

Search optimization through JPEG quantization tables

Sharon Gieske (6167667)

June 3, 2014

Introduction

Taking pictures is very easy and popular in this digital age. The demand for digital cameras was forecast to be 86 million units for 2013 by Futuresource Consulting [2]. And even though a decline in market share is present for digital cameras, due to the proliferation of smartphones (nowadays all equipped with a camera function) the total number of digital images taken each year is very high. Social media sites which have photo upload functions, such as Facebook and Instagram, report significantly huge numbers on the total upload of images. Facebook alone reported in a white paper [1] that more than 250 billion photos are uploaded to their site, with on average a total upload of more than 350 million photos every day. Statistics on Instagram¹ show a total of 20 billion photos shared on Instagram.

Due to this popularity, digital images are often recovered in forensic investigation. For example, in child pornography cases many digital images are present and are important evidence for the investigation. In such a case it can be very important to identify the origin of images to a specific camera or identify images that come from a common source. This can be done by uncovering traces on pictures which are distinguishable for camera models. One of these traces is the JPEG quantization table which is specified as a table of 192 values: set of 8 · 8 values associated with each frequency, for each of three channels (YCbCr)

With over a dozen different camera brands, each developing different models over the years, the number of camera models (and consequently the number of JPEG quantization table) is significantly high. This increases the search space for matching images to camera models. The matching of large databases of images against the camera models will be time costly but needs to be minimized since time is often limited in forensic investigations. This research will focus on optimizing search through the image databases regarding JPEG quantization tables.

Research Question

The research question on which is focused is set as: *‘How can search through JPEG quantization tables be optimized?’*

In order to answer the research question, this research will focus on the following subquestions:

1. What are identifiable parameters of JPEG quantization tables?
2. How can we quickly decrease the search space for JPEG quantization table matches?

¹<http://instagram.com/press/>

Related Work

Approach & Methods

Requirements

Planning

The planning for this research is proposed in Table 1.

Week 1 (2-8/06)	
Literature research	
6	Deadline project proposal
Week 2 (9-15/06)	
Week 3 (16-22/06)	
Week 4 (23-29/06)	
Week 5 (30/06 - 06/07)	
2	Presentation
4	Deadline report

Table 1: Project planning including dates and tasks

Expected Product

References

- [1] *A Focus on Efficiency*. White Paper, September 2013.
- [2] Futuresource Consulting. *Digital cameras in decline, though interchangeable lens segment sees growth*. Press Release, November 2013. <http://www.futuresource-consulting.com/2013-11Cameraspressrelease.html>.