

Image Representation

VALERIE GOUET-BRUNET
CNAM Paris, Paris, France

Definition

In computer science, the representation of an image can take many forms. Mostly it refers to the way that the conveyed information, such as color, is coded digitally and how the image is stored, i.e. how an image file is structured. Several open or patented standards were proposed to create, manipulate, store and exchange digital images. They describe the format of image files, the algorithms of image encoding such as compression as well as the format of additional information often called metadata. The visual content of the image can also take part in its representation. The more recent concept has provided new approaches to representation and new standards, gathered together into the discipline named *content-based image indexing*.

Historical Background

The first use of digital images began in the early 1920s with the technological development of facsimile transmission, and in particular with the Bartlane cable transmission system that was the first system that translated pictures into a digital code for efficient picture transmission. Later, in the 1960s and 1970s, the advent of digital image technology was closely tied to the development of government programs for space exploration and espionage and also to medical research with the invention of computerized axial tomography. With the availability of the CCD image sensors (charge-coupled device), the private sector also began to make significant contributions to the development of digital cameras.

In the mid-1980s, image format TIFF was created by the company Aldus with the aim of agreeing on a common file format for bitmapped images issued from scanners. In parallel, researchers at Xerox PARC had developed the first laser printer and had recognized the need for a standard means of defining page images. After several fruitless attempts, Adobe Systems proposed the PostScript language in 1982, quickly adapted for driving laser printers. Since then, other file formats dedicated to digital images were also proposed, such as GIF (1987), JPEG (1992), PDF (1993), PNG (1995) and SVG (1998). Today, a great effort is made to propose standard formats able to preserve data integrity for archiving purposes, to migrate easily to future technologies as well as to provide efficient compression for access and dissemination.

In 2000, the standard JPEG 2000 was proposed. Moreover, ambitious programs, MPEG-7 and MPEG-21, started in the late 1990s, are focusing on the harmonization of methods for representing, storing, sharing and accessing multimedia contents (text, audio, image and video) in a unified framework.

Foundations

Basics of Image Representation

Digital images can be classified into two main categories: vector graphics and bitmapped images (also called raster images). *Vector images* are geometrical 2D objects created with drawing software or CAD (computer-aided design) systems. They are represented by geometrical primitives such as points, lines, curves, and shapes or polygons, which are all based upon mathematical equations. Unlike bitmaps that are resolution-dependent, vector images are scalable, which means that the scale at which they are shown will not affect their appearance. Such images are dedicated to the representation of images with simple content, such as diagrams, icons or logos.

A *bitmapped image* is composed of a set of dots or squares, called pixels (for picture elements), arranged in a matrix of columns and rows. Each pixel has a specific color or shade of gray, and in combination with neighboring pixels it creates the illusion of a continuous tone image. Unlike human vision, sensors that capture images are not limited to the visual band of the electromagnetic spectrum and digital images can cover almost the entire spectrum, ranging from gamma to radio waves.

Managing bitmapped images requires the choice and manipulation of several parameters such as: *Color model*. A color model is an abstract mathematical model describing the way colors can be represented as tuples of numbers. From this model, the combination of dedicated primary colors provides all the colors possible that are embedded in the corresponding color space. RGB, CMYK, CIELAB and CIELUV are the most known color models and spaces.

Dynamic range. The dynamic range of a digital image (also called color depth) determines the maximum range of gray level or color values carried by each pixel. The number of bits used to represent each pixel determines how many colors can appear in the image.

Photographic-quality images are usually associated with 24-bit dynamic range, such as in the JPEG format. *Resolution*. Resolution expresses the density of elements, pixels for instance, within a specific area. This term does not have any sense when dealing with digital images as files, but it applies when associating a digital image with a physical support, such as display on a screen, printing on a printer or capture with a

scanner. Resolution is classically represented in terms of dpi (dots per inch) unit, which was originally the unit adopted for printing.

Appearance of bitmapped images, which are made up of a fixed grid of pixels, clearly depends on the resolution chosen, unlike vector images that are scalable and then have the same appearance whatever the dimensions chosen for visualization.

Image Compression

Image compression is the process of shrinking the size of digital image files. Compression algorithms are especially characterized by two factors: compression ratio and generational integrity. *Compression ratio* is the ratio of compressed image size to uncompressed size and *generational integrity* refers to the ability for a compression scheme to prevent or mitigate loss of data, and therefore image quality, through multiple cycles of compression and decompression. Lossless compression ensures that the image data is retained, as with Run-length encoding, Huffman coding and LZW coding. On the other hand, lossy compression schemes involve intentionally sacrificing the quality of stored images by selectively discarding pieces of data.

Run-Length Encoding (RLE) is probably the simplest form of lossless data compression: sequences in which the same data value occurs in many consecutive data elements are stored as a single data value and count. Image data is normally run-length encoded in a sequential process that treats the image data as a 1D stream, line by line, column by column or diagonally in a zigzag fashion. Common digital image formats for run-length encoded data include TGA, PCX. It is possible with BMP, TIFF and JPEG.

Huffman Coding, created in 1951 by David A. Huffman, is an entropy encoding algorithm used for lossless data compression. The basic idea of this algorithm is to code with few digits the most common input symbols of a document. Each symbol is encoded by using a variable-length code table, where the codes are defined according to the estimated probability of occurrence for each possible symbol. Today, Huffman coding is often used during the final process of some other compression methods such as JPEG and MP3.

LZW Coding Lempel-Ziv-Welch (LZW) is a lossless data compression algorithm created by Abraham Lempel, Jacob Ziv, and Terry Welch and published in 1984. The compression algorithm builds a string translation table that is based on fixed-length codes (usually 12-bit). As the system character serially examines the document if the string read is not stored in the table, a new code is created in the table and associated with this string. Otherwise, the current string is encoded with an existing code. This algorithm became very widely used after it became part of the GIF image format in 1987. In contrast to other

compression techniques such as JPEG, it allows preserving very sharp edges, suitable for line art images often stored in GIF format.

JPEG Compression JPEG is the most common image format used for compressing and storing by digital cameras and other photographic image capture devices. “JPEG” stands for Joint Photographic Experts Group, the name of the committee that created the standard, which was approved in 1994 as ISO 10918-1 standard. This algorithm stands on the representation of the image in the frequency domain by using a two-dimensional DCT (Discrete Cosine Transform) that describes the variability of the signal in terms of low-level and high-level frequencies. The human eye notices small differences in brightness over a relatively large area, but does not distinguish the exact strength of a high frequency brightness variation very well. Consequently, the amount of information in the high frequency components of the DCT can be neglected without drastically affecting perceptual image quality: the DCT components are divided by factors of a quantization matrix that increase with the spatial frequency, and then rounded to the nearest integer. This is the main lossy operation in the whole process. Typically, many of the higher frequency components are rounded to zero and the other components become small numbers, which take many fewer bits to store.

File Formats

There are a lot of image formats for vector graphics as well as for bitmapped images. Most of them are open standards and patent expired for the others.

Vector image formats contain a geometric description of the objects which can be rendered smoothly at any desired size. Among the most common formats, there are:

- SVG (Scalable Vector Graphics) is an open XML based standard created in 1998 and developed by the World Wide Web Consortium to address the need for a versatile, scriptable and all-purpose vector format for the web and otherwise.

- EPS (Encapsulated PostScript) is a standard file format created by Adobe Systems in the mid-1980s. It follows DSC (Document Structuring Conventions) rules that are a set of standards for PostScript.

File formats abound for *bitmapped images*, but many digital imaging projects have settled on the formula of TIFF, JPEG, GIF and also PNG files.

- TIFF, for Tagged Image File Format, is a file format for storing images such as photographs as well as graphics. It was originally created by the company Aldus, was then under the control of Adobe Systems and is now in the public domain. TIFF supports several lossless and lossy techniques of image compression, such as LZW, Huffman coding and JPEG. The ability to store image data in a lossless format makes TIFF

225 files a useful method for archiving images and preservation purposes.

- JPEG, for Joint Photographers Experts Group, is a file format that was developed specifically for high quality compression of photographic images in a 24-bit RGB color model. It is generally employed for online presentation and dissemination; the associated lossy algorithm for compression makes it inappropriate for archiving purposes. The file format associated is JFIF (JPEG File Interchange

235 Format, 1992), a public domain storage format for JPEG compressed images. Unlike TIFF, JFIF does not allow for the storage of associated metadata, a failing that has led to the development of SPIFF (Still Picture Interchange File Format), which is now the international standard.

- GIF, for Graphics Interchange Format, is an 8-bit image format for indexed colors that was introduced by CompuServe in 1987 and has since come into widespread usage for art images such as diagrams or logos with a limited numbers of colors.

245 In 1995 CompuServe proposed the PNG format (Portable Network Graphics) as a replacement for the GIF format without patent license. PNG offers a better and lossless compression technique called DEFLATE (that combines LZ77 with Huffman coding). Since 250 2003, it has been an international standard.

Today, the status of TIFF as the de facto standard format for archival digital image files is challenged by other formats such as PNG and JPEG 2000 that are able to preserve data integrity as well as to provide efficient compression ratios for access and dissemination.

Metadata Representation

260 Metadata are commonly defined as “data about data.” They constitute the documentation or a structured description associated with a document. Image files automatically include a certain amount of metadata that are stored in an area of the file defined by the file format and called *the header* but information may also be stored externally.

In the widely used TIFF format, the term “tagged” indicates that developers can define and apply dedicated tags to enable them to include their own proprietary information (called “private tags”) inside a TIFF file without causing problems of compatibility. More recently, Exif format (Exchangeable image file format) was created by the Japan Electronic Industries Development Association. The latest version was published in 2002 and while the specification is not currently maintained by any industry or standards organization, its use by camera manufacturers is nearly universal. Exif fields are generated at the creation of the image and should not be modified after with the aim of including additional information like title or

275 280

keywords. To do this, other formats are recommended, such as XMP (eXtensible Metadata Platform). This last is an XML-based standard for creating, processing and storing standardized, extensible and proprietary metadata, created by Adobe Systems in 2001. XMP metadata can be embedded into a significant number of popular file formats. It is used in PDF and other image formats such as JPEG, JPEG 2000, GIF, PNG, TIFF and EPS.

285

290 In parallel to generic standards, standards dedicated to specific image applications also exist, such as DICOM (Digital Imaging and Communications in Medicine). This is a standard created in 1992 and widely adopted by hospitals, for handling, storing, printing and transmitting information in medical imaging. It includes a file format definition and a network communications protocol.

300 Metadata constitute the documentation of all aspects of digital files essential to their persistence, usefulness and access. Images without appropriate metadata may become hard to view, migrate to new technology, or to access among large volumes of images. When annotation is inappropriate or is missing, the representation of the visual content of images by image analysis may be an interesting alternative.

305

Image Content Representation

Born in the early 1990s, Content-Based Image Indexing (CBIR) is a discipline that exploits techniques of image analysis and databases. Indexing an image by its content consists of automatically extracting structures that describe the visual content relevantly for the considered application. These structures can describe the visual content of an image globally or locally by characterizing its distribution of color, shape and texture, or parts or objects of the image. The visual structures exhibited are considered as the index of the image, they are digitally represented with one or several multidimensional vectors called *signature* of the image.

310 315 320

Key Applications

JPEG 2000 is a standard that gathers an image file format and an algorithm of image compression, created by the Joint Photographic Experts Group committee in 2000. The coding algorithm of JPEG 2000 is similar to the JPEG one. It mainly differs in the use of wavelets instead of a DCT. Wavelets provide a decomposition of the image into a pyramid of sub-images which store different levels of resolution of the image. They can be of two types, according to the objective:

(1) a Daubechies wavelet transform that requires quantization to reduce the amount of bits representing data, as JPEG does, and then imposes lossy compression;

325 330 335

(2) a rounded version of Le Gall wavelet transform, that uses only integer coefficients and then does not require quantization, providing lossless coding.

340 The aim of this standard is not only improving compression performance but also adding features, among which are transmission error resilience and region of interest (ROI). This last offers the opportunity of storing parts of the same picture using

345 different quality. Some parts of particular interest such as faces can be stored with higher quality, to the detriment of other ones where low quality/high compression can be tolerated. The JPEG 2000 standard defines two file formats that support embedded XML

350 metadata: JP2, which supports simple XML, and JPX, which has a more robust XML system based on an embedded metadata initiative of the International Imaging Industry Association (the DIG35 specification).

355

Future Directions

Today, there is a need of a unified framework for the creation, representation, storage, access, delivery,

360 management and protection of multimedia contents. New standards for managing these contents are the international standards MPEG-21 and MPEG-7. MPEG-21 is a standard started in 1999 by MPEG

365 (Moving Picture Experts Group) and now normalized as ISO/IEC 21000. Its main objectives are to define an open framework for multimedia applications and more precisely to provide a standardized structure for various media contents and to facilitate their access, delivery, management and protection. MPEG-7 is

370 another ISO/IEC standard started by MPEG in 1998 and formally called Multimedia Content Description Interface. One of its main objectives is to provide unified and efficient searching, filtering and content identification methods for these media.

375

(Abridged)

Recommended Reading

- 380 1. Adobe XMP main page:
<http://www.adobe.com/products/xmp/index.html>. 5
2. Datta R. Joshi D. Li J. and Wang J. Z. Image retrieval: Ideas, influences, and trends of the new age. ACM Comput. Surv. 40 (2), 5 April 2008.
- 385 3. File format info.
<http://www.fileformat.info/format/could.htm>. 3
4. Ian S., Burnett, Pereira F., Van de Walle R., and Koenen R. The MPEG-21 Book. Wiley, 462 pages, 6 May 2006.
- 390 5. Manjunath B.S., Salembier P., and Sikora T. Introduction to MPEG-7: Multimedia Content Description Interface. Wiley & Sons, 6 April 2002.

6. Murray J. D. and van Ryper W. Encyclopedia of Graphics File Formats. O'Reilly, 1152 pages, 2nd edition, 1996.
- 395 7. Oleg S., Pianykh. Digital Imaging and Communications in Medicine (DICOM): A Practical Introduction and Survival Guide. Springer, 2008.
8. Salomon D., Motta G., and Bryant D. Data
- 400 Compression: The Complete Reference. Springer, 4th edition, 2006.
9. Taubman D. S. and Marcellin M. W. JPEG 2000: Image Compression Fundamentals, Standards and Practice. Kluwer International Series in Engineering and Computer Science, Secs 642, 52001.
- 405 10. Wallace G. K. The JPEG still picture compression standard. Commun. of the ACM, 34(4):30–44, 3 April 1991.

Answer the following questions:

- 1) What are the two main categories of digital images and what is the difference between them?
- 2) What are the key factors in determining image quality?
- 3) What is compression and what types of compression are mentioned in the text?
- 4) Name some types of lossless compression.
- 5) What is JPEG and what is it used for?
- 6) Name some vector image formats.
- 7) What are the file formats in which bitmap data can be saved?
- 8) What term is used for “data about data”?
- 9) Why do we need content-based image indexing?
- 10) In what way does JPEG 2000 differ from the JPEF standard?
- 11) What is the standard for multimedia applications?

Match the following terms with their definitions:

- 1) Run-length encoding (RLE)
 - 2) Bitmapped image
 - 3) Content-based information retrieval
 - 4) Color model
 - 5) Lossy compression
 - 6) TIFF
 - 7) JPEG 2000
-
- a) A file format that uses wavelet compression
 - b) An attempt to describe color in a mathematical, predictable and reproducible way
 - c) An image made up of a given number of pixels, each with a specific color value, laid out in a grid
 - d) Technology that is able to retrieve images on the basis of machine-recognizable visual criteria
 - e) Reduction in file size that involves permanent loss of information

- f) Large runs of consecutive identical data values replaced by a simple code with the data value and length of the run
- g) Tagged Image File Format

Mark the following statements as *true* or *false*:

- 1) Vector graphics are images that are completely described using mathematical definitions.
- 2) Vector drawings can usually be scaled without any loss in quality.
- 3) Common raster images include 2- and 3-D architectural drawings, flow charts, logos and fonts.
- 4) The Lempel-Ziv (LZ) compression methods are among the most popular algorithms for lossy storage.
- 5) JPEG 2000 is a newer version of the JPEG format that compresses images via wavelets.
- 6) PNG (Portable Network Graphics) was a predecessor of the GIF format for transfer and display images online.

Vocabulary

advent ['æd.vent], [-vənt] - příchod, nástup
band [bænd] - pásmo
column ['kɒl.əm] - sloupec
consecutive [kən'sek.ju.tɪv] - následný
cosine ['kɒs.aɪn] - kosinus
curve [kɜ:v] - křivka
dedicated ['ded.ɪ.keɪ.tɪd] - zaměřený, věnovaný
detriment ['det.rɪ.mənt] - škoda, újma
dimension [,daɪ'men.tʃən] - rozměr
dissemination [dɪ'sem.ɪ.neɪt] - šíření
edge [edʒ] - hrana
formula ['fɔ:m.ju.lə] - vzorec, program, plán
framework ['freɪm.wɜ:k] - rámec
geometrical primitives [,dʒi:ə'met.rɪk.əl] - geometrické útvary
grid [grɪd] - mřížka
charge coupled device [tʃɑ:dʒ] - zařízení s vázanými náboji
intentionally [ɪn'ten.ʃən.əl] - záměrně
line [laɪn] - čára, přímka
lossless ['lɒslɪs] - bezztrátový
lossy ['lɒsi] - ztrátový
means [mi:nz] - prostředek
objective [əb'dʒek.tɪv] - cíl
occurrence [ə'kʌr.əns] - výskyt
perceptual [pə'sep.tʃu.əl] - vnímavostní
persistence [pə'sɪs.təns] - stálost, vytrvalost
point [pɔɪnt] - bod
polygon ['pɒl.ɪ.gɒn] - mnohoúhelník, polygon
probability [ˌprɒb.ə'bɪl.ɪ.ti] - pravděpodobnost
proprietary [prə'praɪ.ə.tri] - vlastnický, patentovaný
replacement for [rɪ'pleɪs.mənt] - náhrada za
representation [,rep.rɪ.zen'teɪ.ʃən] - zobrazení, vyobrazení
resilience [rɪ'zɪl.i.ənt] - odolnost

row [rəʊ] - řádek
run-length coding [rʌn] - kódování délkou běhu
scalable ['skeɪ.lə.bl] - škálovatelný
scriptable ['skrɪp.tə.bl] - skriptovatelný
shade [ʃeɪd] - odstín
shape [ʃeɪp] - tvar, útvar
smoothly ['smu:ð.li] - hladce
strength [streŋθ] - síla, intenzita
string [strɪŋ] - řetězec
to allow for [ə'laʊ] - počítat s čím
to associate with [ə'səʊ.si.eɪt] - spojovat s
to be made up of [meɪd] - být složen
to capture ['kæp.tʃə] - zachytit
to convey [kən'veɪ] - sdělit, vyjádřit
to differ in st ['dɪf.ə] - lišit se v něčem
to discard [dɪ'skɑ:d] - odhodit
to distinguish [dɪ'stɪŋ.gwɪʃ] - rozlišit
to embed st in st [ɪm'bed] - zapustit, zasadit do
to employ [ɪm'plɔɪ] - použít, využít
to exploit [ɪk'splɔɪt] - využít, zužitkovat
to facilitate [fə'sɪl.ɪ.teɪt] - usnadnit
to focus on st ['fəʊ.kəs] - zaměřit se na
to challenge ['tʃæl.ɪndʒ] - zpochybnit, vznést námitky
to migrate [maɪ'greɪt] - přesunout, přemístit
to mitigate ['mɪt.ɪ.geɪt] - zmírnit
to neglect st [nɪ'glekt] - zanedbat, opomenout
to preserve [prɪ'zɜ:v] - uchovat
to render ['ren.də] - zobrazit
to retain [rɪ'teɪn] - uchovat
to round [raʊnd] - zaokrouhlit
to sacrifice ['sæk.rɪ.faɪs] - obětovat
to settle on st ['set.l] - rozhodnout se pro
to shrink [ʃrɪŋk] - zmenšit
to stand for [stænd] - znamenat
to tie [taɪ] - vázat se, souviset
to treat st as [tri:t] - považovat co za
tuple [tju:pəl] - n-tice
unlike [ʌn'laɪk] - na rozdíl
versatile ['vɜ:s.ə.taɪl] - všestranný, univerzální

Phrases

in a zigzag fashion - klikatě, cikcak

in contrast to - naproti tomu, na rozdíl

in parallel to - souběžně

in particular - obzvláště

widely used - hojně využíváný

with the aim of agreeing - s cílem shodnout se