LOW DISTORTION AND ROBUST STEGANOGRAPHY ON PARALLEL ARCHITECTURE CELL BE USING A SHARED COLOR PALETTE AND A SHARED HASH

Bogdan ŢIGĂNOAIA[1], Francisc IACOB[2]

Această lucrare prezintă o analiză steganografică critică şi un studiu al arhitecturii CELL BE. Este descrisă o tehnică pentru ascunderea informaţiei în imagini color nouă, de capacitate ridicată şi robustă. Distorsiunea imaginii după encodarea mesajului este foarte redusă şi indistinctibilă ochiului uman. Sistemul de codare-decodare foloseşte trei informaţii partajate: un hash al imaginii pentru protecţia împotriva atacurilor steganografice, o paletă de pixeli (aceasta este rezultatul aplicării algoritmului de clusterizare K-means, cu câteva modificări) folosită pentru encodarea mesajului şi dimensiunea unei piese de imagine. Arhitectura paralelă Cell BE este folosită pentru calcularea hash-ului şi a paletei de pixeli. Este propusă o mapare eficientă a aplicaţiei pe arhitectura CELL BE. Sunt folosite noi facilităţi de programare CELL BE cum ar fi: double buffering, instrucţiuni SIMD, mailboxes, evenimente, transferuri DMA.

This paper presents a critical steganographic analysis and a study of the Cell BE architecture. It also describes a new high capacity and robust technique for hiding information on color images. The image distortion after encoding the message is low and indistinguishable to the human eye. The encoding and decoding system uses three shared information: an image hash for protection against steganographic attacks, a pixels palette (this is the result of applying the K-means clustering algorithm, with some modifications) used for encoding the message and the dimension of one piece of the image. Parallel architecture Cell BE is used for the calculation of the hash and the pixels palette. An efficient mapping of the application on the Cell BE architecture is proposed. New Cell BE programming facilities such as double buffering, SIMD instructions, mailboxes, events, DMA transfers are used.

Keywords: steganography, CELL BE, parallel architectures

1. Introduction

Image processing domain presents new challenges for researchers. Parallel architectures offer a high processing power and represent a new and good tool for new studies. Tools for the analysis of the parallel applications represent today an

[1] Ph.D. student, Computer Science Department, University POLITEHNICA of Bucharest, Romania, e-mail: bogdantiganoaia@yahoo.com

[2] Prof., Computer Science Department, University POLITEHNICA of Bucharest, Romania, e-mail: iacob55@yahoo.com

advantage for the development and the diagnosis of the applications with parallel support. This paper presents solutions for the scientific study made on the image processing domain using parallel architecture Cell BE. IH (information hiding) is the process of hiding a secret message (stego-data) on an ordinary item of communication (cover data). A steganographic system uses two entities for communication: an encoder for coding the secret message and a decoder for the extraction of it. Steganography is the "secret communication between two entities" and it is part of the information hiding. Of course, steganography, the art and the science of writing secret messages so that nobody, apart from the sender and the receiver, suspects the existence of such a message(security through obscurity) [Fabien A. P. Petitcolas, Ross J. Anderson and Markus G. Kuhn, 1999], is included in the study. A new model of steganographic system is proposed. Steganography is a new but good and complex domain for researchers. This domain yet offers a lot of research subjects. One research subject is about the development of new steganographic systems with a low sensibility to the steganographic attacks against these secret messages such as filters, rotations, etc. Another part of the information hiding is the steganalysis, the capacity of analyzing an image and finding a secret message. Steganography is used in the fight with the plagiarists by hiding a copyright message in papers, so, it can be a good tool for proving the authenticity and originality. International community made a benchmark for steganographic systems. This benchmark is available on the Internet community site[1]. In this domain of hiding information, the most important request is a high level of imperceptibility followed by a high level of robustness. The capacity of the steganographic system is also an important parameter. Imperceptibility includes the minimization of the visual effect (distortion) and the detectible level of the stego-data. Kawaguchi and Eason proposed a steganographic system in [2]. The image is divided in bit maps and then a transformation from RGB domain to the canonical gray coding domain is made. Kawaguchi and Eason obtained good results but they did not develop a study regarding the level of the imperceptibility of their technique. Also, the algorithm proposed by Seppanen, Makela, Keskinarkaus (SMK) [3] was the start point for different studies [4]. This paper proposed a new steganographic system using three shared information: an image hash for the protection against steganographic attacks, a pixels palette (this is the result of applying the K-means clustering algorithm, with some modifications) used for encoding the message and the dimension of one piece of the image. Parallel architecture Cell BE is used for the calculation of the hash and the pixels palette.

2. Cell BE

2.1. Architectural description

Cell BE architecture consists of a chip with nine processors. The connections to peripherals are made through a high bandwidth bus. In Fig. 1. it is shown the block diagram of a Cell Broadband Engine architecture.

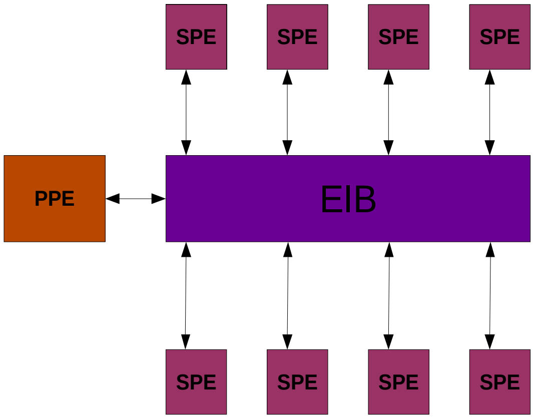


Fig. 1. The block diagram of the Cell BE architecture

There are several important elements:

1) PowerPC Processor Element (PPE)

PPE is a RISC core with 64 bit PowerPC architecture and has a traditional subsystem of virtual memory. It is the main processor and it is responsible for the management of the resources and operating system.

2) Synergistic Processor Elements (SPEs)

There are eight SPEs processors. These processors are optimized for SIMD operations and have an identical RISC architecture with the local data and instructions memory (local store). SPEs also have a set of 128x128 bit registers and a capacity of 256 KB.

SPEs use DMA transfers for moving data and instructions between the main store and the local store.

3) Element Interconnect Bus (EIB)

PPE processor and SPEs can communicate through a high bandwidth bus named EIB (Element Interconnect Bus).This bus has a structure based on four rings for transferring data and a tree structure for commands.

2.2. Working medium

In this working medium, PPE is responsible for the thread allocation and for the resources management between SPEs. The Linux Kernel on the SPEs controls SPE programs. SPE threads model is M:N and this means that M threads are mapped on N processing elements. In a Linux operating system, the main thread of a program is working on PPE and it can create one or many Linux tasks for Cell Broadband Engine. A Linux task for Cell BE has one or many Linux thread(s) that can work on PPE or SPE.A SPE thread is a Linux thread working on SPE. SIMD operations support is present on the Cell BE architecture.

On PPE, this support is represented by the Multimedia Extended Vector instruction set and on SPE by the SPU instruction set.

Communication between PPE-SPE: DMA, mailboxes and events

PPE can communicate with SPE trough the MMIO(memory mapped IO) registers. These registers are accessed by the associated SPE through own channels.

(unidirectional registers and queues). There are three important communication mechanisms:

a) mailboxes

b) events (notification registers by signals)

c) DMA

a) Mailboxes

There are queues that allow the sending of short messages, 32 bit messages. Every mailbox has an own channel associated and a MMIO register. Mailboxes can also be used for synchronization between SPEs.

b) Outbound Interrupt Mailbox - Events

Using Outbound Interrupt Mailbox is for avoiding busy-waiting on PPE and for pointing out when a message arrives from one of the SPEs. The solution for avoiding busy-waiting is the use of the events.

c) DMA

Memory Flow Controller consists of a controller for DMA transfers. DMA is used for moving a high volume of data between PPU and SPU.

Double buffering mechanism**,** which consists of multiple transfers by using a buffer is also used. In the same time, new data is brought on SPE and old data is processed. So, the processing time on SPE is substantially reduced.

3. Steganographic system

3.1. System description

The steganographic system (Fig. 2) has two elements : an encoder and a decoder.

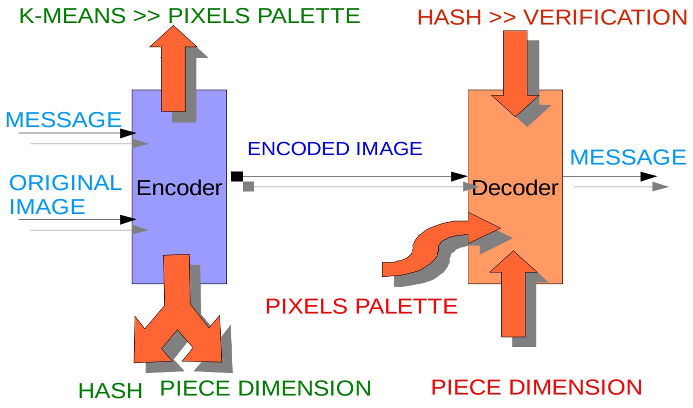


Fig. 2. The block diagram with the steganographic system

The encoder has the main function to hide the message on the image using a LSB algorithm. The pixels palette used for hiding the message is the result of applying the K-means clustering algorithm, with some modifications. This algorithm chooses pixels by calculating the clustroids (these are pixels that have the least distance to the other pixels on a cluster).This algorithm of choosing the pixels offers an uniform palette for encoding. The tests were made for k=2 and the modification of the clustering algorithm consists of the use of a threshold for deciding, at every step, what is the new clustroid. The threshold chosen for tests is 200.For a good evaluation of the distance between two pixels, it is proposed the following formula:

$$D(P,Q)=\sqrt{(P_r-Q_r)^2+(P_g-Q_g)^2+(P_b-Q_b)^2} \qquad (1)$$

The clustering algorithm (with the modification proposed) for finding the pixels palette is (for a piece of the image):

```
0. Input: A piece of the image with quadratic dimension
1. Initialization with two clustroids (big distance between them - clustroid_1,clustroid_2)
2. for every pixel
        calculate the distance between clustroids and the current pixel
        keep the minimum distance
```

```
            if the minimum distance > the threshold chosen then
            clustroid_1=the current pixel
            else clustroid_2=the current pixel
    3. Apply step 2 for the last  two clustroids
    4. Return the final clustroid
```

The encoding algorithm uses this palette. Only the Red component from every pixel is used for encoding. The last five bits are used from every Red component of a pixel, so the distortion is acceptable.

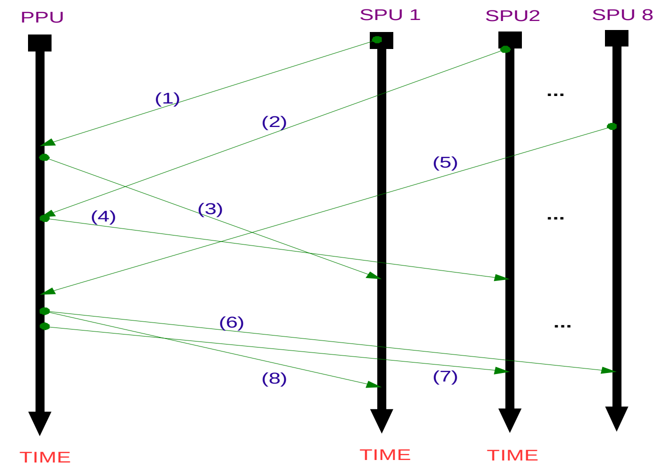
A protection item is used for the robustness of the system . It is about a hash of the encoded image. This hash is defined as the average of the averages of the pixels components for every piece of the image. This hash is useful for pointing out if the encoded image was analyzed by a person and attacked with the purpose of destroying the secret message. The encoded image is sent to the decoder not cryptic and because of that, someone can analyze the image and destroy the message.

The encoder returns three elements: the encoded image, a hash for protection and a pixels palette. These three elements must be shared with the decoder and they must be sent to the decoder through a secured way. The hash and the palette are calculated on Cell BE architecture because there is a high volume of data .The decoder has the main function to find the secret message using three shared elements: the hash, the dimension of one piece of the image and the palette. The decoder needs a high processing power only for verifying, locally, the hash for data integrity. Then, using the pixels palette, it finds the message.

3.2. The mapping of the application on the parallel architecture CELL BE

A parallel architecture is necessary for calculating the hash or the palette, because there is a high volume of data. The mapping of the application on Cell BE is shown in Fig. 3.

The image is initially on the PPU which waits for requests from the SPU. When a request from the SPU arrives, the PPU treats it and sends to the SPU, by mailboxes, an address, address that is used by the SPU to transfer data by DMA from main store to local store. When SPU finishes processing data, it sends this data by DMA to the PPU.



LEGEND : (1), (2) SPU REQUEST ; (3), (4) PPU OFFERS DATA TO SPU; (5) SPU SENDS PROCESSED DATA TO PPU; (6), (7), (8) PPU GIVES FINISH SIGNALS TO SPU

Fig. 3. The time diagram for mapping the application on the Cell BE

It is a cyclic mechanism until the image is processed. When the image is processed, the PPU sends to the SPU the address 0. This means that the SPU can terminate its execution. Finally, the PPU processes data and finishes its execution. Mailboxes, events and DMA transfers are used for communication.

4. The analysis of some performance parameters, contributions and conclusions

4.1. Experimental framework

The dimensions for the piece of the image were 16, 32, 64. The parameter K for K-means algorithm was chosen 2 and the threshold for the clustering algorithm was 200.The images for tests was 256x256, 512x512 and 1024x768.All experiments have a PSNR coefficient up to the 40 dB, so that at this value, there is no visual effect on the encoded image.

4.2. System capacity

The system capacity is measured in bpp (bit per pixel) and varies depending on the number of bits used for encoding and the dimension of the piece of an image.

For a piece of the image, only one pixel is used for encoding (this is the result of the clustering algorithm). The capacity increases with the decreasing of the dimension of the piece of the image (Tables 1 and 2). It also increases with the increasing the number of bits used for encoding from every component of a pixel. It can increase by considering for encoding also the Green and Blue components.

Table 1

The capacity depends on the dimension of a piece of the image (for an imagine of 256x256)

Image Dim.	Piece Dim.	Capacity (bits)	
256x256	8	5*32*32	
	16	5*16*16	
	32	5*8*8	
	64	5*4*4	

Table 2

The capacity depends on the dimension of a piece of the image (for an imagine of 512x512)

Image Dim.	Piece Dim.	Capacity (bits)	
512x512	8	5*64*64	
	16	5*32*32	
	32	5*16*16	
	64	5*8*8	

These are illustrated in Fig. 4.

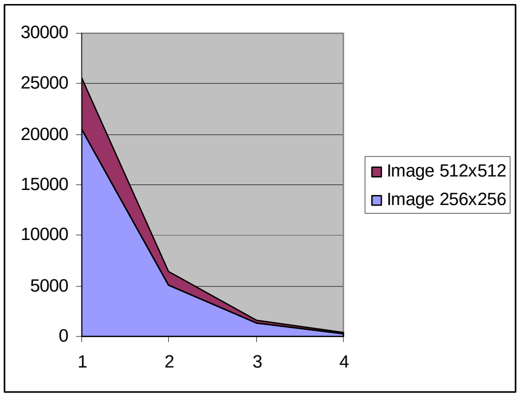


Fig. 4. The capacity depends on the dimension of one piece of the image

4.3. Distortion

The imperceptibility level includes the minimization of the distortion and the detectable level. The distortion is measured in this article by the formula:

$$PSNR = 20 \cdot \log\left(\frac{255}{\sqrt{MSE}}\right) \qquad (2)$$

where MSE is the mean squared error for every pixel in the encoded and original image. For a 40 dB level of this parameter, there is no visual effect between the original image and the encoded image.

4.4. Robustness and data integrity

One solution for detecting if the encoded image was analyzed and modified by someone is proposed in this paper. It is about a hash for protection which gives to the system a high level of robustness. The purpose of someone who analyzes the encoded image is to destroy the secret message if he does not decode it. Verifying the hash on the decoder gives a high level of data integrity.

4.5. Example

One of the example tested uses the following image (Fig. 5) in which is hidden the secret message: “Acesta.este un.mesaj.ascuns”.
There are no visual differences between the encoded (Fig. 6) and original image, but the first one contains the message.

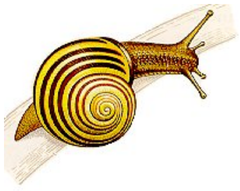


Fig. 5. The encoded image

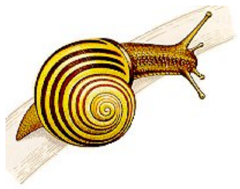


Fig. 6. The original image

4.6. Contributions, future work and conclusions

The paper represents a critical study on steganographic domain and a **new staganographic system** is proposed. An analysis of the most important parameters such as **capacity, distortion, robustness, data integrity** is made. The paper also includes an original and critical analysis of the states-of-the-art in the image processing domain on parallel architectures. A method for the **maintenance of a high level of the data integrity** by calculating the encoded image hash is proposed. For a **dispersed secret message** in the image, it is proposed an algorithm for finding the pixels **palette** used for encoding .The finding of this palette is based on the k-means clustering algorithm with some modifications. Using a parallel architecture (in order to have processing power for a high volume of data) for the development of such a system is **another new element**. The system presents a small disadvantage: three elements must be shared between the encoder and the decoder. **Another contribution consists of designing and implementation** of a new steganographic system on the parallel architecture Cell BE from IBM. New concepts are also used: **double buffering, mailboxes, events, DMA transfers, SIMD instructions.** The paper offers **an efficient mapping of the application** on the architecture Cell BE, so that the resources and the programming facilities are used optimally. **Future work** is referring to the evaluation of other parameters and other encoding schemes.

REFERENCES

[1]. Steganography international community, http://www.petitcolas.net/fabien/watermarking/stirmark/

[2]. *E. Kawaguchi, R.O. Eason,* Principle and applications of BPCS-steganography, Int Symp. On Voice, Video and Data Communications, SPIE, 1998

[3]. *T. Seppnnen, K. Makela, A. Keskinarkaus,* Hiding information in color images using small color palettes. Proc. Information Security, Third Int. Workshop, ISW, December 2000, pp 69-81

[4]. *G. Brisbane, R. Safavi-Naini, P. Ogunbona,* High-capacity steganography using a shared color palette In: IEE Proceedings on Vision, Image and Signal Processing, 9 December 2005, 152(6), 787-792. Copyright IEEE 2005

[5]. *C. Cachin,*:An information-theoretic model for steganography, Inf. Hiding, 1998, pp.306-318

[6]. *B. Girod,* What's wrong with mean-squared error?, *Vis*. Factors Electron. Image Commun., 1993, pp. 207-220

[7]. Steganography Analysis and Research Center, http://www.sarc-wv.com/

[8]. Cell Broadband Engine http://www-01.ibm.com/chips/techlib/techlib.nsf/products/Cell_Broadband_Engine

[9]. Arhitectura sistemelor de calcul, Note de curs (Calculus systems architecture, course notes), Computer Science Department, University POLITEHNICA of Bucharest, https://anaconda.cs.pub.ro/asc/index.php/Laboratoare/

[10]. Workshop Cell BE, IBM, 2007.