31

A New Approach to Integrity of Digital Images

D. Storck
Institute for Computer Graphics, Fraunhofer Gesellschaft
Wilhelminenstr. 7, 64283 Darmstadt, Germany,
+49 6151 155 418 (444 fax),
email: schoko@igd.fhg.de, URL: http://www.igd.fhg.de/~schoko

Abstract

This paper describes a new approach to integrity for digital images. Different to the most mechanisms, this approach describes a method which is resistant against a loss of integrity by slight modifications like compression or conversion. The suggestion to archive this goal is a transformation to the frequency-domain similar to the jpeg-compression.

Keywords

Integrity to digital images, frequency-domain

1 INTRODUCTION

Ever since the invention of the photograph, people have believed in what they saw in pictures. The same goes for newspapers and, to a much higher degree, for TV news coverage. Supported by existing techniques and more and more affordable hardware to perform these on, creating photo-realistic scenes which have never happened or manipulating real pictures becomes more and more easy. This may be a great opportunity for entertainment. Think about a movie with Tom Hanks and Marilyn Monroe. Two people who've never acted together in the same movie star in a perfect illusion (as has already been shown in „Forrest Gump" or, even before that, in Woody Allen's „Zelig"). On the other hand, the door for misuse is ajar. Pictures of famous people can be modified to provide a different information. *„Our lack of clarity produces both overly optimistic (trusting) and overly pessimistic perceptions. In the world of digital information, the tools and mechanisms of ensuring integrity are complex and exotic, and our unfamiliarity with these tools leads us to distrust their efficacy. Thus we regard the digital information environment as basically lacking integrity. These feelings are supported by, for example, the arresting examples of image manipulation in the digital domain, which run counter to our assumptions about photographic images as recorded vision*

J. L. Encarnação et al. (eds.), *Mobile Communications*

(Mitchell 1992)." (Clifford 1994). To avoid such forgery and fraud, certain mechanisms must be used to check the integrity and origin of pictures. Common techniques used for textual data may also be used for images, but there are several disadvantages involved in this approach. The traditional techniques used to provide integrity detect every modification in the image data to some extent. This means a scaled or gamma-corrected image will lose its integrity, which may not be the intention of the sender or originator of the picture. Integrity should only be corrupted if the semantic information of the picture is modified.
The department Security Technology for Graphics and Communication Systems of the Fraunhofer Institute of Computer Graphics is developing a method to detect manipulations of images which modify the semantic information of the images. This work is part of the PLASMA-project, an object-oriented security-platform for multi-media data. Modifications which affect the quality (like gamma-correction, scaling or compression) should not necessarily trigger a detection mechanism. The new approach makes it necessary to transform the image data into the frequency domain. This is done by a Discrete Cosine Transformation (DCT) of 8x8 pixel blocks. The coefficients resulting from this transformation have different influence on the picture. The first value (DC coefficient) and the next few AC coefficients are the most important factors in the appearance of the picture. The higher AC values have only little perceptual influence on the picture and are often modified by compression standards to get good compression rates. Therefore these values of the higher frequencies are irrelevant for the integrity check. Severe manipulation of the picture always affects the lower frequencies and the DC value, therefore the protection of integrity is based on these coefficients.

2 RELATED WORK

A very good overview to the new problems of integrity of digital information is given in (Clifford 1994). Many parts of the next two sections are taken from this paper.
One of the key advantages of digital information is that each copy is of equal quality to the original. Indeed the copies are identical to and indistinguishable from the original digital object, whether the object represents text, images audio video, or executable computer software. But anyone who is involved with computers quickly learns that it is simple to alter bits or characters in a file and pass it to others as it would be original. The amount of tools allowing to change files in a sophisticated manner has increased significantly over the last ten years.
To avoid such fraud, a number of „digest" algorithms have been developed which compute a relatively brief „hash" or digest of a given digital file, typically considerably less than 1000 bytes in size. Most famous algorithms were developed by Rivest. They are known by the identifier MD4 and MD5 (Rivest 1992).
The characteristics of these algorithms are such that it is extremely improbable for any two files to generate the same digest, particularly if the two files are „close" in content. It is possible to approximate mathematically the likelihood that two different files would produce the same digest; this probability is very small. The problem with most of these algorithms is that they only work if one assumes that the integrity is maintained only if there is a bit-for-bit equivalence between the copy and the original data. Unfortunately, this definition is often too limited to be of much use. With the continued proliferation of storage and interchange standards for various types of digital objects, and the development and deployment of

increasingly intelligent systems to convert them from one format to another, bit-by-bit equivalence will lose much of its value as a definition of „identical" objects in the future (Clifford 1994). As Clifford describes in his paper: *„We have very few tools to help us deal with digital objects at the level of abstraction of intellectual content. Our understanding of digest algorithms, for example does not extend beyond bit-level equivalence to deeper notions of document content. The development of digest algorithms at the intellectual content level (or even heuristic digest algorithms that measure similarity rather than precise equivalence, and that perhaps only work on limited classes of material, or only work on some rather than all instances of such objects) would represent a substantial breakthrough. In my view, the development of such methods, as well as a general theoretical framework within which to categorize them and to analyze their performance, represents a challenging but important area for ongoing research."*
Surely, a realization of the approach described in this paper will not be the perfect solution for all these problems. But it may be a first step towards an algorithm for integrity of digital images which is robust against several modifications. In the following section some requirements are spelled out. They describes modifications to pictures which destroy the bit-to-bit equivalence between copy and original image, but do not destroy the semantic information of the image.

3 REQUIREMENTS TO AN ALGORITHM FOR THE INTEGRITY OF DIGITAL IMAGES

As described above, in the world of digital information algorithms are often applied to digital images to translate them from one form to another. These algorithms change the bytes in an digital image but do not touch the semantic information. They can be roughly categorized into two classes: lossless and lossy. In the case of lossless algorithms, no information is lost, and normally the original data can be reconstructed from the converted file. In the case of lossy conversion some information of the image data gets lost. But if it is only a loss of the coefficients from high frequencies, the difference between the original and the altered copy of the image can only be detected by the most discerning eye. In the most cases not even by any human being.
For example the popular JPEG compression algorithm supports some lossy compression options which discard information. In fact, typically JPEG compressors let the user manipulate settings that control the degree of loss permitted in compression, with higher settings for permissible loss producing a greater degree of compression. Whether information is actually lost in a specific case depends on the settings used in JPEG compression and the content of the original image. But it may be difficult or impossible for the user receiving a JPEG-compressed image to be aware of or to detect this loss of information.
According to these given parameters, the integrity of an image is not destroyed by this methods and an algorithm for the integrity of digital images should be robust against conversion and lossy compression (especially the JPEG compression).
Due to the differences of the most computer monitors it may be necessary to correct the brightness of a received image. There are many tools available to adjust the brightness of an image to a specific CRT. These gamma-corrections do not destroy the semantic information and therefore should not affect the integrity of the image.

If an image is scaled or enhanced by interpolation to a higher pixel density there is no loss of information (at least no loss of semantic information, only if the image is scaled down to an icon) if one knows how the interpolation is done, since it should be possible to reconstruct the original. Due to this it should be possible, that the integrity is not lost by (at least) scaling an image.

3.1 Requirements - Summary

An algorithm that provides integrity checks for digital images should detect any manipulation that destroys or changes the semantic information of the image. The following manipulations do not effect the semantic information and there should not be a loss of integrity if they are used.

- Format-conversion
- Compression (especially JPEG)
- Gamma-Correction
- Scaling

4 SPECIFICATION

4.1 The frequency domain and the JPEG-compression

As described at the beginning, with the JPEG-compression the pixel information of an image is transformed into the frequency domain. The resulting coefficients affect the appearance of the image in different ways of which the DC-value and the coefficients of the low frequencies are the primary influences. The coefficients of the medium and the high frequencies have only little effect to the information of the image (Hung 1993).
These different influences are important for a first approach to integrity of digital images. It seems to be sufficient to use only the DC-value and perhaps the coefficients of the lowest frequencies for the integrity (see figure 1). The final solution will be determined by future tests and fine-tuning. In this first approach (only the DC-value is used) the robustness against compression on the high frequencies or other low-pass filtering should be achieved. Further tests have to been made to check if this approach is also robust against manipulations. To avoid that color modification like gamma correction destroy the integrity it is necessary to transform the RGB-values after the Discrete Cosine Transformation into YUV-components. The Y-values are not affected by changes in the color space, so they are used to build the message which is protected by a common integrity mechanism.
If already slight modifications to the image changes the DC-value of the Y-components it becomes necessary to encode this value in a special way. A possible solution for this encoding is shown in the next section.

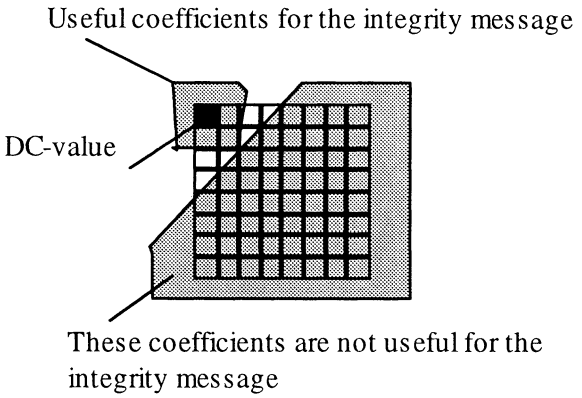

Figure 1 Not all coefficients are useful to built the integrity check message. The grid shows the DC-value and the 63 AC-values. Only the DC-value and few of the lowest frequency coefficients are used in the approach described.

4.2 Building the integrity check message

To create the integrity check message the common techniques are used (Schneier 1994). But different to the common approaches not the complete data is used to build the message. As shown in figure 2, only some data is extracted from the image and used with the classic algorithms to create a signature which allows to check the integrity of the image.

Encoding of the coefficients

If the described approach is not robust enough against slight modifications by using the DC-values directly it becomes necessary to encode the coefficients. Once this has been done, the hash-value (using MD4 or MD5) is not built directly from the DC-values, but rather built from the result of the *integer*-function of the difference between two following DC-values. Even if slight modifications of the image affect the difference between two DC-values, it is not to expected that the sign of the difference between the DC-values of two following DCT-blocks is changed. This robustness assumes that the modification is done to the complete image. If the modification is confined to a region of the image or to manipulation of some objects from the image, it can be expected that some differences change completely and the integrity will be lost.

In this way, it is possible to protect an image against the loss of integrity if there is a manipulation of the semantic information of the image (This assumes that the semantic information is related to the shape and the position of the object and not to their color.). Figure 3 illustrates the encoding of the DC-values.

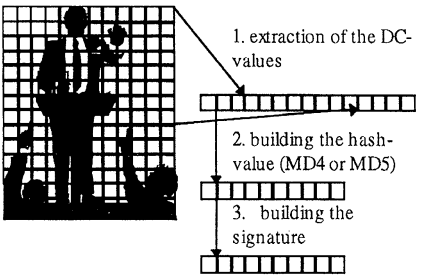

Figure 2 Creation of the signature for protection of integrity. First the DC-values are extracted. These values are the input for a MD4 or MD5 hashing function. The hash value can be signed with any common public-key algorithm like the RSA.

Integrity and scaling

An integrity protected image with the techniques described above, would clearly lose its integrity by scaling. This can be avoided by adding the original size as part of the signed information if the integrity of a picture should be checked. It will become necessary to restore the original image size before checking the integrity. The following example shows the possible structure of an integrity check message for images.

```
typedef struct {
  int width, height;
  BOOL difference_flag;
  Signature SizeSig;
  Signature ImageSig;
}ImageSignature;
```

The *difference_flag* will be set to TRUE if the image signature is calculated from the results of the *integer*-function from two following DC-values. If the signature is calculated from the DC-values directly the flag is set to FALSE.

Integration of the image signature

The signature of the image should not appear as a separate file and has to been integrated directly into the image file. The different file formats offer several ways to integrate additional data. In GIF-files the signature can be put in the **application extension block** and in JFIF-files (the common exchange format for JPEG-compressed images) the signature can be integrated after a so called **APP0-marker**.

Sadly, some imaging tools do not write back these values if an image is stored. So the integrity of an image may be lost because the signature was not stored by an application even if there were no changes to the image. In this case the signature in a separate file may be helpful.

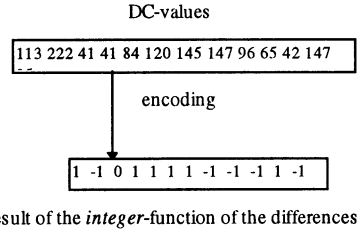

Figure 3 The encoding of the DC-values may become necessary if already slight modifications produce a loss of integrity.

5 CONCLUSION

This new approach may be a step towards a new understanding of integrity for digital information. Even if the first tests produced the expected (promising) results, a lot of work still has to be done. The first realization works fine with simple pictures as shown in figure 4. The different possibilities to create the integrity check message (selection of the coefficients to calculate the signature, calculate the differences between two values, etc...) allow to tune the mechanisms for special requirements.
Future test will show if it is possible to create completely different pictures which produce the same signature.

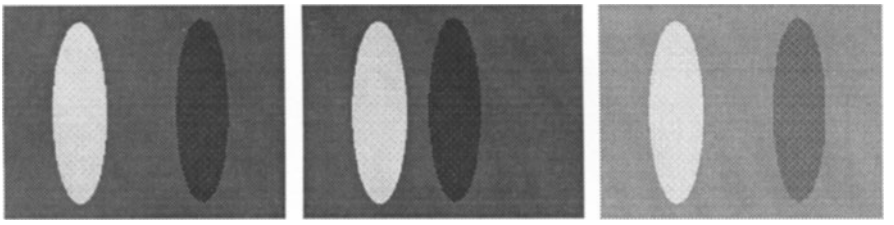

Figure 4 The new approach already produces the expected results with simple images. Part a) shows the original image. Part b) shows the same image, but the distance between the two objects was reduced. Part c) is a brightened version of the original image. The tests show that the integrity is still correct in image c). The integrity check for image b) fails because 20% of the calculated message differ from the message produced by the original image. With common mechanisms like the MD5 hashing function and the RSA algorithm the integrity check would fail for both images.

6 REFERENCES

Clifford, A.L. (1994) The Integrity of Digital Information: Mechanics and Definitional Issues. *Journal of the American Society for Information Science* 737-744, 45(10).

Hung, A..C. (1993) PVRG-JPEG Codec 1.1. Portable Video Research Group. Stanford University

Mitchell, W.J. (1992) The reconfigured eye: Visual truth in the post photographic era. Cambridge, MA: MIT Press

Mitchell, W.J. (1994) When is seeing believing? *Scientific American*, Feb 1994.

Rivest, R.L. (1992) RFC 1321: MD5 message digest algorithm

Schneier, B. (1994) Applied Cryptography. *John Wiley and Sons.*

7 BIOGRAPHY

Dietmar Storck has earned a Dipl.-Inform. from the Technical University Darmstadt in 1993. Since then he is employed at the Fraunhofer Institute of Computer Graphics. He is a member of the department „Security Technology for Graphic- and Communication-Systems". His research activities are dealing with encryption of multi-media data and special methods for the security of image data. Further interests includeaccess control and billing of multimedia services via the internet.