Preface

—...δεῖν γὰρ ἐπ′ ἐκείνων καὶ αὐτὸν παρακελεύεσθαι τὸν Πυθαγόραν ζητεῖν ἐξ ἐλαχίστων καὶ ἁπλουστάτων ὑποθέσεων δεικνύναι τὰ ζητούμενα...

Pythagorians

Since the day that Johann Gutenberg invented the modern typography in Germany in 1450 we are inundated with reams of paper documents. Although the manufacturing of paper was known in China since the second century and high-quality paper was already available in the fifth century, the availability of printed material began to increase significantly only after the invention of Gutenberg, who introduced the metallic elements in typography.

In less than 50 years after the printing of the first book, more than 1,700 printing machines were used in Europe, through which 40,000 different print jobs had already been completed (Johnson 1973). By 1975, about 400 years later, it was estimated that approximately 50,000,000 printing works had been completed, while according to the 1998 estimate there are over 10^{12} documents in the world with a provision for doubling every 3 years (Open Archive Systems 1998). During the beginning of the twenty-first century the trend was still mainly oriented towards the production and reproduction of printed documents at a rate faster than the rate of digitization.

Nowadays, the main issues regarding the printed documents are not related with the processes of paper or document production, but rather with the huge physical storage requirements. On the other hand, the storage requirements for digital images are significantly lower. Indicatively, in a 4.7 GB DVD one may store about 4,320 uncompressed A4 pages digitized at a resolution of 300 dpi in binary form (1 Bits Per Pixel (bpp)—black and white), or about 540 gray-level images (8 bpp), or about 180 full color (24 bpp). This DVD including the protective sleeve requires about 160 cubic centimeters or 1.6×10^{-4} m^3 of physical storage space. A simple comparison to the respective paper volume for the same amount of material (using a typical 80 gr. paper), which gives an estimate of about 432×10^{-4} m^3 for the case of the 4,320 pages, shows that the digitized material requires 270 times less storage volume.

In order to cope with the storage issue that arises from the huge amount of produced digital documents by improving the data compression ratio, numerous compression methods for digitized documents have been studied and developed over the recent years, either case-specific or generic. It is noticeable that by applying an efficient method of digital data compression, a volume compression for 2,000,000 books of 200 pages to less than 0.2 m^3 can be achieved.

Even though these numbers are impressive, this is not only a storage issue. With the advent of the new communication technologies through the pervasive presence of the Internet and the mobile networks, this issue is directly reflected on the management and transmission of the digitized information that a color document represents. Considering the data archival, document digitization contributes significantly to more efficient data management, since search and retrieval can be faster and more intuitive (using semantics). The ability to efficiently transmit digitized documents through wired or wireless networks could boost a fast building trend towards tele-working, tele-cooperation and true cybernetics (digital governance) and civil services. These highlight one of the main issues in information transmission, transmission speed. The provided table shows indicative transmission times for an ISO A4 page digitized at 300 dpi (which is about 25 MB uncompressed) through LAN, DSL, ISDN and PSTN technologies, reported in bpsec and Bps.

Channel bandwidth (approximate true speed)	**Approximate transmission time**
PSTN 28.8 Kbps (0.0035 MBps)	2 h
PSTN 33.6 Kbps (0.0041 MBps)	1.5 h
PSTN 56 Kbps (0.0068 MBps)	1 h
ISDN 128 Kbps (0.0156 MBps)	25 min
DSL 384 Kbps (0.0469 MBps)	10 min
Dedicated Line 2 Mbps (0.25 MBps)	1.5 min
LAN 10 Mbps (1.25 MBps)	20 s
LAN 100 Mbps (12.5 MBps)	2 s

Apparently, in all cases except the fast LAN networks (or the modern high bandwidth WiFi and mobile networks), transmission times can be rather disappointing and, in some cases, impractical. A straightforward way to tackle with this issue is to reduce the amount of data being transferred using a data compression mechanism. The application of compression is not only an economic solution—since it is an algorithmic response to the problem and, in most cases, does not demand any infrastructural changes—it is also the only meaningful choice towards the quantification of the information within a transmitted message and a distinction between the useful part of the information and the inherent redundancy (since, in most cases, data compression exploits statistical characteristics of the data to reduce the redundancy). Otherwise, the only alternative would be to change all network infrastructures to meet the constantly growing bandwidth demands, which would eventually bring higher investments and, consequently, higher costs for the consumers. Considering also the fact that digital image compression has already a long history of success and can be considered a mature technological approach to solve problems of the modern society, its choice as a solution to this problem seems most appealing. Indicatively, the application of JPEG image compression (which is not the most appropriate method of MRC compression, though) on a typical ISO A4 page digitized at 150 dpi and 24 bpp (true color) can lead to a reduction from 6 MB to 200 KB attaining a 1:30 compression ratio with unnoticeable quality degradation (30.7 dB PSNR) and a data transmission time reduction through a 56 Kbps channel from 15 min to 30 s.

A digitized document is a digital image that represents an 'exact' replica of an original document (Lynn 1990). Digital images can be considered 'superior' to printed documents for a number of reasons, including that they can be efficiently stored, efficiently and nonlinearly searched and retrieved, reproduced and transmitted with accuracy or controlled distortion.

The treatment of digitized documents as generic digital images is not efficient for data compression. This is due to the fact that the digitized documents include mixed visual content with largely differentiating characteristics, including parts with text, parts with graphics and parts with natural images in shades of gray or various colors. This is why they are usually referenced as mixed raster content or MRC (ITUT 1999). Each of the individual parts of such mixed images exhibits different statistical and semantic characteristics and need to be treated differently to achieve efficient data compression. For decades these image characteristics have been the center of attention in the field of image compression, and specific methods have been discovered (and continue being developed) to treat each of the cases. In recent years there has also been an attempt to identify these different parts in mixed images and apply the appropriate data compression method to each identified part. MRC compression specialists have to deal with notions such as segmentation and layered representation, sparse matrix compression, error resilience during transmission through noisy and congested networks, reconstruction of lost data, etc. All these notions represent the central considerations in this treatise, so methods to treat them are presented in detail, based on extended research over a number of years (Pavlidis and Chamzas 2005; Pavlidis et al. 2001, 2002, 2003, 2004, 2005; Politou et al. 2004), which departed significantly from other approaches basically regarding the overall scope.

Most of the works in the field of MRC processing and compression start by adopting some significant assumptions for the type of the characteristics in the images that are to be processed. In most cases, the compression of MRC refers to binary images (1 bpp) as they usually represent black text characters on white background. In addition, in many cases segmentation of MRC is studied without considering compression or even the efficiency in transmission or error correction in noisy environments. These considerations are detailed by a study of segmentation of color MRC images in order to decorrelate their basic structural elements into at least three layers for their independent compression, while keeping a 'loose' compatibility with the ITU Recommendations T.44 (ITUT 1999) for MRC.

This is complemented by a study on compression methods, which by exploiting segmentation operate in an optimal way in the different layers of information targeting an improvement of the total compression ratio and perceived quality. Another approach that completes the picture is presented as a study and simulation of efficient transmission of compressed MRC images through noisy and congested channels, in order to analyze and propose reconstruction mechanisms in an integrated management framework.

The presentation of the main concepts in this treatise is divided into two main parts, the first of which includes introductory chapters to cover the scientific and technical background aspects, whereas in the second part there is a set of research and development approaches in MRC proposed by the author to tackle the issues in MRC segmentation, compression and transmission from a novel point of view. The first chapter reviews *the color theory* and *the mechanism of color vision* in humans; the scope is to form a stable ground for a color representation theory and to identify the main characteristics and limitations of visual perception. The *information theory and data coding* is introduced in the second chapter, and a brief presentation of the most widely used compression methods is included to set the background for data compression and to highlight the complexity involved in dealing with MRC. In this chapter, a contribution of the author is included that involves the study of the new image compression standard, JPEG2000, which led to the development of a novel system for progressive transmission of color images in cultural multimedia databases. The third chapter introduces the *segmentation of images* through an extensive literature review, which highlights the differences regarding the approaches to tackle MRC segmentation. Different approaches are being presented in order to emphasize that segmentation of images is a case-specific (or context-specific) mechanism and there is not one-for-all solution available. In the second part, and specifically in the fourth chapter, the *segmentation of color images for optimized compression* is introduced and analyzed, including a multi-layered decomposition and representation of MRC and the processes that may be adopted to optimize the coding rates of those different layers. A nonlinear projections method is presented, in which the background layer in the multi-layered representation is optimized, to tackle one of the major issues in MRC image compression, that is, the data filling. In the final chapter the *segmentation of color images for optimized transmission* is introduced and analyzed, including two distinctive approaches that are based on segmenting the image data into significant and complementary; the first, a global approach based on the application of multi-layered coding and progressive transmission using typical coding standards and labeling of data as significant and additional, in addition to applying a segmentation method based on the clustering of data from a differential image; the second and latter, the exploitation of the JPEG2000 standard for color image transmission using a selective data segmentation into significant and complementary parts, which provides the functionality to apply cost policies.

Xanthi, Greece George Pavlidis

http://www.springer.com/978-981-10-2829-8

Mixed Raster Content
Segmentation, Compression, Transmission
Pavlidis, G.
2017, XXVIII, 354 p. 260 illus., 103 illus. in color.,
Hardcover
ISBN: 978-981-10-2829-8