B. Justrell, Digital imaging in archives, Arh. vjesn.,
god. 41 (1998), str. 23–28

Börje Justrell

Državni arhiv Švedske

DIGITAL IMAGING IN ARCHIVES[1]

UDK 681.3:655.2
930.25:681.3

Stručni članak

Na početku članka autor definira digitalne slike kao "elektroničke fotografije", koje se sastoje od mreže točaka od kojih svaka nosi i podatak o boji, a spremljene su u digitalnom binarnom kodu (u obliku nula i jedinica). Digitalne slike, za razliku od tekstovnih datoteka, nemoguće je pretraživati.

Prednosti digitalnih slika u odnosu na analogne su sljedeće: kvaliteta digitalnih slika se ne mijenja s vremenom i ne umanjuje se upotrebom. One se mogu vrlo lako premještati i umnažati u neograničenom broju primjeraka. Također se mogu prenositi putem računalnih mreža. Trenutačno postoje i veći nedostaci digitalizacije: nužno je redovito kopiranje slika, a važno je i u kakvim se uvjetima čuvaju. Problem predstavlja i zaštita autorskih prava i nepostojanje standardnih formata za čuvanje.

Za upotrebu u arhivima često se predlaže "hibridni sustav", koji uključuje mikrofilmiranje radi zaštite i digitalizaciju za korištenje.

Bitne komponente programa digitalizacije slika su konverzije, upravljanje zbirkom, predstavljanje sadržaja zbirke te održavanje i osiguravanje dugoročnog pristupa.

Za kvalitetu digitalizirane slike ključan je sam postupak digitalizacije. Najvažniji su elementi: razlučivost, broj bita potrebnih za spremanje pojedine točke, poboljšanje slika, sažimanje, korištena oprema i njene karakteristike, prosudba i skrb operatera.

Razlučivost (rezolucija) je gustoća točaka od kojih se slika sastoji. Nužno je odrediti koja je razlučivost dostatna da se obuhvate svi važni detalji u zapisu koji digitaliziramo. Preporučljivo je odrediti razlučivost za cijelu grupu zapisa.

---

1 Izlaganje s godišnjeg savjetovanja Hrvatskog arhivističkog društva i Odbora za informatičku tehnologiju MAV-a, održanog u Bizovcu u listopadu 1997. s temom Elektronički zapisi i informatizacija arhiva (*Electronic Records and Archival Automation*).

B. Justrell, Digital imaging in archives, Arh. vjesn.,
god. 41 (1998), str. 23–28

Brojem bita koji se koriste za spremanje pojedine točke određuje se ukupni broj raspoloživih boja. Tri osnovne skupine su: crno-bijeli prikazi kod kojih se za jednu točku koristi jedan bit; razne nijanse sive boje, za što se koristi 8 bita za jednu točku; korištenje većeg broja bita za definiranje raznih boja (korištenje 24 bita omogućuje prikazivanje 16,7 milijuna boja). Preporučuje se da se pri određivanju ovog elementa uzmu u obir svojstva originala: za crno-bijele zapise koristit ćemo 1 bit za točku, a boju samo kod dokumenata kod kojih boja daje neku informaciju.

Postupci poboljšanja slika mogu povećati kvalitetu, ali se postavljaju problemi vjernosti i autentičnosti. Zbog toga se poboljšanja treba koristiti oprezno i dokumentirati sve takve postupke.

Kompresija se koristi za smanjivanje veličine datoteke. Tehnike kompresije mogu utjecati na kvalitetu slike. Preporuča se korištenje onih tehnika kod kojih se ne gube podaci.

Oprema za digitalizaciju i njena svojstva u velikoj mjeri utječu na kvalitetu slike. Razna oprema može dati razne rezultate. Zbog toga treba oprezno pristupiti izboru opreme.

Golem je utjecaj prosudbe i skrbi operatera na kvalitetu slike. Zbog toga treba stalnim programom kvalitete provjeravati rezultate procesa digitalizacije.

Sažetak izradio Tomislav Ćepulić

Today, digital imaging has grown to one of the most expanding field in the computer world. Digital imaging technology has been available for some time, but it was not until 1990s that it became widespread when technical advances had led to improved image capture, lower costs and greater accessability. This technology holds great promises for the society and therefore also for archival management. However, there are drawbacks. The major problem with digital technology from a preservation perspective is the obsolescence caused by a rapidly changing technology and the costs associated with long term preservation of digital files.

What are digital images?

A digital image is an "electronic photo" taken of a scene or scanned from an original source material, which could be a document, a microfilm or an analogue photo.

A digital image is sampled and mapped as a grid of dots or picture elements, so called "pixels". Each pixel is given a tonal value (black, white, shades of grey or colour) and digitally represented in binary code (zeros and/or ones).

Digital images are also known as "raster" or "bit mapped" images, and they are different from other types of electronic files, for example vector files, in which graphic information is represented as mathematically defined lines and curves.

Digital images are, unlike alphanumeric text files, "dump" files because the information in them cannot be searched.

Advantages and disadvantages of digital imaging

Digital images offer many advantages over analogue counterparts like photocopy and microfilm in terms of capture, duplication, storage and transmission. The technology has the potential to create a higher quality reproduction of a deteriorating original then conventional copying processes. This technology is also very flexible. A page may be inserted or replaced with ease, and a digital image can, unlike microfilm or photocopies, be reproduced over and over again and all subsequent copies of that digital image will retain the same quality. Unlike microfilm and paper, digital images do not decay with use, and they offer several alternatives in output – paper, microfilm, digital files – for distribution to users. Digital images can also be transferred over networks for remote, simultaneous and multiple access. But – as mentioned above – there are a number of drawbacks that, at least currently, limit the utility of digital imaging in archival management. Besides the problems of obsolescence and that special recopying, storage and migration requirements must be considered when planning for a digital imaging system, there are legal constraints, including copyright, lack of standards and wide variations in quality and capacity of hardware and software. Like other recently introduced technologies digital imaging also suffers from a frustrating lack of vendor support and stability.

All these disadvantages can – or will be possible in the future – to overcome. However, many micrographic and imaging experts are arguing for a so called "hybrid system" approach which means using the best out of two technologies: microfilm for preservation and digital images for access. Today a growing number of archival institutions are initiating digital imaging programmes: National Archives of Canada are digitizing its collection of photo negatives on nitrate film, Archivo de Indias in Seville in Spain has up until today digitized more than 11 million pages from records of the Spanish colonisation of America, State Library in New South Wales in Australia has digitized about 8 000 pages of manuscripts from the Joseph Banks collection and they are now available on the Internet – just to mention some archival initiatives. The important thing with these programmes is that they have to be carefully defined to guarantee success and minimize failure. Before starting such a programme realistic expectations have to be set up, built on an understanding of the attributes of the documents which are to be converted but also of the conditions for digital capture and of the whole range of uses the digital images will support.

The substance of a digital imaging programme

The key components of an imaging programme are:
- conversion
- collection management
- presentation
- maintaining and long term access.

All these components are equally important – the chain is not stronger then its weakest part – but in this paper I will focus on conversion.

The goal of any imaging programme should be to capture and present in new formats the significant informational content contained in a collection of documents. For that reason the quality assessments of the digital images have to be based on a comparison between those digital images and the original source documents, not on some vaguely defined concept of what is good enough to serve immediate needs. What is sufficient for today's purposes will probably be inadequate tomorrow.

This will raise the question of image quality. My point is that the problem is not to capture an image at the highest quality possible, but to match the conversion process to the informational content of the original – no more, no less.

What governs digital image quality at capture? Basically it is
- resolution
- bit depth
- image enhancement
- compression
- equipment used and its performance
- operator judgement and care.

I will look closer to each of these factors.

Resolution is determined by the number of pixels used to present the image, expressed in dots per inch (dpi) or pixels per inch (ppi). Increasing the number of pixels used, will result in a higher resolution and a greater ability to record fine details, but just continuing to increase resolution will not result in better quality, only in a larger file size. The key is to determine the point at which sufficient resolution has been used to capture all significant details in the source document.

The recommendation is to choose a resolution that is sufficient to capture the finest significant details in the group of documents that are intended to be scanned.

Bit depth is determined by the number of bits used to define each pixel. The greater bit depth used, the greater number of grey and colour tones can be represented. There are three kinds of scanning techniques:
- bitonal scanning using one bit per pixel to represent black or white

– greyscale scanning using multiple bits per pixel to represent shades of grey; the preferred level of grey scale is 8 bits per pixel, and at this level the image displayed select from 256 different levels of grey which is the maximum of average human perception

– colour scanning using multiple bits per pixel to represent colour; 24 bits per pixel is called true colour level, and it makes a possible selection among 16.7 million colours.

The recommendation is to choose a bit depth that is in accordance with the characteristics of the source documents: bitonal scanning for document consisting of black ink on white paper, greyscale (8 bits) scanning for documents containing significant grayscale information and colour scanning for documents containing colour information.

Image enhancement processes can be used to improve image capture, but the use of them raises concerns about fidelity and authenticity. Typical enhancement features include filters, tonal reproduction curves and colour management tools. The recommendation is to apply enhancement processes cautiously and document all processes that have been used.

Compression is normally used to reduce file size for processing, storage and transmission of digital images. The quality of an image can be affected by the compression techniques used and applied level of compression. Compression techniques can be either "lossless" when no information is thrown away in reducing the file size, or "lossy" when the least significant information is averaged or discarded in this process.

The recommendation is to use loss less, standard compression techniques which today mean ITU Group 4 or JBIG compression for 1-bit images and lossless JPEG or LZW for multi-bit images.

The equipment used and its performance has an important impact on image quality. Different equipment can perform differently even if they offer the same technical capability. The recommendation is therefore that manufacturers'claims of system capabilities should be investigated carefully and confirmed through sampling and references.

Operator judgement and care always have a tremendous impact on image quality. In the end decisions taken by humans decides which quality will be achieved. A continuos quality programme is recommended to verify consistency of output from the scanning process.

Conclusion

I have in this paper pointed out some fundamental aspects on digital imaging. Archives are initiating digital imaging programmes to meet urgent needs in archival

management, but doing this just by trying out the technology is a dangerous way. Before starting a digital imaging programme the needs of the users of digital images must be clearly defined and the attributes of the documents to be converted closely examined. This will, together with an appropriate technical infrastructure, ensure the utility of the programme.