



(22) Date de dépôt/Filing Date: 1995/10/11

(41) Mise à la disp. pub./Open to Public Insp.: 1996/05/09

(45) Date de délivrance/Issue Date: 2001/12/18

(30) Priorité/Priority: 1994/11/08 (08/335,651) US

(51) Cl.Int.⁶/Int.Cl.⁶ H04N 7/50

(72) Inventeurs/Inventors:

Vuong, Bao, US;

Hoogenboom, Chris, US

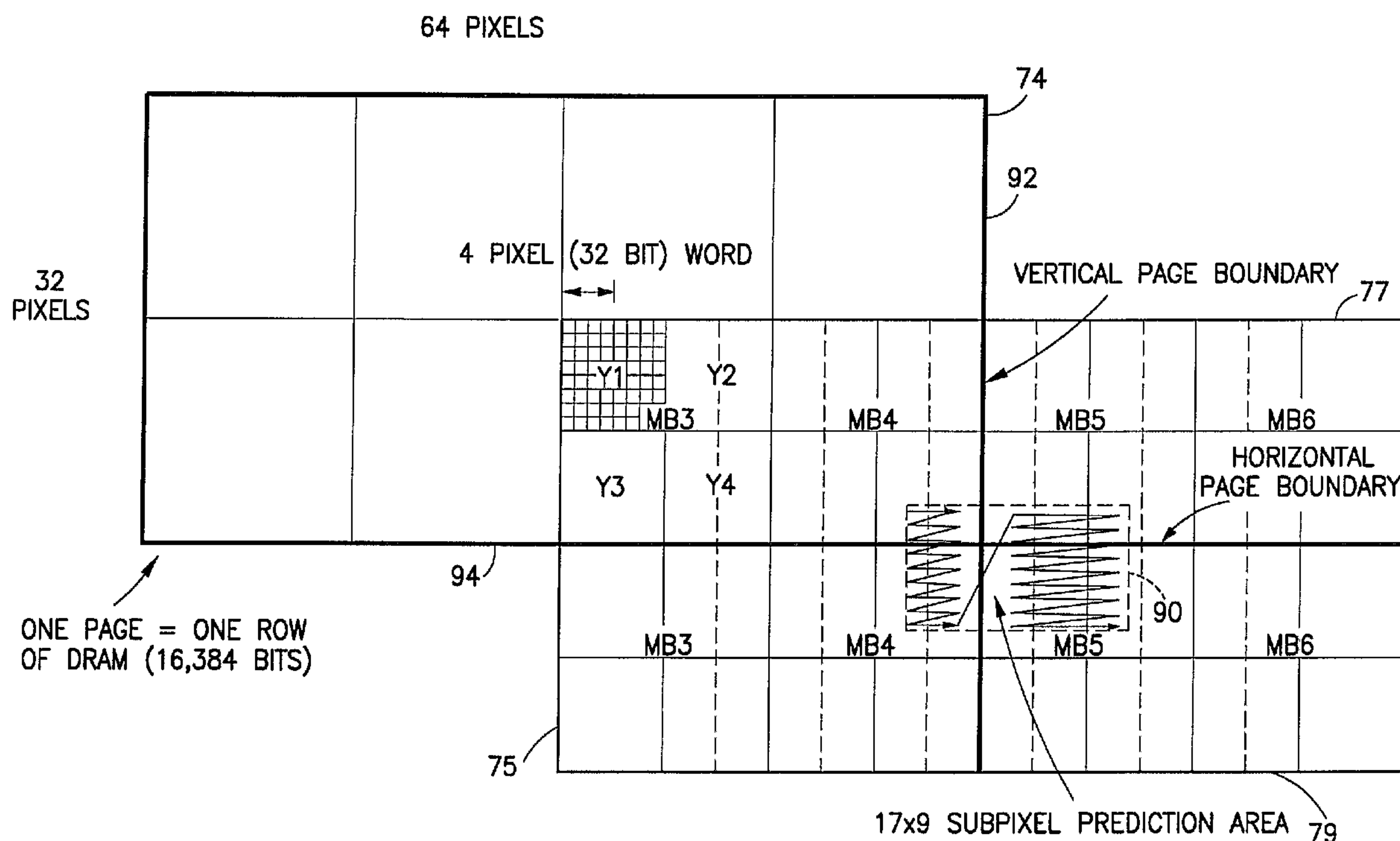
(73) Propriétaire/Owner:

GENERAL INSTRUMENT CORPORATION OF
DELAWARE, US

(74) Agent: RIDOUT & MAYBEE

(54) Titre : FILTRE D'INTERPOLATION DE PIXELS POUR PROCESSEUR DE DECOMPRESSION VIDEO

(54) Title: PIXEL INTERPOLATION FILTERS FOR VIDEO DECOMPRESSION PROCESSOR



(57) Abrégé/Abstract:

A method and apparatus are disclosed for interpolating pixels to obtain subpels for use by a video decompression processor. A prediction area is defined from which subpels are necessary to decompress a portion of a video image. Instead of reading all of the pixels from the prediction area and then processing them together to perform the necessary interpolation, portions of the pixel data are read and simultaneously averaged using in-place computation in order to reduce hardware requirements. Rounding of subpixel results is achieved using the carry input of conventional adders to add a binary "1" to the averaged pixels, which are subsequently truncated to provide the interpolated subpels.

ABSTRACT

5 A method and apparatus are disclosed for
interpolating pixels to obtain subpels for use by a
video decompression processor. A prediction area is
defined from which subpels are necessary to
decompress a portion of a video image. Instead of
reading all of the pixels from the prediction area
and then processing them together to perform the
necessary interpolation, portions of the pixel data
10 are read and simultaneously averaged using in-place
computation in order to reduce hardware
requirements. Rounding of subpixel results is
achieved using the carry input of conventional
adders to add a binary "1" to the averaged pixels,
15 which are subsequently truncated to provide the
interpolated subpels.

PIXEL INTERPOLATION FILTERS FOR VIDEO
DECOMPRESSION PROCESSOR

5 The present invention relates to a video
decompression processor, and more particularly to an
efficient scheme for providing horizontal, vertical
and/or bidirectional interpolation of prior frame
pixel data necessary to reconstruct a current video
frame.

10 Digital transmission of television signals can
deliver video and audio services of much higher
quality than analog techniques. Digital
transmission schemes are particularly advantageous
for signals that are broadcast via a cable
15 television network or by satellite to cable
television affiliates and/or directly to home
satellite television receivers. It is expected that
digital television transmitter and receiver systems
will replace existing analog systems just as digital
20 compact discs have replaced analog phonograph
records in the audio industry.

 A substantial amount of digital data must be
transmitted in any digital television system. In a
digital television system, a subscriber receives the
25 digital data stream via a receiver/descrambler that
provides video, audio and data to the subscriber.
In order to most efficiently use the available radio
frequency spectrum, it is advantageous to compress
the digital television signals to minimize the
30 amount of data that must be transmitted.

The video portion of a television signal comprises a sequence of video "frames" that together provide a moving picture. In digital television systems, each line of a video frame is defined by a sequence of digital data bits referred to as "pixels." A large amount of data is required to define each video frame of a television signal. For example, 7.4 megabits of data is required to provide one video frame at NTSC (National Television System Committee) resolution. This assumes a 640 pixel by 480 line display is used with eight bits of intensity value for each of the primary colors red, green and blue. At PAL (phase alternating line) resolution, 9.7 megabits of data is required to provide one video frame. In this instance, a 704 pixel by 576 line display is used with eight bits of intensity value for each of the primary colors red, green and blue. In order to manage this amount of information, the data must be compressed.

Video compression techniques enable the efficient transmission of digital video signals over conventional communication channels. Such techniques use compression algorithms that take advantage of the correlation among adjacent pixels in order to derive a more efficient representation of the important information in a video signal. The most powerful compression systems not only take advantage of spacial correlation, but can also utilize similarities among adjacent frames to further compact the data. In such systems, differential encoding is usually used to transmit

only the difference between an actual frame and a prediction of the actual frame. The prediction is based on information derived from a previous frame of the same video sequence.

5 Examples of video compression systems using motion compensation can be found in Krause, et al. U.S patents 5,057,916; 5,068,724; 5,091,782; 5,093,720; and 5,235,419. Generally, such motion compensation systems take advantage of a block-
10 matching motion estimation algorithm. In this case, a motion vector is determined for each block in a current frame of an image by identifying a block in a previous frame which most closely resembles the particular current block. The entire current frame
15 can then be reconstructed at a decoder by sending the difference between the corresponding block pairs, together with the motion vectors that are required to identify the corresponding pairs. Often, the amount of transmitted data is further
20 reduced by compressing both the displaced block differences and the motion vector signals. Block matching motion estimating algorithms are particularly effective when combined with block-based spatial compression techniques such as the
25 discrete cosine transform (DCT).

Each of a succession of digital video frames that form a video program can be categorized as an intra frame (I-frame), predicted frame (P-frame), or
30 bidirectional frame (B-frame). The prediction is based upon the temporal correlation between successive frames. Portions of frames do not differ

from one another over short periods of time. The encoding and decoding methods differ for each type of picture. The simplest methods are those used for I-frames, followed by those for P-frames and then B-frames.

I-frames completely describe a single frame without reference to any other frame. For improved error concealment, motion vectors can be included with an I-frame. An error in an I-frame has the potential for greater impact on the displayed video since both P-frames and B-frames are predicted from an I-frame.

P-frames are predicted based on previous I or P frames. The reference is from an earlier I or P frame to a future P-frame and is therefore called "forward prediction." B-frames are predicted from the closest earlier I or P frame and the closest later I or P frame. The reference to a future picture (i.e., one that has not yet been displayed) is called "backward prediction." There are cases where backward prediction is very useful in increasing the compression rate. For example, in a scene in which a door opens, the current picture may predict what is behind the door based upon a future picture in which the door is already open.

B-frames yield the most compression but also incorporate the most error. To eliminate error propagation, B-frames may never be predicted from other B-frames. P-frames yield less error and less compression. I-frames yield the least compression,

but are able to provide random access entry points into a video sequence.

5 One standard that has been adopted for encoding digital video signals is the Motion Picture Experts Group (MPEG) standard, and more particularly the MPEG-2 standard. This standard does not specify any particular distribution that I-frames, P-frames and B-frames must take within a sequence. Instead, the standard allows different distributions to provide 10 different degrees of compression and random accessibility. One common distribution is to have I-frames about every half second and two B-frames between successive I or P frames. To decode P frames, the previous I-frame must be available. 15 Similarly, to decode B frames, the previous and future P or I frames must be available.

Consequently, the video frames are encoded in dependency order, such that all pictures used for prediction are coded before the pictures predicted 20 therefrom. Further details of the MPEG-2 standard (and the alternative DigiCipher®II standard) and its implementation in a video decompression processor can be found in document MC68VDP/D, a preliminary data sheet entitled "MPEG-2/DCII Video Decompression 25 Processor," © Motorola Microprocessor and Memory Technologies Group, 1994, incorporated herein by reference.

30 In order to implement video compression in practical systems, a video decompression processor is required for each digital television receiver. The development of very large scale integration

(VLSI) integrated circuit chips is currently underway to implement such video decompression processors. In consumer products such as television sets, it is imperative that the cost of the system components be kept as low as possible. One of the significant costs associated with a video decompression processor is the random access memory (RAM) required to (i) buffer the compressed data prior to decompression, (ii) store the previous frame data necessary to predict a current frame using motion estimation techniques, and (iii) buffer the decompressed data prior to its output to a video appliance such as a television set, video tape recorder or the like. Another significant cost of a decompression processor is in the hardware necessary to calculate the predictions of current frame data from prior frame data, especially when interpolation is necessary among adjacent pixels to provide subpixel ("subpel") data required for the prediction.

The efficient utilization of the random access memory referred to above by a video decompression processor, and typically implemented in external DRAM, requires a scheme that can utilize a minimal amount of memory while maintaining the required data access rates (i.e., memory bandwidth). DRAMs are typically organized as an array of rows (also referred to as "pages") and columns. One of the rules of DRAM operation is that a change in row address results in a slow access for the first data of the new row. Thus, in order to maximize DRAM I/O

bandwidth, it is desirable to read data such that it causes the minimum number of changes in the row address. Thus, it is advantageous to tailor the memory map to minimize row changes. It is further
5 advantageous to sequentially access the data stored in the memory. Such sequential accesses are fast and therefore desirable. Random accesses, on the other hand, which may require frequent changes in the row address, are slow and therefore undesirable.

10 In a video decompression processor, such as one conforming to the MPEG (Motion Picture Experts Group) or DigiCipher®II (DCII) standards, various processes, including prediction calculation, require DRAM access. When the prediction of a current frame
15 block from a previous frame is good, i.e., the prediction frame bears a close resemblance to the frame to be transmitted, only a small amount of residual error remains for transmission. This leads to a high compression efficiency. If a bad
20 prediction is made, then the residual error may be so large that the compression efficiency is adversely affected. Thus, an accurate prediction of the frame-to-frame movement in a video sequence is essential in achieving a high compression ratio.

25 For a typical video sequence, the scene may contain many objects that move independently at various speeds and directions. In order to ease hardware implementation and limit the amount of information needed to represent each movement, a
30 frame of video is often segmented into rectangular blocks. One then assumes that only the blocks are

moving with independent speeds and directions. In order to reduce system complexity and increase speed, the area which is searched for the best match between a current frame block and the previous frame may be limited to the neighborhood around the target block. This limitation in the search area is usually acceptable because the movement of an object in most typical video sequences is seldom fast enough to create a large displacement from one frame to the next. With a limited search area, it is possible to efficiently perform an exhaustive search to find the best match. Once the best match is found, the prediction frame is constructed by assembling all the best matching blocks together. To implement this in hardware, the previous frame is stored in a random access memory and the prediction frame is generated block by block from the memory by reading one pixel at a time using the proper displacement vector for that block.

20 This method produces a good prediction frame when the objects in a video sequence are displaced both vertically and horizontally by an integer number of pixels. However, for a typical video sequence, the object movements are not usually an integral number of pixels in distance. For those cases where the displacement falls between two pixels, a better prediction frame can be generated by using values that are interpolated from adjacent pixels. If one considers only the midpoints between pixels, there are three possible modes of interpolation, i.e., horizontal, vertical and

diagonal. Horizontal interpolation consists of taking the average of two horizontally adjacent pixels. Vertical interpolation is generated by computing the average between two vertically adjacent pixels. Diagonal interpolation requires the averaging of four neighboring pixels. An example of a half-pixel interpolation processor for a motion compensated digital video system can be found in commonly assigned U.S. Patent No. 5,398,079.

10

The prediction calculation required in a video decompression processor using motion compensation is one of the most difficult decoding tasks, particularly where interpolation is required. Ideally, a VLSI design for such a decompression processor will be fast, small, simple and memory bandwidth efficient. Conceptually, the easiest approach for implementing a prediction calculation function would be to read in all of the data necessary to compute the prediction in a simple sequential order, and then perform whatever interpolation filtering is required. However, such a simplistic approach is disadvantageous for various reasons. If the hardware waits until after all of the data has been read in before starting the filtering function, large amounts of storage will be required. Further, only a fixed time is available in which to calculate the prediction. If it is necessary to read in all of the data before performing the filtering, only a short period of

15

20

25

30

time is left to do the filtering itself. As a rule, more hardware is necessary when less time is available to perform a calculation. Still further, if data is read in a sequential order, many row changes and poor DRAM I/O bandwidth will result.

It would be advantageous to provide a video decompression processor in which DRAM I/O bandwidth is improved and which does not require a large amount of complex hardware in order to calculate the prediction data necessary for motion compensation. It would be further advantageous to provide efficient and compact filters for providing horizontal, vertical and bidirectional subpel interpolation, which filters can be easily implemented in a practical VLSI design. The present invention provides subpel filters having the aforementioned advantages for use in a video decompression processor.

provide the broadest possible interpretation of the invention disclosed herein.

In the drawings,

Figure 1 is a block diagram of a video decompression processor in accordance with the present invention;

5 Figure 2 is a diagrammatic illustration showing how one luminance video frame is divided into pages in accordance with the present invention;

Figure 3 is a diagrammatic illustration of one luminance page;

10 Figure 4 is a diagrammatic illustration of one chrominance page;

Figure 5 is a diagrammatic illustration of a worst case condition wherein the reading of data from a prediction area requires four different
15 memory rows to be accessed;

Figure 6 is a block diagram of a horizontal interpolation filter in accordance with the present invention;

20 Figure 7 is a block diagram of a filter for providing vertical and bidirectional interpolation in accordance with the present invention;

Figure 8 is a diagrammatic illustration of the operation of the horizontal subpel filter of Figure 6 in a situation where a prediction area has no
25 vertical page break;

Figures 9 and 10 are diagrammatic illustrations of the operation of the horizontal subpel filter of Figure 6 in a situation where a prediction area includes a vertical page break; and

Figure 11 is a diagrammatic illustration of the operation of the vertical subpel filter of Figure 7.

