUS 750 – Canon IXUS 800 IS – Canon PowerShot A20 – Canon PowerShot A70 – Canon PowerShot A700 – Canon PowerShot G2 – Canon PowerShot G3 – Canon PowerShot G5 – Canon PowerShot G6 – Canon PowerShot Pro1 – Canon IXUS v – Canon IXUS 300 – Canon IXUS 330 – Canon PowerShot S40 – Canon IXUS 400 – Canon PowerShot S45 – Canon PowerShot S50 – Canon PowerShot S60 – Canon IXUS II – Canon IXUS 40
23	Kodak P850 – Nikon E5400 – Olympus C2100UZ – Olympus C3030Z – Olympus C40Z,D40Z – Olympus C5050Z – OlympusD C700UZ – Olympus C8080WZ – Pentax Optio 550 – Pentax Optio 750Z – Sony DSC-F505V – Sony DSC-F707 – Sony DSC-F717 – Sony DSC-H1 – Sony DSC-H2 – Sony DSC-H5 – Sony DSC-P150 – Sony DSC-P200 – Sony DSC-R1 – Sony DSC-S90 – Sony DSC-V1 – Sony DSC-V3 – Sony DSC-W7
16	Canon EOS 10D – Canon EOS 20D – Canon EOS 300D – Canon EOS 30D – Canon EOS 350D – Canon EOS D30 – Canon EOS D60 – Sony MVC-CD200 – Sony MVC-CD300 – Sony DSC-P1 – Sony DSC-P5 – Sony DSC-P71 – Sony DSC-P9 – Sony DSC-S70 – Sony DSC-S75 – Sony DSC-S85
8	Konica Minolta DiMAGE Z2 – Minolta DiMAGE F100 – Nikon E2500 – Nikon E4500 – Nikon E5000 – Nikon E5700 – Nikon E8800 – Nikon E990
7	Nikon D100 – Nikon D1H – Nikon D200 – Nikon D2H – Nikon D50 – Nikon D70 – Panasonic DMC-FX01
6	Kodak P880 – Nikon D2X – Olympus C3040Z – Olympus C70Z,C7000Z – Olympus SP310 – Olympus SP500UZ
6	Konica Minolta DiMAGE A2 – Minolta DiMAGE 5 – Minolta DiMAGE 7 – Minolta DiMAGE 7i – Minolta DiMAGE S304 – Minolta DiMAGE S404
5	Kodak DC4800 – Kodak DX4900 – Kodak V610 – Kodak Z650 – Kodak Z740
5	Nikon E4800 – Nikon E775 – Nikon E8400 – Nikon E8700 – Sigma SD10
4	Canon IXUS – Casio EX-P700 – Casio EX-Z750 – Pentax Optio S
3	Minolta DiMAGE X – Nikon E7900 – Nikon E950
3	Nikon E3100 – Nikon E880 – Nikon E995
3	Olympus E-300 – Olympus E-330 – Olympus E-500
3	Panasonic DMC-FX7 – Panasonic DMC-FZ20 – Panasonic DMC-FZ5
2	Canon EOS-1D – Canon EOS-1DS
2	Canon EOS-1Ds Mark II – Canon EOS 5D
2	Canon PowerShot G2 – Canon PowerShot Pro90 IS
2	Epson PhotoPC 3000Z – Konica Minolta DiMAGE A200
2	Fuji FinePix S7000 – Fuji FinePix2700
2	Hewlett-Packard PhotoSmart C850 – Hewlett-Packard PhotoSmart C935
2	Minolta DiMAGE 7Hi – Minolta DiMAGE A1
2	Nikon E5200 – Nikon COOLPIX P3
2	Panasonic DMC-FZ30 – Panasonic DMC-LX1
2	Pentax ist D – Pentax ist DS
2	Sony DSC-F828 – Sony DSC-F88

Table 2: Each entry corresponds to cameras with identical JPEG quantization tables.