Title: TRANSFORM COEFFICIENT COMPRESSION USING MULTIPLE SCANS

Abstract: A transform coefficient block (210) of a frequency domain representation of a digital image is processed by performing scans (212, 214, 216) on at least three different regions of the block (210).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:
— with international search report
TRANSFORM COEFFICIENT COMPRESSION USING MULTIPLE SCANS

BACKGROUND

[0001] Data compression is used for reducing the cost of storing large data files on computers, as well as reducing the time for transmitting large data files between computers. In the so-called “transform methods” data is transformed into coefficients that represent the data in a frequency domain. Coefficients may be quantized (lossy compression) without significantly affecting the quality of data that is reconstructed from the quantized coefficients. Redundancy in the coefficients may then be reduced or eliminated without affecting quality of the reconstructed data (lossless compression).

[0002] One class of transforms is the discrete cosine transform. The DCT puts most of the image information in a small number of coefficients. The majority of the coefficients can be quantized to smaller bit sizes in order to gain compression.

[0003] The DCT is fast to calculate. However, performing lossless compression on the DCT coefficients can be expensive and complex.

SUMMARY

[0004] According to one aspect of the present invention, a transform coefficient block of a frequency domain representation of a digital image is processed by performing scans on at least three different regions of the block. Other aspects and advantages of the present invention will become apparent from the following detailed description, taken in conjunction with the accompanying drawings, illustrating by way of example the principles of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0005] Figure 1 is an illustration of a method of compressing a digital image;

[0006] Figure 2 is an illustration of a transform coefficient block of a
frequency domain representation of the digital image.

[0007] Figure 3 is an illustration of a method of performing context-based coding on a block of the frequency domain representation.

[0008] Figure 4 is an illustration of a method of reconstructing a digital image from a bitstream.

[0009] Figure 5 is an illustration of apparatus for performing compression and reconstruction of a digital image.

DETAILED DESCRIPTION

[0010] As shown in the drawings for purposes of illustration, the present invention is embodied in a method for compressing digital images. The method is especially efficient for compressing digital images containing text and other shapes having horizontal and vertical edges. The method may be used by printers and other machines having separate pipelines for text and graphics.

[0011] The method will be described in connection with the discrete cosine transform. However, the method is not limited to DCT transforms. It may be used with Wavelets-based transforms and other transforms in which energy is concentrated (e.g., most of the energy in the low frequency components).

[0012] Reference is made to Figure 1, which shows a method of compressing a digital image. The digital image includes an array of pixels. In the spatial domain, each pixel is represented by an n-bit word. In a typical 24-bit word representing RGB color space, for instance, eight bits represent a red component, eight bits represent a green component and eight bits represent a blue component.

[0013] The digital image is transformed from the spatial domain to a frequency domain (110). A discrete cosine transform may be used to transform blocks of pixels in the spatial domain to blocks of DCT coefficients in the frequency domain. For example, 8x8 blocks of pixels may be transformed to 8x8 blocks of DCT coefficients.

[0014] Lossy compression is performed on the blocks of transform coefficients (112). For example, the DCT coefficients may be quantized.
Quantization rounds off the DCT coefficients to zero and non-zero values.

[0015] Additional reference is now made to Figure 2, which shows an 8x8 block 210 of DCT coefficients. The DC coefficient is in the upper left hand corner, and frequency increases towards the lower right hand corner. Typically, the quantized higher frequency coefficients will be equal to zero.

[0016] Lossless compression of the transform coefficients is then performed (114). Scans 212, 214 and 216 are performed on three different regions of each transform coefficient block (116). The first region includes, and the first scan 212 covers, those coefficients lying along a diagonal of the transform coefficient block 210. The second region includes, and the second scan 214 covers, those coefficients lying above the first region. The third region includes, and the third scan 216 covers, those coefficients lying below the first region. The second scan 214 (covering the coefficients in the second region) tends to cover horizontal edges, whereas the third scan 216 (covering the coefficients in the third region) tends to cover vertical edges.

[0017] Each scan may progress from the low frequency coefficients to the high frequency components. Typically the DC coefficient is not scanned because it is coded separately. Preferably, each scan 212, 214 and 216 covers the same number of coefficients. In the 8x8 block of transform coefficients shown in Figure 2, each scan 212, 214 and 216 covers twenty one coefficients.

[0018] The coefficients are coded, one block at a time (118). Moreover, the scans of each block are coded separately. For example, the DC coefficient is coded and added to an output bitstream, the coefficients covered by the first scan 212 are coded and added to the bitstream, then the coefficients covered by the second scan 214 are coded and added to the bitstream, and then the coefficients covered by the third scan 216 are coded and added to the bitstream. The coding reduces the number of bits without reducing image information. The coding may be performed in any number of ways. As examples, the coefficients in each scan may be coded by conventional Huffman coding followed by run-length encoding, or they may be coded by entropy encoding or arithmetic coding.

[0019] Reference is now made to Figure 3, which shows yet another
coding method: context-based coding. The context-based coding is based on the assumption that the coefficients in a scan will typically have different distributions. The context-based coding assigns different codebooks to different distributions. For example, a first codebook is assigned to coefficients displaying a narrow distribution centered about zero, and a different codebook is assigned to coefficients displaying a wide distribution centered about zero. This approach tends to be more efficient than using the same codebook for the different distributions.

[0020] The context-based coding may be performed on each block as follows. The DC coefficient is coded and added to the bitstream (312). If all coefficients in all scans are equal to zero (314), a special symbol indicating such is added to the bitstream (316), and the coding is finished. If all coefficients in all scans are not equal to zero (314), a special symbol indicating such is added to the bitstream (318), and the first scan is examined (326).

[0021] The last non-zero coefficient in the scan is found, and its position is coded and added to the bitstream (320). Then, the coefficients in the scan are processed (322) in reverse order, from the last non-zero coefficient in the scan to the first. If a scan contains all zero coefficients, the position of the last non-zero coefficient may be coded as a zero, and no coefficients would be processed. Another scan is examined (326) until all scans have been coded (324).

[0022] The coefficients in a scan may be processed (322) by using the \( n^{th} \) coefficient in the scan as context for the \( n-1^{th} \) coefficient in the scan. The \( n^{th} \) coefficient is used to select one of multiple codebooks for the \( n-1^{th} \) coefficient, and the selected codebook is used to provide a codeword for the \( n-1^{th} \) coefficient. Path length and magnitude of each coefficient may be coded. The codeword corresponding to the \( n-1^{th} \) coefficient is added to the bitstream.

[0023] Consider the following example of coefficients in a scan: 153, -41, -8, -1, -1, 1, 0, 1, 0, 0, ... 0, 0. Now consider the following rule for assigning codebooks: a codebook \( c_0 \) is assigned to a coefficient preceding a 0, a codebook \( c_1 \) to a coefficient preceding a \( \pm1 \), a codebook \( c_2 \) to a
coefficient preceding a ±2, a codebook c₃ to a coefficient preceding a ±3 or ±4, a codebook c₄ to a coefficient preceding a ±5 or ±6 or ±7 or ±8, and codebook c₅ to all other coefficients. The codebooks are assigned as shown below in Table 1. A codeword for 153 is taken from codebook c₅, a codeword for -41 is taken from codebook c₄, a codeword for -8 is taken from codebook c₁, and so on.

<table>
<thead>
<tr>
<th>Coeff. No.</th>
<th>Value</th>
<th>Codebook Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>153</td>
<td>Assign codebook c₅</td>
</tr>
<tr>
<td>2</td>
<td>-41</td>
<td>Assign codebook c₄</td>
</tr>
<tr>
<td>3</td>
<td>-8</td>
<td>Assign codebook c₁</td>
</tr>
<tr>
<td>4</td>
<td>-1</td>
<td>Assign codebook c₁</td>
</tr>
<tr>
<td>5</td>
<td>-1</td>
<td>Assign codebook c₁</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>Assign codebook c₀</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>Assign codebook c₁</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Start assigning here. Assign codebook c₀</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

[0024] The compression method was just described for a single color channel. For a color digital image having multiple color channels (e.g., RGB, YUV), the method is performed on each color channel. Resulting are nine scans per block, which are coded separately. Context from the luminance channel may be used to code the chrominance channels. If a luminance value is 0, it may be assumed that the chrominance component is also zero.

[0025] Reference is now made to Figure 4. A digital image is reconstructed by decoding a bitstream into frequency domain coefficients (410); filling in at least three different regions of each transform coefficient block with the decoded frequency domain components to produce a frequency domain representation (412); and performing an inverse transform on the frequency domain representation (414).
[0026] Reference is now made to Figure 5, which shows a machine 510 that performs one or both of the compression and reconstruction methods described above. The machine 510 includes a processor 512 and memory 514. The memory 514 stores a program 516 that, when executed, causes the processor 512 to compress or reconstruct the digital image as described above.

[0027] The compression method is not limited to the number of scan patterns and the shape of the scan patterns described above. Thus the compression method is not limited to three scan patterns having zig-zag shapes. The shapes of the scan patterns may be selected according to properties of the digital image.

[0028] Different scans may be non-overlapping, or they may overlap certain coefficients. Different scans may cover different numbers of transform coefficients, or they may cover the same number of coefficients.

[0029] More than three scans may be used. However, increasing the number of scans reduces the number of coefficients in each scan.

[0030] The method is not limited to 8x8 blocks of transform coefficients. Blocks of other sizes may be used.

[0031] The present invention is not limited to the specific embodiments described and illustrated above. Instead, the present invention is construed according to the claims that follow.
TRANSFORM COEFFICIENT COMPRESSION USING MULTIPLE SCANS

THE CLAIMS:

1. Apparatus (510) for processing a frequency domain representation of a digital image, the apparatus (510) comprising a processor (512) for performing scans (212, 214, 216) on at least three different regions of at least one block (210) of the frequency domain representation.

2. The apparatus of claim 1, wherein the scans (212, 214, 216) cover the same number of coefficients.

3. The apparatus of claim 1, wherein the regions are selected according to properties of the image.

4. The apparatus of claim 1, wherein the regions are optimized for edges in the image.

5. The apparatus of claim 1, wherein the regions are non-overlapping.

6. The apparatus of claim 1, wherein a first scan (212) is performed on a first region, a second scan (214) is performed on a second region, and a third scan (216) is performed on a third region, the first region being along a diagonal of a block, the second and third regions being on opposite sides of the first region.

7. The apparatus of claim 1, wherein a zig-zag scan is performed in each region.

8. The apparatus of claim 1, wherein the processor (512) also codes the coefficients covered by the scans (212, 214, 216), the scans (212, 214, 216) being coded separately.
9. The apparatus of claim 8, wherein for each scan (212, 214, 216), the last non-zero coefficient is found, and the coefficients are coded in reverse order from the last non-zero coefficient to the first coefficient.

10. The apparatus of claim 9, wherein the processor (512) chooses from different codebooks to select codewords for the coefficients, a codebook for the n\textsuperscript{th} coefficient being selected according to the n-1\textsuperscript{th} coefficient in the scan (212, 214, 216).
FIG. 1

110 TRANSFORM DIGITAL IMAGE TO FREQUENCY DOMAIN

112 PERFORM LOSSY COMPRESSION

114 PERFORM LOSSLESS COMPRESSION

116 PERFORM SCANS ON DIFFERENT REGIONS OF EACH TRANSFORM COEFFICIENT BLOCK

118 PERFORM CODING
FIG. 3

START

312

CODE DC COEFFICIENT

314

YES

ALL COEFF. IN ALL SCANS = 0?

NO

318

ADD SYMBOL

316

ADD SYMBOL

324

MORE SCANS?

YES

DONE

326

GO TO NEXT SCAN

320

CODE POSITION OF LAST NON-ZERO COEFFICIENT IN SCAN

322

PROCESS COEFFICIENTS IN SCAN, WORKING IN REVERSE ORDER FROM THE LAST NON-ZERO COEFFICIENT TO THE FIRST COEFFICIENT, USING THE nTH COEFFICIENT AS CONTEXT FOR THE n-1TH COEFFICIENT

- Assign codebook
- Select codeword
- Add codeword to bitstream
FIG. 4

1. Decode bitstream into frequency domain coefficients (410)
2. Fill in different regions of blocks with frequency domain coefficients (412)
3. Perform inverse transform (414)

FIG. 5

Machine

- Processor (512)
- Memory (514)
- Program (516)
INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G06T 9/00 H04N 7/26

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G06T H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ, INSPEC, IBM-TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
</table>

Further documents are listed in the continuation of box C. Patent family members are listed in annex.

* Special categories of cited documents:
  *A* document defining the general state of the art which is not considered to be of particular relevance
  *E* earlier document but published on or after the international filing date
  *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
  *O* document referring to an oral disclosure, use, exhibition or other means
  *P* document published prior to the international filing date but later than the priority date claimed
  *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
  *X* document of particular relevance; the claimed invention cannot be considered without it
  *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is taken alone
  *A* document member of the same patent family

Date of the actual completion of the international search: 21 March 2003
Date of mailing of the international search report: 28/03/2003

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer
Herter, J
<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>GROSSE H-J ET AL: &quot;Improved coding of transform coefficients in JPEG-like image compression schemes&quot; PATTERN RECOGNITION LETTERS, NORTH-HOLLAND PUBL. AMSTERDAM, NL, vol. 21, no. 12, November 2000 (2000-11), pages 1061-1069, XP004237631 ISSN: 0167-8655 abstract; figures 1,2</td>
<td>1-10</td>
</tr>
<tr>
<td>A</td>
<td>PANCHANATHAN S ET AL: &quot;JPEG based scalable image compression&quot; COMPUTER COMMUNICATIONS, ELSEVIER SCIENCE PUBLISHERS BV, AMSTERDAM, NL, vol. 19, no. 12, 1 October 1996 (1996-10-01), pages 1001-1013, XP004052784 ISSN: 0140-3664 the whole document</td>
<td>1-10</td>
</tr>
</tbody>
</table>