INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

T.88

Amendment 1
(06/2003)

SERIES T: TERMINALS FOR TELEMATIC SERVICES

Information technology – Lossy/lossless coding of bi-level images

Amendment 1: Encoder

ITU-T Recommendation T.88 (2000) – Amendment 1

INTERNATIONAL STANDARD ISO/IEC 14492
ITU-T RECOMMENDATION T.88

Information technology – Lossy/lossless coding of bi-level images

Amendment 1

Encoder

Summary

The objective of this amendment is to extend the ITU-T Rec. T.88 | ISO/IEC 14492 ('JBIG2') standard for compression of bi-level images, currently specified in terms of compliant data streams and decoding, to also cover compliant encoding. In particular, this amendment:

1) specifies a compliant architecture for JBIG2 encoding; and
2) illustrates example encoding methods for each optional architecture component.

Specification of the compliant architecture facilitates applications needing the benefits of a combined encoding/decoding standard, while still permitting competition in the encoding method chosen for any component. The compliant encoding method(s) chosen as examples for each architectural component were selected to target expired (or expiring) patents, openly published methods or royalty-free patents.

Selection of a representative method for each optional component should result in reasonable encoder performance in both compression and quality. Using the full combination of components, together with their example encoding methods, provides new users with a default encoder design and a benchmark for JBIG2 encoding that illustrates reasonable near-lossless quality for documents with 300 dpi, or higher, resolution.

Source

Amendment 1 to ITU-T Recommendation T.88 (2000) was approved on 29 June 2003 by ITU-T Study Group 16 (2001-2004) under the ITU-T Recommendation A.8 procedure. An identical text is also published as ISO/IEC 14492, Amendment 1._This amendment includes the modifications introduced by erratum 1 on 16 December 2004.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

Introduction

In this amendment, the following new materials are added:

a) new clauses 9, 10, and 11 to describe the required architecture and procedures for JBIG2 encoding; and

b) a new Annex J to document optional JBIG2 encoding methods.

The encoding procedures in clauses 9 and 10 are essentially the inverse of the decoding procedures already described in clauses 6 and 7 of ITU-T Rec. T.88 | ISO/IEC 14492. To simplify the required new documentation, description of each of the encoding procedures is given by referring to the corresponding decoding procedures in clauses 6 and 7, wherever applicable. Clause 11 and Annex J, however, are new material and thus contain more detailed documentation. In clause 11 (although the encoding complements that of clause 8 of ITU-T Rec. T.88 | ISO/IEC 14492), JBIG2 encoding architecture as well as its technical components are described, and their corresponding implementation methods are given by reference. In J.1, compliant example encoding methods are summarized in table form.

INTERNATIONAL STANDARD
ITU-T RECOMMENDATION

Information technology – Lossy/lossless coding of bi-level images

Amendment 1

Encoder

1) New clauses 9, 10, and 11

Add the following clauses:

9 Encoding procedures

The encoding procedures in this clause are essentially the inverse of the decoding procedures already described in clause 6 and will not be duplicated here. The inverse of generic region encoding is described in 6.2. The inverse of generic refinement encoding is described in 6.3. The inverse of text region encoding is described in 6.4. The inverse of symbol dictionary encoding is described in 6.5. The inverse of halftone region encoding is described in 6.6. The inverse of pattern dictionary encoding is described in 6.7.

10 Control encoding procedures

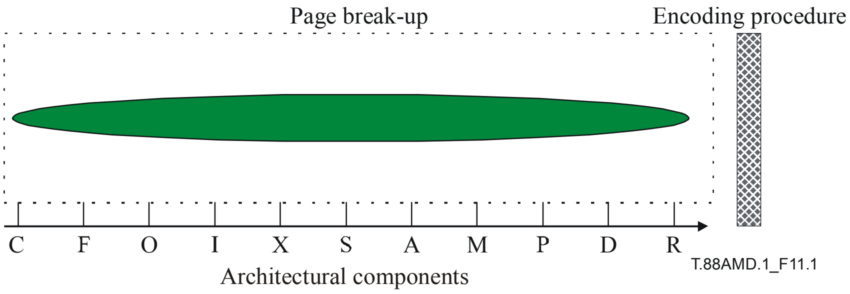
The control encoding procedures in this clause are essentially the inverse of the decoding control procedures already described in clause 7 and will not be duplicated here. The inverse of segment header syntax encoding is described in 7.2. The inverse of segment type encoding is described in 7.3. The segment types syntax for the region segment information field, symbol dictionary segment, text region segment, pattern dictionary segment, halftone region segment, generic region, generic refinement segment, end of page segment, end of stripe segment, end of file segment, profiles segment, code table segment and extension segment are described in detail in 7.4.1 to 7.4.15 respectively.

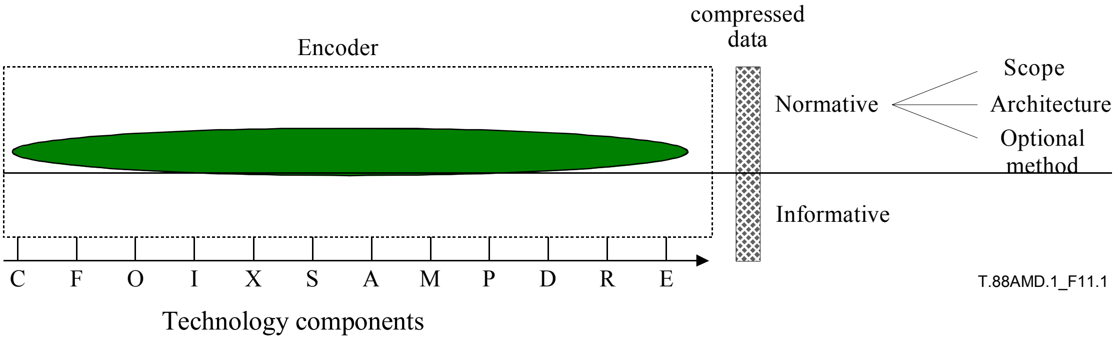
11 Page break-up

The page break-up ("Front end") procedures in this clause ~~9~~ are conceptually the inverse of the page make-up ("Back end") procedures already described in clause 8. However, page~~Page~~ break-up also requires additional page and document decomposition steps prior to encoding.

11.1 Page break-up architecture~~Encoder model~~

This clause describes the JBIG2 ~~encoding architecture~~encoder break-up defined by compliant, but optional, technical 'components' (with a range of 'algorithms' possible to implement each of these components). These JBIG2 page break-up~~encoding~~ components are a set of processing steps labelled: Capture, Filter, Orient (de-skew), Identify, eXtract, Screen, Align (register), Match, Post-match, Dictionary (optimize), and Refine~~ and Encode (bitstream generation)~~. An example sequence of this component set is illustrated in the ~~Encoder~~ Architectural~~e~~ Components figure below as the horizontal axis with abbreviated labels C F O I X S A M P D R ~~E~~ (leading from input on the left to a compressed data stream on the right). The vertical dimension above each label represents the range of possible algorithms that may be used to implement each component. The horizontal band illustrates an example JBIG2 compliant ~~encoding~~ page break-up method, using some algorithm for each architectural component ~~of this JBIG2 encoding~~ and spanning over~~all~~ these components.





A compliant JBIG2 encoder need not include all architectural components, nor use them in exactly the above sequence.

11.2 Page image decomposition

A page image is decomposed into several groups of sub-images such as marks [J2], line-arts, residues and halftones. Each group is identified and then compressed using an appropriate set of processing (architectural components) from those summarized in 11.2.1 to 11.2.1~~2~~1. Processing may include one or more of these component techniques prior to bitstream creation. The specific algorithm selected for each processing step is left up to the implementer but compliant examples for each processing step are provided in J.1. Implementing a full combination of these components, each using a compliant example encoding method, will result in an encoder capable of producing reasonable near-lossless quality for most 300 + dpi images.

11.2.1 Capture (rasterize)

Capture (rasterization) is a process by which an image source is converted into a two-dimensional bi-level raster image. This is done by mapping a region of the image source to a set of pixels of the raster image, and then assigning a 1-bit colour value to each pixel. In the scope of this amendment, two types of images are defined: generated and scanned images. A generated image is an image converted from a computer-generated metafile or vector graphic (e.g., a bitmap rasterized from a document created using a typical word processor), whereas a scanned image is an image obtained from a paper document by means of imaging hardware such as a scanner or facsimile.

11.2.2 Filter

In most cases, a scanning process is noisy, and the resulting scanned image may contain random pixel values not representative of the original source. These pixels or small groups of pixels are called flyspecks. It is often desireable to remove flyspecks in a scanned image to improve compression efficiency as well as visual quality of the reconstructed image. A scanned image also contains quantization errors, i.e., identical marks in the original image may be slightly different in the scanned image. Smoothing the edges of the marks helps to recover the equivalence of such identical marks in the scanned image and also improve compression efficiency. These filtering techniques are shown as a reference in J.1. Filtering is seldom required for generated images although these techniques may still be applied.

11.2.3 Orient (de-skew)

A scanned image may be skewed when it is scanned or photocopied at a slight angle, and it is often beneficial to identify and adjust any skew prior to compression. In most texts, marks (characters) are aligned in straight lines, and examining the slope of these lines that align pairs of marks yields the skew angle. Several methods of de-skewing are shown as a reference in J.1.

11.2.4 Identify

Identification of sub-image categories involves two processes: segmentation and classification. First an image is segmented into groups of sub-mages or 'regions' having similar characteristics. These regions (segments) are then classified into pre-defined categories such as textual data, line-art~~s~~ and halftones, to which appropriate compression methods are applied.

11.2.5 Extract

A symbol (character) is a mark consisting of black pixels. A symbol boundary is first traced by observing the connectivity of black pixels, and the adjacent black pixels are extracted to form a symbol. Although simply extracting all the pixels confined by the boundary may work in most cases, it does not handle nested marks. Several methods are shown as a reference in J.1.

11.2.6 Screen

Comparing an extracted mark against all the symbols in the dictionary is inefficient especially when the dictionary size is large, and relatively complex matching criteria as described in 11.2.8 are used. Simple methods, such as restricting comparisons to only be made between marks and dictionary symbols with similar width and height, can be used to find possible matching candidates. More detailed approaches are shown as a reference in J.1.

11.2.7 Align (register)

Symbols are often aligned (registered) in the dictionary using the same criteria selected for the screening method in 11.2.6. When distribution of black pixels is tested against symbols in a dictionary to find matching candidates, aligning symbols along their centroids can enhance the screening rate. More detailed approaches are shown as a reference in J.1.

11.2.8 Match

Marks are extracted from a region containing textual data and compared with existing symbols in a dictionary, in order to exploit any similarities between them for better compression. Basically, each mark is tested to determine whether it is similar enough to be considered a 'match' to one of the existing symbols. One way of matching is to first obtain a difference bitmap between the mark and a symbol and test the number of black pixels in the difference bitmap to a pre-defined threshold. Giving more weight to the clustered black pixels in the difference bitmap usually improves matching results. When a close match is found, a reference to the matching symbol in the dictionary is coded. When there is no close match, the extracted mark is stored as a new symbol in the dictionary.

11.2.9 Post-match

Several additional criteria and processing steps may be applied to the symbol dictionary to improve image quality. A best dictionary symbol shape may be determined by examining several similar symbols, which have already passed the matching step. Direct encoding of a symbol or alignment of symbol bottoms may also be used to improve symbol dictionary accuracy.

11.2.10 Dictionary (optimize)

After a symbol dictionary has been generated, it may be examined further to identify any 'singletons' [J2]. Singletons are symbols in the dictionary that have not been referenced by ~~any other symbols~~more than one mark. One may sometimes wish to remove such symbols from the dictionary and place them back into the 'residue' sub-images (which contain any residual marks). Such a residue image is compressed using a JBIG2 'generic' entropy encoder.

11.2.11 Refine

Encoded image (or symbol) bitmaps may also be subsequently 'refined' to similar (but different) bitmaps [J1], [J4]. For example, where images are first encoded in a near-lossless manner (e.g., when scanned image symbols are encoded using dictionaries), they can be subsequently 'refinement' encoded to a fully lossless representation of the original image. Also, successive dictionary symbols may be more efficiently encoded as 'refinements' of symbols encoded previously.

~~11.2.12 Encode (bitstream generation)~~

~~Bitstream encoding is the creation of the bitstream, consisting of one or more segments, representing regions, dictionaries and their file header. Bitstream syntax is described in this Recommendation | International Standard. Note that this Recommendation | International Standard does not require that any one particular bitstream be adopted for each image or sub-image.~~

11.3 Multi-page document composition

An encoder may organize multi-page document segments using a sequential, random or embedded organization as described in Annex D (File Formats). Dictionary segments may be organized into one global segment, one or more segments per page or stripe, or a combination of global and page-specific dictionary segments. Dictionary segment organization dramatically effects decoder performance and should be carefully selected for a specific application.

2) New Annex J

Add the following annex:

Annex J

Compliant example encoding methods

(This annex forms an integral part of this Recommendation | International Standard)

A normative JBIG2 encoder should perform the following:

- lossy encoding mode is mandatory – lossless is optional;
- minimum resolution of 300 dpi (to reduce matching error and accommodate a number of encoder system simplifications, which would otherwise be required for 200-dpi images);
- two or more stripes;
- scanned or generated image source.

The following table contains compliant example references to a range of 'algorithms' useable to implement each of the optional components of a compliant JBIG2 encoding architecture as described in clause 11. These example encoding methods target expired (or expiring) patents, openly published methods and royalty-free patents. For detailed descriptions of each technique, refer to the appropriate document, patent or book. This is not an exhaustive list of algorithms and/or component methods. Selecting one of each method ~~will~~ should result in reasonable encoder performance.

J.1 List of JBIG2 encoding components and corresponding algorithms

Architectural component		~~Norm/ Inform~~	Component method	Method reference (Found in J.2)
Scanned	Capture (rasterize) – Resolution – Striping	~~N~~	Digitize from sensor to bi-tonal bitmap – 300 dpi – 2 or more	– None – None
	Filter – Fly-speck remove – Quantization error remove	~~N~~	– Isolated mismatched pixels method – Single protruding pixel method	– Figure 3 of [J1] – Page 217 of [J1]
	Orient (de-skew)	~~N~~	Adjacent mark-based slope detection, Hough transform	– Pages 357-372 of [J2]
	Identify (Region) – Segment – Classify	~~N~~	– Recursive x-y cut – Texture analyzer	– Pages 372-384 of [J2] – Pages 385-388 of [J2]
	Extract – Isolate (black shapes) – Truncate (by size) – Force (to avoid generic region encoding)	~~N~~	– 8 connected boundary trace/rubout-based region-fill approach – Min/max of shape w & h – Not required (default)	– Pages 320-325 of [J2], [J3] – [J3] – None
	Screen (means of speeding up match)	~~N~~	Comparison of height, width and/or distribution of black pixels	– Pages 332-333 of [J2]
	Align (register)	~~N~~	Alignment of symbols using centroid	– Pages 332-333 of [J2]
	Match – Lossy/Lossless – Lossless	~~N~~	– Weighted XOR (WXOR) plus CSIS (combined size-independent strategy)[b] /CTM (compression-based template matching)[c] – XOR (Hamming distance = 0)	– Pages 325-332 of [J2], [J3] and [J4] – None

Architectural component		~~Norm/ Inform~~	Component method	Method reference (Found in J.2)
	Post-match – Symbol optimize – Symbol encoding – Symbol placement optimization	~~N~~	– 'Best' dictionary symbol shape (not just simple average of the marks) – Direct encoding (no refinement) – Align symbol bottoms	– None – None – None
	Dictionary – Singletons – Page spanning	~~N~~	– In dictionary[d] (default) – Sequentially incremental[e] (default)	– None – None
	Refine	~~N~~	Arithmetic (inverse of JBIG2 decoder process)	– [J4], ITU-T Rec. T.88 | ISO/IEC 14492
	Encode (bitstream generation)	~~N~~	Arithmetic or Huffman (inverse of JBIG2 decoder process)	– ITU-T Rec. T.88 | ISO/IEC 14492
Generated	Capture (rasterize) – at page or character level – Resolution – Striping	~~N~~	Rasterize from generated to bi-tonal bitmap – at page or character level – 300 dpi – 2 or more	– None – None
	Filter	~~N~~	N/A	N/A
	Orient (de-skew)	~~N~~	N/A	N/A
	~~Region~~ Identify (Region) – Segment – Classify	~~N~~	– Recursive x-y cut – Texture analyzer	– Pages 372-384 of [J2] – Pages 385-388 of [J2]
	Extract – Isolate (black shapes) – Truncate (by size) – Force (to avoid generic region encoding)	~~N~~	– 8 connected boundary trace/rub-based region-fill approach – Min/max of shape w & h – Not required (default)	– Pages 320-325 of [J2], [J3] – [J3] – None
	Screen (means of speeding up match)	~~N~~	Comparison of height, width and/or distribution of black pixels	Pages 332-333 of [J2]
	Align (register)	~~N~~	Identical (Bounding box is identical)	None
	Match[a]	~~N~~	XOR (Hamming distance = 0)	None
	Post-match – Symbol optimize – Symbol encoding – Symbol placement optimization	~~N~~	– N/A – Direct encoding (no refinement) – N/A	– N/A – None – N/A
	Dictionary – Singletons – Page spanning	~~N~~	– In dictionary[d] (default) – Sequentially incremental[e] (default)	– None – None
	Refine	~~N~~	N/A	N/A
	Encode (bitstream generation)	~~N~~	Arithmetic or Huffman (inverse of JBIG2 decoder process)	– ITU-T Rec. T.88 | ISO/IEC 14492

[a] For some cases of JBIG2 encoding from Generated data, it may be practical to also eliminate the Extract, Screen~~ing~~ and Align ~~Registration~~ components by implementing the Match~~ing~~ component before ~~rasterization, which is typically part of~~ the Capture (rasterize) component.

[b] Add CSIS to reduce substitution errors of WXOR.

[c] Select WXOR+CSIS for less computation than CTM.

[d] Minimizing number of Singletons in the dictionary will reduce dictionary memory requirements, and optimizing the sizes (h, w) of Singletons in the dictionary may increase the compression.

[e] Random page access is ~~best~~ enabled if, for example, one uses a single multi-page-spanning 'common' dictionary, plus multiple 'page unique' dictionaries.

J.2 Method references

[J1] HOWARD (P.): Lossless and Lossy Compression of Text Images by Soft Pattern Matching, *Proc. 1996 IEEE Data Compression Conf. (DCC)*, pages 210-219, Snowbird, Utah, March 1996.

[J2] WITTEN (I.H.), MOFFAT (A.), BELL (T.C.): Managing Gigabytes, *Morgan Kaufmann Publishers*, Second Edition, 1999.

[J3] US Patent, 4410916, Compression Labs, Inc., *Dual mode facsimile coding system and method*, October 1983.

[J4] MOHIUDDIN (K.), RISSANEN (J.), ARPS (R.): Lossless binary image compression based on pattern matching, *Proc. Int. Conf. on Computers, Systems, and Signal Processing*, pages 447-451, 1984.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	**Terminals for telematic services**
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communications
Series Y	Global information infrastructure and Internet protocol aspects
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2005