

7609774-6 2803

REPLACED SHEETS		SG
☒	Description : paged 1 to paged 14	
☒	Claim(s) : paged 15 to paged 23	
☒	Drawing(s) : paged 48 to paged 98	
☒	Abstract : paged 23 to paged 27	
DATE OF RECEIPT		

METHODS OF ESTIMATING MOTION IN IMAGE DATA AND APPARATUS FOR PERFORMING SAME

Field of the Invention

5 The present invention relates to image data methods and apparatus, and more particularly to methods and apparatus for processing full motion image data.

Background of the Invention

10 Video communication which is in widespread use can be classified generally into two categories including analog and digital image communication. Analog image communication has been used for video conferencing, for example, but is generally limited by poor resolution of the transmitted and received image data. In contrast, digital image communication of full motion image data generally has significantly improved fidelity relative to analog image communication in part because digital
15 information can be generally transmitted and received in large quantities at higher data rates than analog image data.

20 As will be understood by those skilled in the art, to represent image data digitally, an image area or frame is typically made up of an array of pixels (e.g., 352 rows x 244 lines) with a digital number representing the color, luminance and chrominance of each pixel. Pixel color information typically consists of three digital values including one digital value for red, one for green and one for blue. Thus, a

This Sheet Has Been Replaced



large quantity of information is needed to describe a pixel which means that digital representations of complete image frames requires exceptionally large quantities of digital data. For example, memory capacity of about three million bytes is required to process color video of 1000x1000 pixels, which means that about ninety (90) million bytes of digital information must be processed every second if a dynamic full motion picture image of 30 frames per second is to be generated. In particular, in full motion video, not only are large blocks of data required to describe each individual picture image, but a new image or frame must be presented to the viewer at approximately thirty new images per second to create the appearance of continuous motion. Moving these large quantities of video image data across digital networks or phone lines is simply infeasible given available bandwidth.

To address this problem, digital data compression has been explored extensively as a technique for reducing the number of bits required to send a given message and have that message received and reproduced with sufficient fidelity. Data compression typically utilizes either a single shorthand notation to signal a repetitive string of bits or omits data bits from the transmitted message. This latter form of compression is typically referred to as "lossy" compression and capitalizes upon on the human mind to provide the omitted data. In still video, the JPEG standard is used for data compression and defines the method by which the still image is to be compressed. In motion video, much of the picture data remains constant from frame to frame. Therefore, the video data may be compressed by first describing a reference frame and describing subsequent frames in terms of the change from the reference frame.

Compression of digital data by performing motion estimation can also be accomplished by comparing video blocks within a previous image frame with video blocks within a present frame (both of which are created by pixel movement in the vertical and/or horizontal directions) and then finding the video block having the smallest difference signal. As will be understood by those skilled in the art, the difference signal can be obtained by determining a sum of absolute difference between blocks. Since it is difficult to compare each block in a large image with blocks from a previous frame, a search area having a predetermined size corresponding to each block



is typically set for the previous frame. The motion vector is then determined by searching in the predetermined search area in the present frame. The smaller the size of the block, the more accurate the search for movement is. However, use of a smaller block size typically increases the size of the motion vectors because as the search area increases in size, the larger the movement can be which will be found in the search area and therefore the larger the motion vector.

There are a number of methods for estimating motion. For example, there is the matching method, a gradient method and a block matching method. In particular, FIG. 1 illustrates a conventional 4x4 blocking matching method. Referring to FIG. 1, a screen can include 244 lines and 352 pixels per line which is an international standard for dynamic picture imaging for communication over telephone lines. Each pixel has a corresponding brightness and the brightness can be expressed digitally as a number in the range of 0-255 and motion on the screen can be estimated by evaluating a change in brightness of each pixel. Preferably, in searching motion, a block(s) of pixels is identified and motion is determined by searching according to a change in brightness of the block. This is achieved by assigning motion of a typical pixel in the block to motion of the block on the assumption that the block has only one motion. To illustrate, FIG. 1 illustrates a block of 4x4 or sixteen (16) pixels (Block A).

As will be understood by those skilled in the art, since the motion of the block is two-dimensional, it can be expressed by a two-dimensional motion vector according to Equation 1.

$$V(i,j) = \{V_x(i,j), V_y(i,j)\} \quad (1)$$

In addition to the above methods for estimating motion, there is also a full search method, an M-step search method, a hierarchical search method and a predictive search method. The full search method, which is illustrated by FIG. 2, is the most typically used method for estimating motion. FIG. 3 also illustrates the characteristics of a conventional sum of absolute difference (SAD) graph in the full search method, as known to those skilled in the art. In particular, when a position of an object



corresponds to A at the present time and corresponds to B at a just previous time, a block having the least SAD is determined according to the following equation and the brightness signal of the block is used to estimate the motion of the block A1 of the object, so that the position of the object is estimated based on the position of the block A1.

$$\sum_{x=0}^{N(m)-1} \sum_{y=0}^{M(m)-1} |I_i(x,y,m) - I_{i-1}(x-v_x(k), y-v_y(k), m)|$$

For example, when a search size for searching with the motion vector is the size of ± 1 pixel, nine (9) SADs are obtained and a position having the least value among the nine SADs is used as an estimate of the position to which the object has moved. These nine SADs are: $\{V(x,y): (-1,1), (0,1), (1,1), (-1,0), (0,0), (1,0), (-1,-1), (0,-1), (1,-1)\}$.

However, since the motion may not be properly estimated if the motion is beyond the search range of ± 1 pixel, the search range is typically extended to the size of ± 16 pixels but this significantly increases the number SADs that need to be checked for each of the possible motion vectors in the search range. Accordingly, while the full search method can be effective in dynamic image compression, the amount of calculation that is necessary to obtain an estimate can be excessive, so other search methods have been developed to address this limitation in the full search method. As mentioned above, these other methods include the M-step search method, the hierarchical search method and the predictive search method.

Referring now to FIG. 4A-C, a prior art M-step search method is illustrated. In the M-step search method, motion is estimated by searching with an accuracy of $2^{(M-1)}$ pels and the outskirts of the searched motion are again searched with an accuracy of $2^{(M-2)}$ pels, and then these operations are repeated. For example, when the M-step method is a 3-step method, motion is estimated first by forming the size of the block in 8 pixels as illustrated in FIG. 4A, and next the motion is estimated by forming the size



of the block in 4 pixels as illustrated in FIG. 4B. Next, the motion is estimated by forming the size of the block in 2 pixels as illustrated in FIG. 4C. Accordingly, the motion can be estimated by dividing an estimation step into a series of steps and then gradually estimating the motion in greater detail to achieve a final estimate.

5 Unfortunately, the M-step method has a disadvantage in that if the motion is estimated in error in one step, the motion search may be collected to a local minimum ("lm") as illustrated by FIG. 5.

10 In the case of a hierarchical search method, motion can be searched in a coarser resolution image of 88x61 pixels, which is a reduction from a full resolution image of 352x244 pixels, and then motion is again searched in a finer resolution image of 176x122 pixels which enlarges the outskirts of the searched motion to a greater degree. Thus, if motion is searched in the size of 4 pixels in the least finer image, it has an effect of searching the motion in the size of 16 pixels in the full resolution image.

15 Then, the motion can be estimated by searching in the size of 2 pixels in the next finer resolution image which has the effect of searching motion in the size of 4 pixels in the full resolution image. As will be understood by those skilled in the art, these steps minimize the amount of calculation necessary to estimate motion in the full resolution image. Unfortunately, the hierarchical search method has a disadvantage in that if the motion is estimated in error in the coarser resolution image, the motion search may be
20 collected to a local minimum point ("lm").

25 In the predictive search method, motion in the reference block to be searched is predicted from the motion of circumferential blocks and the outskirts around a prediction value are searched. However, the predictive search method has a disadvantage in that if the prediction value is not accurate, the motion search may be collected to a local minimum point ("lm").

30 In the above described methods, a motion vector field is used as a prediction model and a 2-D spatial model and 3-D spatial model based on a coarser resolution image are also used. In the 3-D spatial model, the motion in the current reference block is effectively predicted by effectively combining information about the motion vector field in the coarser resolution image and information about the motion vector



field in the circumferential blocks. The 3-D spatial model also includes multiple elementary models to effectively treat motion discontinuity. The 3-D spatial model is more fully described in a doctoral dissertation entitled "3-D Kalman Filter for Video and Motion Estimation", by coinventor Jac-Min Kim, Rensselaer Polytechnic Institute (1994), the disclosure of which is hereby incorporated herein by reference. However, the 3-D spatial model has disadvantages in that the model parameter set is very complex and because it is difficult to implement the motion estimation since the algorithm is complex as well and the amount of necessary calculation is large when a Kalman filter is used to estimate motion.

Thus, notwithstanding the above described methods of estimating motion in image data, there still continues to be a need for methods and apparatus for efficiently processing full motion image data in compressed format with reduced error.

Summary of the Invention

It is therefore an object of the present invention to provide improved methods and apparatus for processing full motion image data.

It is further object of the present invention to provide methods and apparatus for processing full motion image data in compressed format with reduced error.

It is another object of the present invention to provide methods and apparatus for processing and communicating full motion image data using reduced bandwidth.

These and other objects, features and advantages of the present invention are provided by methods and apparatus for estimating motion in video data so that efficient communication of video data in real time with high fidelity is feasible. Using this method, full motion image data is processed in 3-dimensions using multiple models and by adjusting a search region based on prediction errors from the models to more accurately estimate motion. In particular, according to a preferred embodiment of the present invention, a step is performed to produce a first frame of image data from a received video signal containing information describing an image data frame of predetermined resolution (e.g., 352x244 pixels). However, the first frame of image data is of coarser resolution (e.g., 176x122 pixels) relative to the image data frame of predetermined resolution. In addition, a second image data frame of coarser resolution



(e.g., 88x61 pixels) relative to the first image data frame is also produced so that a more efficient routine for detecting motion can be achieved using 3-dimensional processing.

Once the images of reduced resolution have been produced, an estimate of motion in the second image data frame is then produced and used in the estimation of motion in the first image data frame of higher resolution. The estimation of motion in the first image data frame is then used to obtain an estimation of motion in the image data frame of predetermined resolution. In particular, an estimate of motion in the second image data frame, for example, can be determined by reading a plurality of motion vectors (e.g. nine for a search size of ± 1 pixel) from an image block in a previous version of the second image data frame and reading image data in a reference block in a current version of the second image data frame. An optimum motion vector is also determined by predicting a plurality of motion vectors based on a corresponding plurality of motion models and selecting the best model and vector which yields the smallest sum of absolute difference error. Once an optimum motion vector has been predicted, the plurality of motion vectors previously read are adjusted relative to the optimum motion vector and then a determination is made as to which vector yields the smallest relative motion error so that this vector can be used as an estimation of motion from the previous version of the second image data frame to the current version of the second image data frame. These steps are then performed again for the first image data frame of higher resolution, using the estimation obtained from the second image data frame, and then performed again so that a highly accurate estimation of motion in the image data frame of predetermined resolution can be achieved. The present invention also includes apparatus for performing the above described operations.

Brief Description of the Drawings

FIG. 1 illustrates a pictorial representation of a prior art 4x4 block matching method of motion estimation.

FIG. 2 illustrates a pictorial representation of a prior art full search method of motion estimation.



FIG. 3 illustrates a sum of absolute difference graph for a full search method of motion estimation.

FIG. 4 illustrates a pictorial representation of a prior art 3-step search method of motion estimation.

FIG. 5 illustrates a sum of absolute difference graph including local minimums ("lm") for a M-step search method of motion estimation.

FIG. 6 is a block diagram illustrating an apparatus for performing motion estimation according to the present invention.

FIG. 7 is a flow chart illustrating operations for estimating motion in image data according to an embodiment of the present invention.

FIG. 8 is a flow chart illustrating operations for estimating motion in an image data frame of full or relatively coarse resolution.

FIG. 9 is a flow chart illustrating operations for predicting an optimum motion vector based on a plurality of models.

FIG. 10 illustrates a pictorial representation showing the full, first, second and third resolution frames used to determine a motion estimation according to the present invention.

Description of Preferred Embodiments

The present invention will now be described more fully hereinafter with reference to the accompanying drawings, in which preferred embodiments of the invention are shown. This invention may, however, be embodied in different forms and should not be construed as limited to the embodiments set forth herein. Rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art. Like number refer to like elements throughout.

Referring now to FIGS. 6-10, the present invention will be described. In particular, FIG. 6 illustrates a preferred hardware configuration for an apparatus for estimating motion in image data according to the present invention. The apparatus may include an input/output unit 10 for receiving an external video signal containing image data and transmitting an estimated motion signal determined by the operations described



herein. A search memory 20 is also provided for storing the received external video signal and results obtained during intermediate stages of the preferred operations for performing motion estimation according to the present invention. A microcontroller 30 is also provided for producing, among other things, image data frames of multiple resolution relative to an image data frame of predetermined full resolution and a signal representing an estimate of motion in the received image data. A block matching circuit 50, which is responsive to the microcontroller 30, is provided for performing block matching of data preferably stored in the search memory 20. In addition, an interpolation circuit 40 for interpolating a value of the data stored in the search memory 20 in response to a signal from the microcontroller 30 and producing an interpolated signal, is provided.

Information of a general nature relating to other apparatus for performing motion estimation can be found in U.S. Patent Nos. 5,365,281, entitled *Motion Signal Detecting Circuit*; 5,392,073 entitled *Motion Estimating Device For Variable Length Encoded Data*; 5,467,086 entitled *Apparatus and Method of Coding/Decoding Video Data*; 5,489,949 entitled *Method and Apparatus for Motion Estimation*; 5,497,153 entitled *System for Variable-Length-Coding and Variable-Length-Decoding Digital Data for Compressing Transmission Data*; 5,493,344 entitled *Motion Vector Estimation Method and Apparatus Having Half-Pixel Accuracy*; 5,508,745 entitled *Apparatus for Controlling A Quantization Level to be Modified by a Motion Vector*; and 5,510,857 entitled *Motion Estimation Coprocessor*, all of which are assigned to the assignee of the present application and hereby incorporated herein by reference.

Referring now to operations S100-S900 of FIG. 7, the microcontroller 30 reads the external video signal received by the input/output unit 10, operation S100. This signal contains information to construct an image data frame of predetermined full resolution (e.g., 352x244 pixels). Referring now to operation S200, the microcontroller 30 performs the operation of producing a first image data frame from the image data frame of predetermined resolution, so that the first image data frame is of coarser resolution (e.g., 176x122 pixels), as illustrated best by FIG. 10. This



operation is then followed by a step of storing information relating to the first image data frame in the search memory 20. In operation S300, a second image data frame is produced so that the second image data frame is of coarser resolution (e.g., 88x61 pixels) relative to the first image data frame and then this frame is stored in search memory 20. Next, in operation S400, a third image data frame is produced so that the third image data frame is of coarser resolution (e.g., 44x30.5 pixels) relative to the second image data frame and then this frame is also stored in search memory 20.

Referring still to FIG. 7, the microcontroller 30 then estimates motion after a predetermined amount of time from the third image data frame to obtain a first estimation, as illustrated by operation S500. This first estimation is then stored in search memory 20. Next, the microcontroller 30 estimates motion from the second image data frame on the basis of the motion estimated in operation S500, to obtain a second estimation which is then stored in search memory 20, as illustrated by operation S600. Following this step, the microcontroller 30 estimates motion from the first image data frame on the basis of the motion estimated in operation S600, to obtain a third estimation which is stored, as illustrated by operation S700. Finally, the microcontroller 30 estimates motion in the image data frame of predetermined full resolution on the basis of the motion estimated in operation S700 and then this estimate is stored, as illustrated by operation S800. The microcontroller 30 then causes the production of a video signal containing the motion estimation corresponding to the full resolution image data frame for transmission by the input/output unit 10, operation S900.

Referring now to FIG. 8, the operation S500 for estimating motion in the third data frame to obtain an estimation of motion will be described, however the steps performed by this preferred operation are also applicable to operations S600-S800. In operation S500, the microcontroller 30 reads the motion vectors (V_s) of the blocks from a just previous image data frame (S510) and reads image data in a reference block in a current image data frame (S520). During operation S500, a optimum motion vector is also predicted from a motion model (S530) and then image data in a search



region within the current image data frame and around the predicted motion vector, is read (S540). This reading step (S540) is then followed by the step S550 of adjusting or changing the motion vectors obtained from (S510), around or relative to the predicted motion vector obtained from (S530).

5 As illustrated by steps S560-S599, a sum of absolute difference (SAD) between the reference block and block of image data at an adjusted motion vector is determined for each of the adjusted motion vectors obtained from (S550) and the lowest determined SAD and the adjusted motion vector yielding the lowest determined SAD are stored in memory, for use during subsequent operations.

10 Referring now to FIG. 9, the operation S530 for predicting an optimum motion vector according to the present invention will be described. Using this operation S530, the search region can be preferably adjusted based on prediction errors derived by analyzing multiple motion models and minimizing these errors by using a feedback loop to determine the best model to most accurately estimate motion using the steps of FIG.
15 8. In particular, at step S531, a motion vector is predicted using a first motion model and then at step S532, a sum of absolute difference is determined between a reference block and a prediction block when the reference block of a current frame is predicted from a previous video frame using the predicted motion vector based on the first
5 motion model. However, rather than using only one model, a plurality of different
20 models are used to reduce the likelihood of occurrence of a local minimum ("lm") in the difference graph of FIG. 5, which can be problematic using the methods of the prior art. For example, steps S533-S535 illustrate that the model yielding the motion vector corresponding to the lowest SAD, when step S532 is performed, will provide an
10 optimum motion vector to be used in step S530.

25 As illustrated below with respect to Table 1, when compared to the full search method and hierarchical search method described in the background section of the present application, the present invention is preferable to both methods based on the average value of the mean square error (MSE) of temporal interpolation in thirty(30) frames, when a flower garden sequence and mobile sequence are used are used.
30



	<u>Flower Garden Sequence</u>	<u>Mobile Sequence</u>
	Full Search 278	361
5	Hierarchical Search 282	273
	Present Invention 232	239

10

Table 1

In addition, Table 2 illustrates an average value of the mean square error (MSE) of motion compensated prediction in thirty (30) frames, using the above search methods. Based on the MSE of motion compensated prediction, the full search method is the most effective, but the method according to the present invention is almost as good as the full search method.

15

	<u>Flower Garden Sequence</u>	<u>Mobile Sequence</u>
	Full Search 226	296
20	Hierarchical Search 338	372
	Present Invention 245	303

25

Table 2

Referring now to Table 3, the times needed to calculate the final motion estimation by each method is illustrated. Based on time, the hierarchical search method is the most effective, but the method according to the present invention is almost as efficient based on time.

30

	<u>Flower Garden Sequence</u>	<u>Mobile Sequence</u>
	Full Search 105 sec.	105 sec.
35	Hierarchical Search 6.3 sec.	6.3 sec.

40



Present
Invention

6.5 sec.

6.5 sec.

5

Table 3

10

15

20

25

30

As described above, the microcontroller 30, in conjunction with the other structural elements of FIG. 6, is preferably designed to perform the above operations of FIGS. 7-9. Accordingly, an apparatus for compressing motion video data by motion estimation according to the present invention includes means for receiving a video signal containing information to construct an image data frame of predetermined resolution; means, coupled to said receiving means, for producing information to construct a first image data frame of coarser resolution relative to the image data frame of predetermined resolution and information to construct a second image data frame of coarser resolution relative to the first image data frame; means, coupled to said information producing means, for estimating motion in the second image data frame to obtain a first estimation and for estimating motion in the first image data frame, based on the first estimation, to obtain a second estimation; and means, responsive to said estimating means, for transmitting a video signal containing an estimate of motion occurring in the image data frame of predetermined resolution, based on the second estimation.

According to the present invention, the apparatus for estimating motion preferably includes means for reading a plurality of motion vectors and reading image data in a reference block in the second image data frame; means for predicting an optimum motion vector based on a motion model; means, responsive to said reading means and predicting means, for reading image data in a first search region in the second image data frame, adjacent the predicted optimum motion vector; means for adjusting one of said plurality of motion vectors relative to the predicted optimum motion vector, to obtain a first motion vector; and means, responsive to said adjusting means, for determining a first sum of absolute difference between the reference block and a block of image data at the first motion vector.



Finally, the apparatus for predicting an optimum motion vector preferably includes means for predicting a motion vector using a first model; means for determining a sum of absolute difference between a reference block and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the first model; means for predicting a motion vector using a second model; means for determining a sum of absolute difference between a reference block and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the second model; and means, responsive to said predicting and determining means, for determining the predicted motion vector based on the first model as an optimum motion vector, if the sum of absolute difference determined using the predicted motion vector based on the first model is less than the sum of absolute difference determined using the predicted motion vector based on the second model.

In the drawings and specification, there have been disclosed typical preferred embodiments of the invention and, although specific terms are employed, they are used in a generic and descriptive sense only and not for purposes of limitation, the scope of the invention being set forth in the following claims.



THAT WHICH IS CLAIMED IS:

1. A method of estimating motion in image data, comprising the steps of:
producing a first image data frame from an image data frame of predetermined resolution, so that the first image data frame is of coarser resolution relative to the image data frame of predetermined resolution;
producing a second image data frame from the image data frame of predetermined resolution or the first image data frame, so that the second image data frame is of coarser resolution relative to the first image data frame;
estimating motion in the second image data frame to obtain a first estimation;
estimating motion in the first image data frame based on the first estimation, to obtain a second estimation; and
estimating motion in the image data frame of predetermined resolution based on the second estimation.
2. The method of Claim 1, wherein said step of estimating motion in the second image data frame comprises the steps of:
predicting a motion vector based on a motion model;
reading image data in a first search region within the second image data frame, based on the predicted motion vector;
determining a first motion vector based on the predicted motion vector; and
determining a sum of absolute difference based on the first motion vector.
3. The method of Claim 1, wherein said step of estimating motion in the second image data frame comprises the steps of:
reading a plurality of motion vectors;
reading image data in a reference block in the second image data frame;
predicting an optimum motion vector based on a motion model;
reading image data in a first search region adjacent the predicted optimum motion vector;



adjusting one of said plurality of motion vectors relative to the predicted optimum motion vector, to obtain a first motion vector; and

determining a first sum of absolute difference between the reference block and a block of image data at the first motion vector.

4. The method of Claim 3, wherein said step of reading image data in a first search region adjacent the predicted optimum motion vector comprises reading image data in a first search region surrounding the predicted optimum motion vector.

5. The method of Claim 3, further comprising the steps of:

adjusting another of said plurality of motion vectors relative to the predicted optimum motion vector, to obtain a second motion vector;

determining a second sum of absolute difference between the reference block and a block of image data at the second motion vector; and

5 determining the first motion vector as an estimate of motion in the second image data frame instead of the second motion vector, if the first sum of absolute difference is less than the second sum of absolute difference.

6. The method of Claim 3, wherein said step of predicting an optimum motion
10 vector based on a motion model comprises the steps of:

predicting a motion vector using a model; and

determining a sum of absolute difference between a reference block and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector.

15

7. The method of Claim 6, wherein said step of predicting a motion vector using a model comprises predicting a plurality of motion vectors using a corresponding plurality of models.



8. The method of Claim 7, further comprising the steps of determining a plurality of sums of absolute difference between the reference block and a plurality of prediction blocks in the second image data frame which are determined from the reference block and corresponding ones of the plurality of motion vectors.0

5

9. The method of Claim 8, further comprising the step of determining the lowest value of the plurality sums of absolute difference.

10. The method of Claim 3, wherein said step of predicting an optimum motion vector based on a motion model comprises the steps of:

predicting a motion vector using a first model;

determining a sum of absolute difference between a reference block and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the first model;

predicting a motion vector using a second model;

determining a sum of absolute difference between a reference block and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the second model; and

determining the predicted motion vector based on the first model as an optimum motion vector, if the sum of absolute difference determined using the predicted motion vector based on the first model is less than the sum of absolute difference determined using the predicted motion vector based on the second model.

25

11. A method of estimating motion in image data, comprising the steps of:

producing a first image data frame from an image data frame of predetermined resolution; so that the first image data frame is of coarser resolution relative to the image data frame of predetermined resolution;



producing a second image data frame from the image data frame of predetermined resolution or the first image data frame, so that the second image data frame is of coarser resolution relative to the first image data frame;

5 producing a third image data frame from the image data frame of predetermined resolution, the first image data frame or the second image data frame, so that the third image data frame is of coarser resolution relative to the second image data frame;

estimating motion in the third image data frame to obtain a first estimation;

estimating motion in the second image data frame based on the first estimation, to obtain a second estimation;

10 estimating motion in the first image data frame based on the second estimation, to obtain a third estimation; and

estimating motion in the image data frame of predetermined resolution based on the third estimation.

15 12. The method of Claim 11, wherein said step of estimating motion in the third image data frame comprises the steps of:

predicting a motion vector based on a motion model;

reading image data in a first search region within the third image data frame, based on the predicted motion vector;

20 determining a first motion vector based on the predicted motion vector; and determining a sum of absolute difference based on the first motion vector.

13. The method of Claim 11, wherein said step of estimating motion in the third image data frame comprises the steps of:

reading a plurality of motion vectors;

25 reading image data in a reference block in the third image data frame;

predicting an optimum motion vector based on a motion model;

reading image data in a first search region adjacent the predicted optimum motion vector;

30 adjusting one of said plurality of motion vectors relative to the predicted optimum motion vector, to obtain a first motion vector; and



determining a first sum of absolute difference between the reference block and a block of image data at the first motion vector.

14. The method of Claim 13, wherein said step of reading image data in a first search region adjacent the predicted optimum motion vector comprises reading image data in a first search region surrounding the predicted optimum motion vector.

15. The method of Claim 13, further comprising the steps of:
adjusting another of said plurality of motion vectors relative to the predicted optimum motion vector, to obtain a second motion vector;
determining a second sum of absolute difference between the reference block and a block of image data at the second motion vector; and
determining the first motion vector as an estimate of motion in the third image data frame instead of the second motion vector, if the first sum of absolute difference is less than the second sum of absolute difference.

16. The method of Claim 13, wherein said step of predicting an optimum motion vector based on a motion model comprises the steps of:
predicting a motion vector using a model; and
determining a sum of absolute difference between a reference block and a prediction block in the third image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector.

17. The method of Claim 16, wherein said step of predicting a motion vector using a model comprises predicting a plurality of motion vectors using a corresponding plurality of models.

18. The method of Claim 17, further comprising the steps of determining a plurality of sums of absolute difference between the reference block and a plurality of

19

This Sheet Has Been Replaced



* C L M S *

prediction blocks in the third image data frame which are determined from the reference block and corresponding ones of the plurality of motion vectors.

19. The method of Claim 18, further comprising the step of determining the
5 lowest value of the plurality sums of absolute difference.

20. The method of Claim 13, wherein said step of predicting an optimum motion vector based on a motion model comprises the steps of:

10 predicting a motion vector using a first model;
determining a sum of absolute difference between a reference block and a prediction block in the third image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the first model;

15 predicting a motion vector using a second model;
determining a sum of absolute difference between a reference block and a prediction block in the third image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the second model; and

20 determining the predicted motion vector based on the first model as an optimum motion vector, if the sum of absolute difference determined using the predicted motion vector based on the first model is less than the sum of absolute difference determined using the predicted motion vector based on the second model.

21. An apparatus for compressing motion video data by motion estimation,
25 comprising:

means for receiving a video signal containing information to construct an image data frame of predetermined resolution;

means, coupled to said receiving means, for producing information to construct a first image data frame of coarser resolution relative to the image data frame of

20

This Sheet Has Been Replaced



* C L M S *

predetermined resolution and information to construct a second image data frame of coarser resolution relative to the first image data frame;

means, coupled to said information producing means, for estimating motion in the second image data frame to obtain a first estimation and for estimating motion in the first image data frame, based on the first estimation, to obtain a second estimation; and

means, responsive to said estimating means, for transmitting a video signal containing an estimate of motion occurring in the image data frame of predetermined resolution, based on the second estimation.

10

22. The apparatus of Claim 21, wherein said means for estimating motion comprises:

means for reading a plurality of motion vectors and reading image data in a reference block in the second image data frame;

15 means for predicting an optimum motion vector based on a motion model;

means, responsive to said reading means and predicting means, for reading image data in a first search region in the second image data frame, adjacent the predicted optimum motion vector;

20 means for adjusting one of said plurality of motion vectors relative to the predicted optimum motion vector, to obtain a first motion vector; and

means, responsive to said adjusting means, for determining a first sum of absolute difference between the reference block and a block of image data at the first motion vector.

25 23. The apparatus of Claim 22, wherein said means for predicting an optimum motion vector comprises:

means for predicting a motion vector using a model; and

30 means for determining a sum of absolute difference between a reference block and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector.

21

This Sheet Has Been Replaced



* C L M S *

24. The apparatus of Claim 22, wherein said means for predicting an optimum motion vector comprises:

means for predicting a motion vector using a first model;

5 means for determining a sum of absolute difference between a reference block and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the first model;

means for predicting a motion vector using a second model;

10 means for determining a sum of absolute difference between a reference block and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the second model; and

means, responsive to said predicting and determining means, for determining the predicted motion vector based on the first model as an optimum motion vector, if the
15 sum of absolute difference determined using the predicted motion vector based on the first model is less than the sum of absolute difference determined using the predicted motion vector based on the second model.

This Sheet Has Been Replaced



METHODS OF ESTIMATING MOTION IN IMAGE DATA AND APPARATUS FOR PERFORMING SAME

Abstract of the Disclosure

5 Methods of estimating motion in image data include the steps of producing a first frame of image data from a received video signal containing information of an image data frame of higher predetermined resolution (e.g., 352x244 pixels) and producing a second image data frame of coarser resolution (e.g., 88x61 pixels) relative to the first image data frame. An estimate of motion in the second image data frame is then produced and used to determine a estimate of motion in the first image data frame of higher resolution. The estimate of motion in the first image data frame is then used to obtain an estimate of motion in the received image data frame of predetermined full resolution. The estimate of motion in the second image data frame can be determined by reading a plurality of motion vectors from an image block in a previous version of the second image data frame and reading image data in a reference block in a current version of the second image data frame. To obtain this estimate, an optimum motion vector is determined by predicting a plurality of motion vectors based on a corresponding plurality of motion models and selecting the best model and vector which yields the smallest difference signal. Once an optimum motion vector has been predicted, the plurality of motion vectors previously read are adjusted relative to the optimum motion vector. Then a determination is made as to which one of the plurality of vectors yields the smallest relative motion error so that this vector can be used as an estimation of motion from the previous version of the second image data frame to the current version of the second image data frame. These steps are then performed again for the first image data frame of higher resolution, using the estimation obtained from the second image data frame, and then performed again so that a highly accurate estimation of motion in the image data frame of predetermined resolution can be achieved.

#24613

30

This Sheet Has Been Replaced



FIG. 1 (Prior Art)

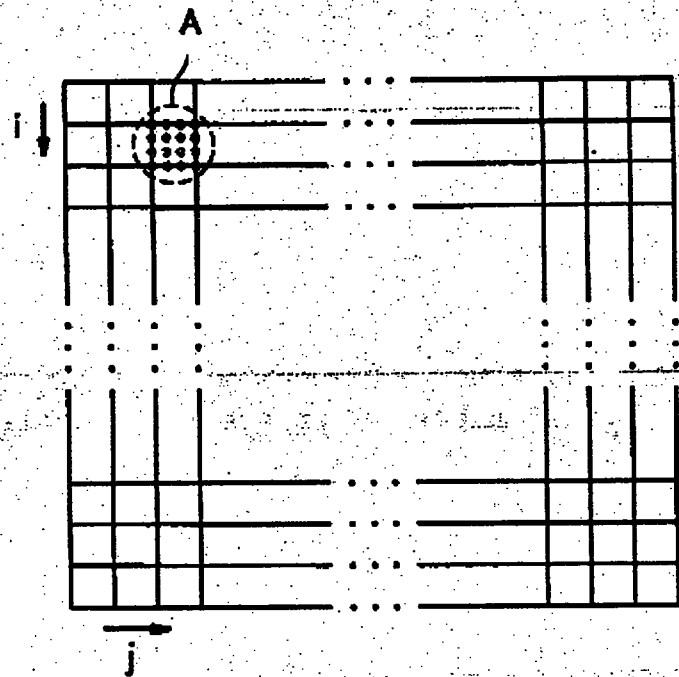
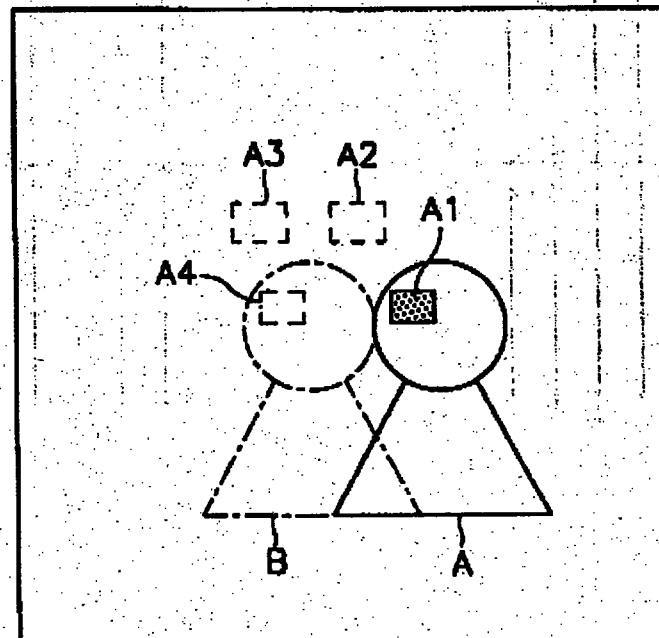


FIG. 2 (Prior Art)



This Sheet Has Been Replaced



FIG.3(Prior Art)

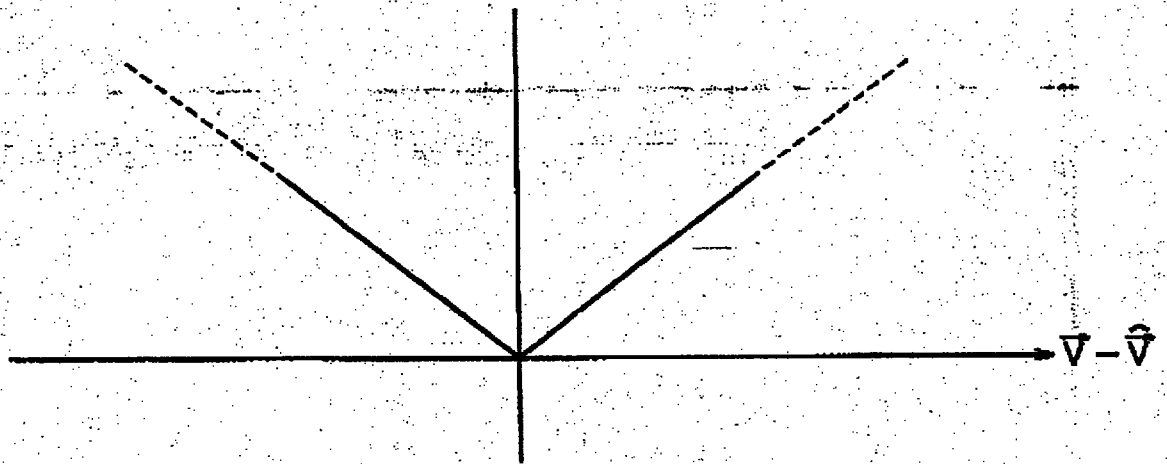
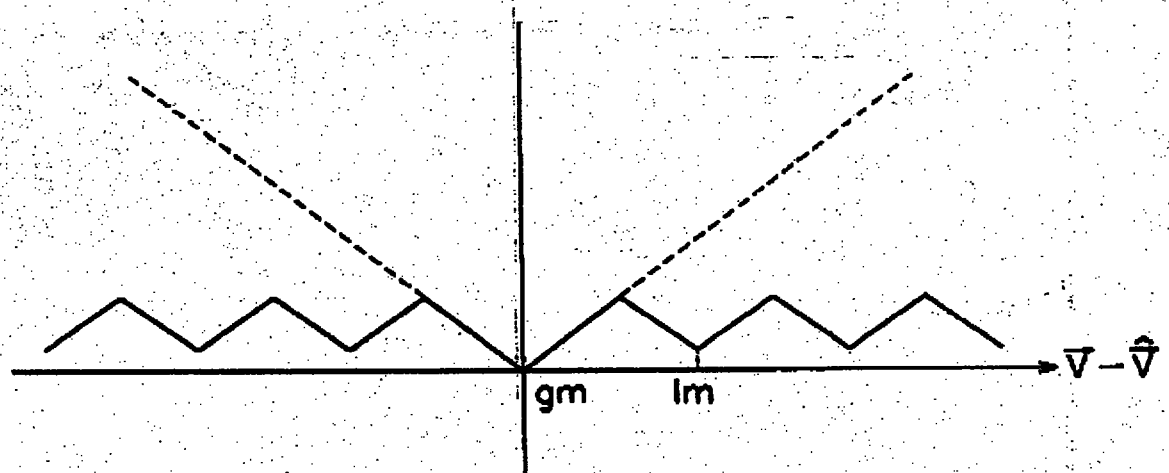


FIG.5(Prior Art)



This Sheet Has Been Replaced



FIG.4A (Prior Art)

8pixels

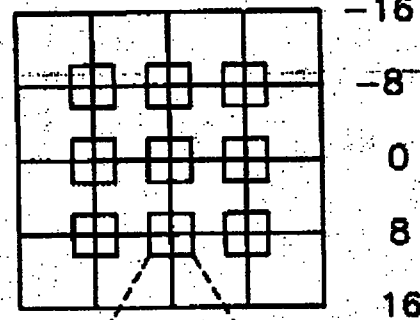


FIG.4B (Prior Art)

4pixels

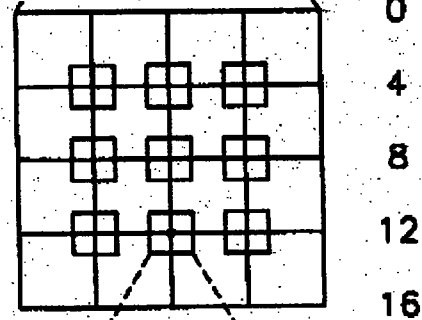
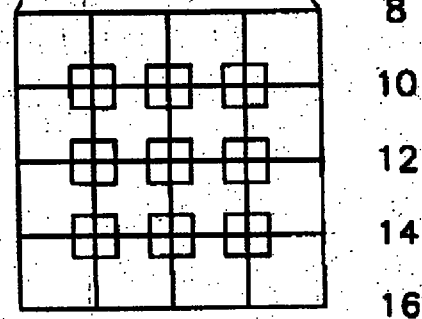


FIG.4C (Prior Art)

2pixels



This Sheet Has Been Replaced



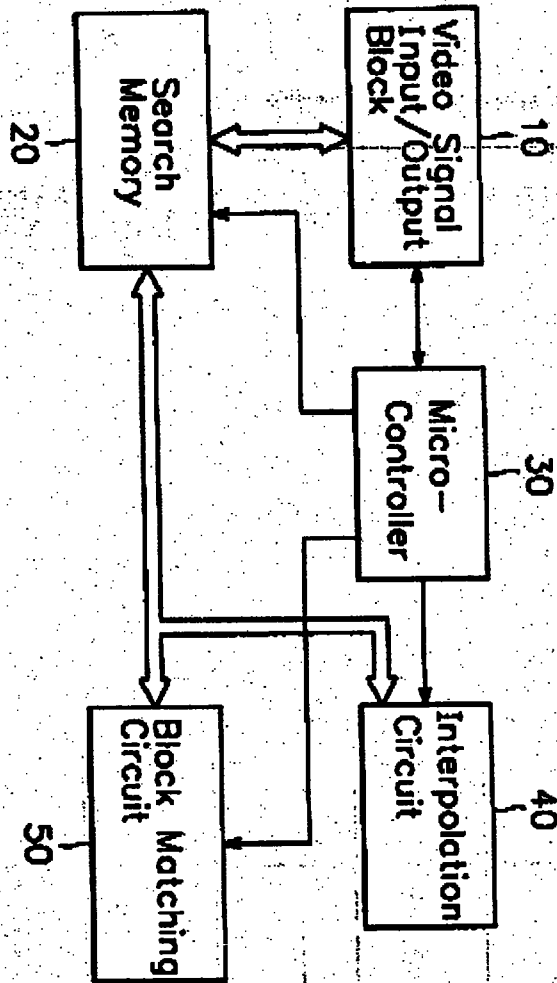
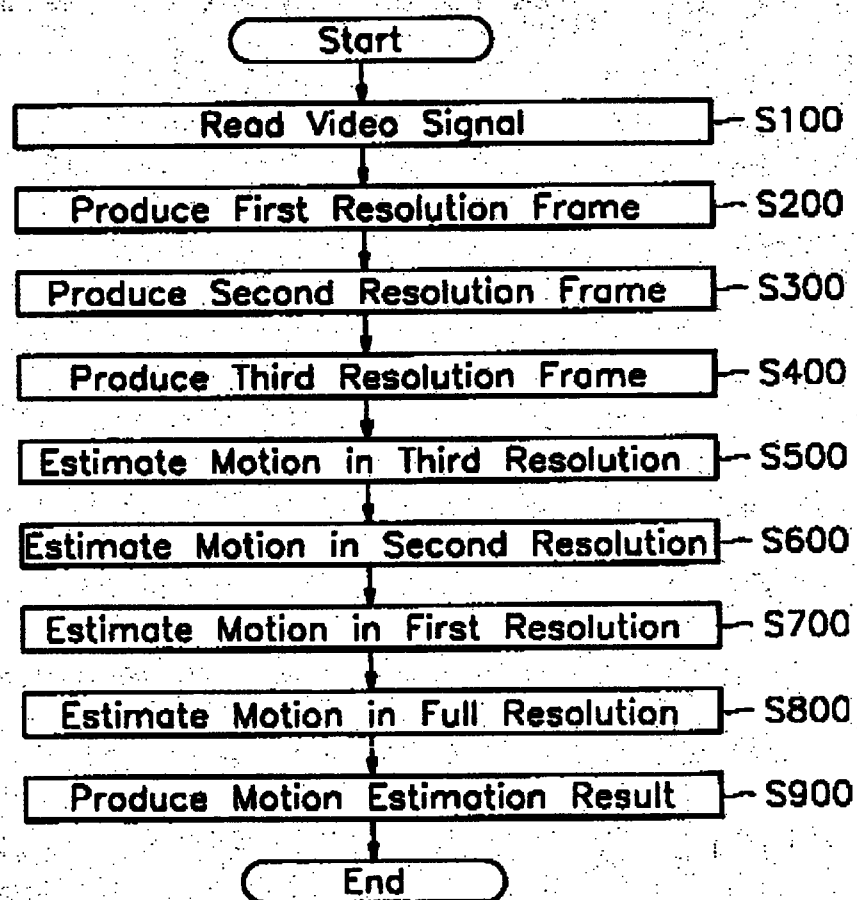


FIG. 6

This Sheet Has Been Replaced



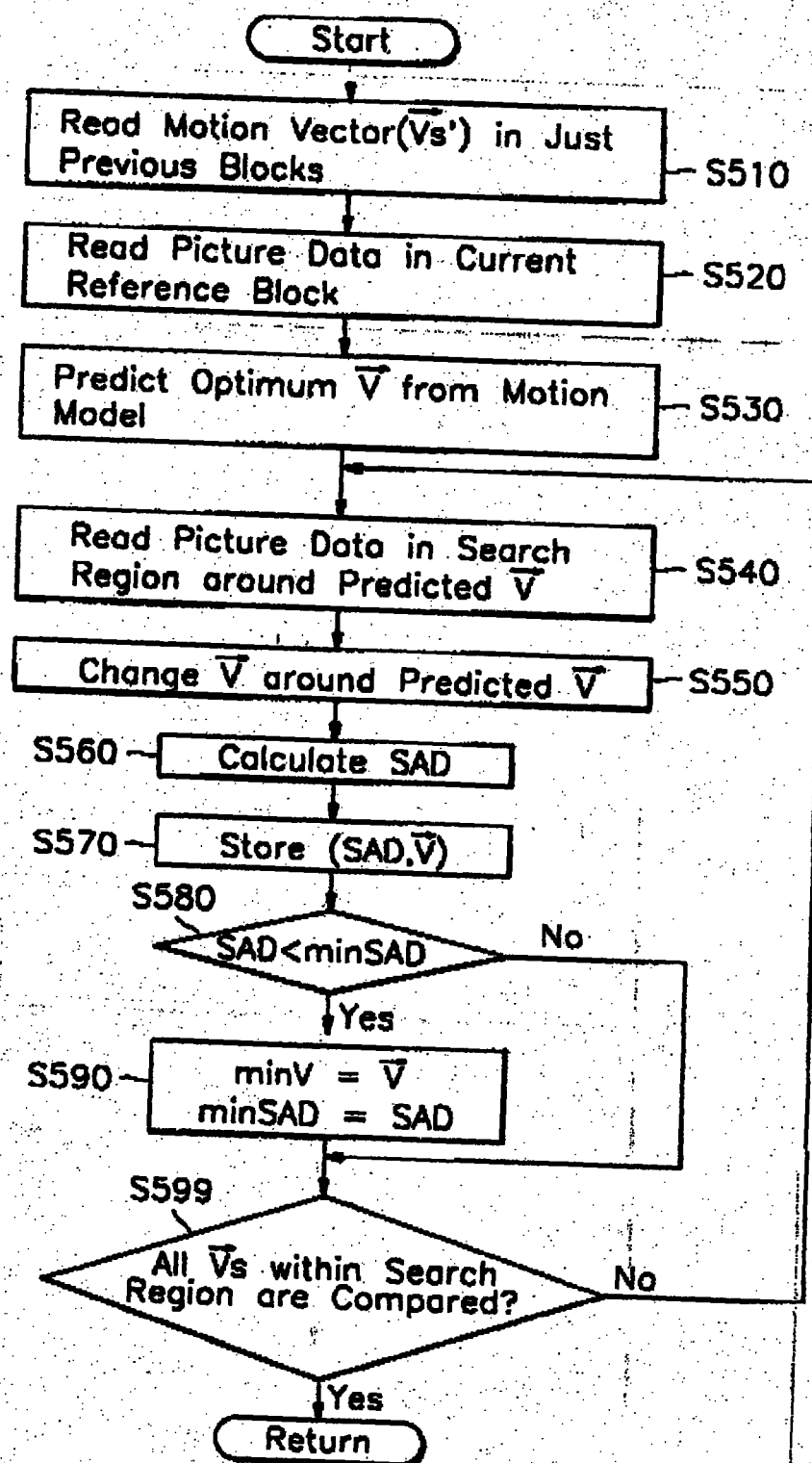
FIG. 7



This Sheet Has Been Replaced



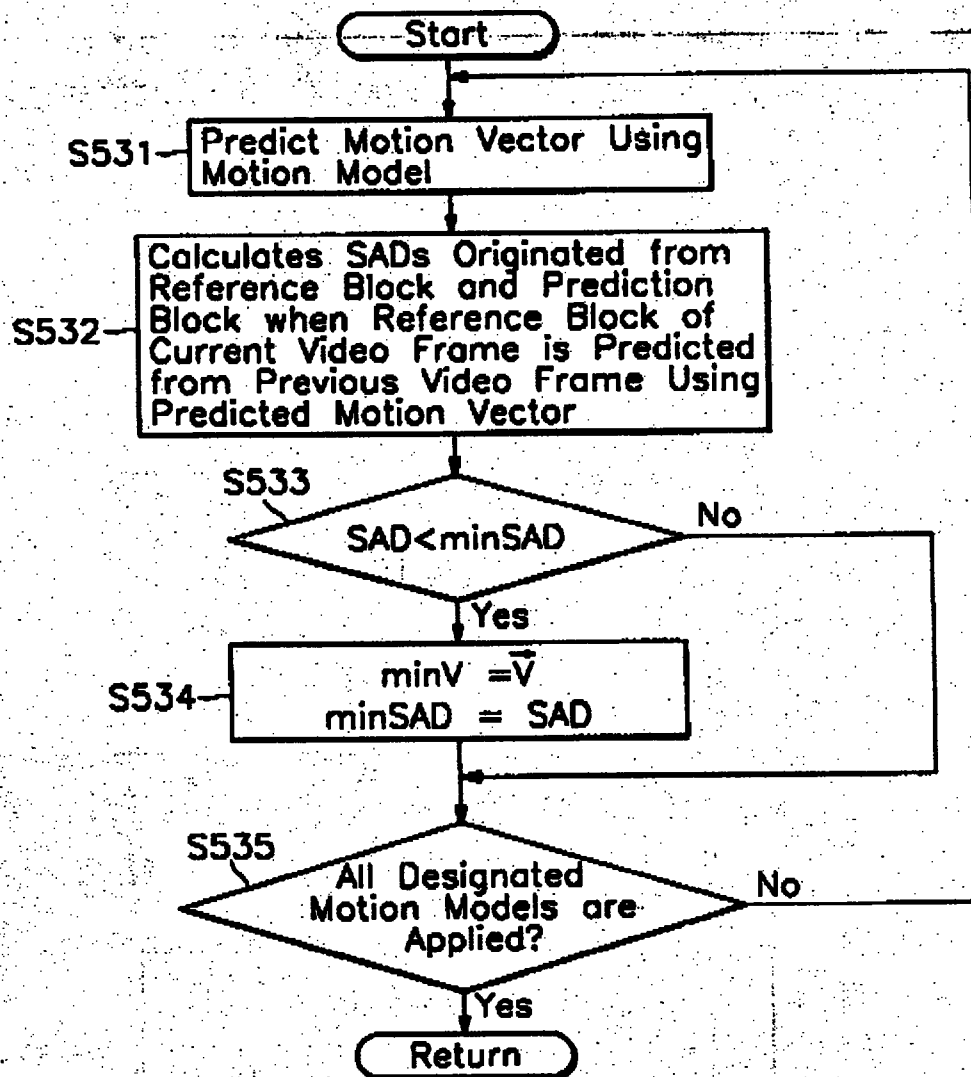
FIG.8



This Sheet Has Been Replaced



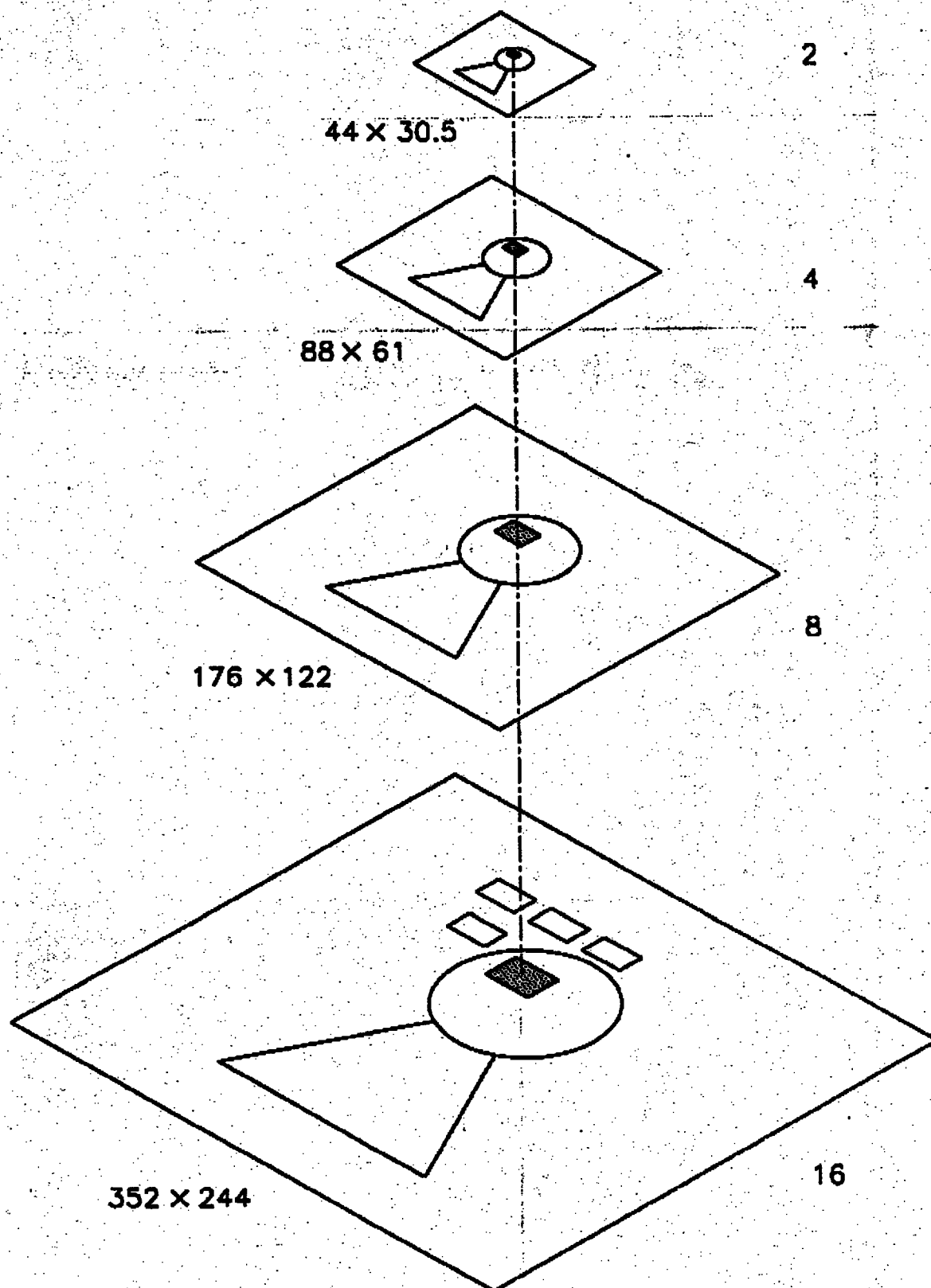
FIG. 9



This Sheet Has Been Replaced



FIG.10



This Sheet Has Been Replaced



REPLACEMENT SHEETS		SG
☒	Description : paged .1. to paged .18.	
☒	Claim(s) : paged .15. to paged .22.	
☒	Drawing(s) : paged .14. to paged .18.	
☒	Abstract : paged .23. to paged .24.	
DATE OF RECEIPT		12/12/91

METHODS OF ESTIMATING MOTION IN IMAGE DATA AND APPARATUS FOR PERFORMING SAME

Field of the Invention

5 The present invention relates to image data methods and apparatus, and more particularly to methods and apparatus for processing full motion image data.

Background of the Invention

10 Video communication which is in widespread use can be classified generally into two categories including analog and digital image communication. Analog image communication has been used for video conferencing, for example, but is generally limited by poor resolution of the transmitted and received image data. In contrast, digital image communication of full motion image data generally has significantly improved fidelity relative to analog image communication in part because digital
15 information can be generally transmitted and received in large quantities at higher data rates than analog image data.

 As will be understood by those skilled in the art, to represent image data digitally, an image area or frame is typically made up of an array of pixels (e.g., 352 rows x 244 lines) with a digital number representing the color, luminance and
20 chrominance of each pixel. Pixel color information typically consists of three digital values including one digital value for red, one for green and one for blue. Thus, a

This is the replacement sheet

Date: - 6 DEC 1996

large quantity of information is needed to describe a pixel which means that digital representations of complete image frames requires exceptionally large quantities of digital data. For example, memory capacity of about three million bytes is required to process color video of 1000x1000 pixels, which means that about ninety (90) million bytes of digital information must be processed every second if a dynamic full motion picture image of 30 frames per second is to be generated. In particular, in full motion video, not only are large blocks of data required to describe each individual picture image, but a new image or frame must be presented to the viewer at approximately thirty new images per second to create the appearance of continuous motion. Moving these large quantities of video image data across digital networks or phone lines is simply infeasible given available bandwidth.

To address this problem, digital data compression has been explored extensively as a technique for reducing the number of bits required to send a given message and have that message received and reproduced with sufficient fidelity. Data compression typically utilizes either a single shorthand notation to signal a repetitive string of bits or omits data bits from the transmitted message. This latter form of compression is typically referred to as "lossy" compression and capitalizes upon on the human mind to provide the omitted data. In still video, the JPEG standard is used for data compression and defines the method by which the still image is to be compressed. In motion video, much of the picture data remains constant from frame to frame. Therefore, the video data may be compressed by first describing a reference frame and describing subsequent frames in terms of the change from the reference frame.

Compression of digital data by performing motion estimation can also be accomplished by comparing video blocks within a previous image frame with video blocks within a present frame (both of which are created by pixel movement in the vertical and/or horizontal directions) and then finding the video block having the smallest difference signal. As will be understood by those skilled in the art, the difference signal can be obtained by determining a sum of absolute difference between blocks. Since it is difficult to compare each block in a large image with blocks from a previous frame, a search area having a predetermined size corresponding to each block

This is the replacement sheet

Date: 6 DEC 1996

is typically set for the previous frame. The motion vector is then determined by searching in the predetermined search area in the present frame. The smaller the size of the block, the more accurate the search for movement is. However, use of a smaller block size typically increases the size of the motion vectors because as the search area increases in size, the larger the movement can be which will be found in the search area and therefore the larger the motion vector.

There are a number of methods for estimating motion. For example, there is the matching method, a gradient method and a block matching method. In particular, FIG. 1 illustrates a conventional 4x4 blocking matching method. Referring to FIG. 1, a screen can include 244 lines and 352 pixels per line which is an international standard for dynamic picture imaging for communication over telephone lines. Each pixel has a corresponding brightness and the brightness can be expressed digitally as a number in the range of 0-255 and motion on the screen can be estimated by evaluating a change in brightness of each pixel. Preferably, in searching motion, a block(s) of pixels is identified and motion is determined by searching according to a change in brightness of the block. This is achieved by assigning motion of a typical pixel in the block to motion of the block on the assumption that the block has only one motion. To illustrate, FIG. 1 illustrates a block of 4x4 or sixteen (16) pixels (Block A).

As will be understood by those skilled in the art, since the motion of the block is two-dimensional, it can be expressed by a two-dimensional motion vector according to Equation 1.

$$V(i,j) = \{V_x(i,j), V_y(i,j)\} \quad (1)$$

In addition to the above methods for estimating motion, there is also a full search method, an M-step search method, a hierarchical search method and a predictive search method. The full search method, which is illustrated by FIG. 2, is the most typically used method for estimating motion. FIG. 3 also illustrates the characteristics of a conventional sum of absolute difference (SAD) graph in the full search method, as known to those skilled in the art. In particular, when a position of an object

corresponds to A at the present time and corresponds to B at a just previous time, a block having the least SAD is determined according to the following equation and the brightness signal of the block is used to estimate the motion of the block A1 of the object, so that the position of the object is estimated based on the position of the block A1.

$$\sum_{x=0}^{N(m)-1} \sum_{y=0}^{M(m)-1} | I_t(x,y,m) - I_{t-1}(x-v_x(k), y-v_y(k), m) |$$

For example, when a search size for searching with the motion vector is the size of ± 1 pixel, nine (9) SADs are obtained and a position having the least value among the nine SADs is used as an estimate of the position to which the object has moved. These nine SADs are: $\{V(x,y): (-1,1), (0,1), (1,1), (-1,0), (0,0), (1,0), (-1,-1), (0,-1), (1,-1)\}$.

However, since the motion may not be properly estimated if the motion is beyond the search range of ± 1 pixel, the search range is typically extended to the size of ± 16 pixels but this significantly increases the number SADs that need to be checked for each of the possible motion vectors in the search range. Accordingly, while the full search method can be effective in dynamic image compression, the amount of calculation that is necessary to obtain an estimate can be excessive, so other search methods have been developed to address this limitation in the full search method. As mentioned above, these other methods include the M-step search method, the hierarchical search method and the predictive search method.

Referring now to FIG. 4A-C, a prior art M-step search method is illustrated. In the M-step search method, motion is estimated by searching with an accuracy of $2^{(M-1)}$ pels and the outskirts of the searched motion are again searched with an accuracy of $2^{(M-2)}$ pels, and then these operations are repeated. For example, when the M-step method is a 3-step method, motion is estimated first by forming the size of the block in 8 pixels as illustrated in FIG. 4A, and next the motion is estimated by forming the size

of the block in 4 pixels as illustrated in FIG. 4B. Next, the motion is estimated by forming the size of the block in 2 pixels as illustrated in FIG. 4C. Accordingly, the motion can be estimated by dividing an estimation step into a series of steps and then gradually estimating the motion in greater detail to achieve a final estimate.

5 Unfortunately, the M-step method has a disadvantage in that if the motion is estimated in error in one step, the motion search may be collected to a local minimum ("lm") as illustrated by FIG. 5.

10 In the case of a hierarchical search method, motion can be searched in a coarser resolution image of 88x61 pixels, which is a reduction from a full resolution image of 352x244 pixels, and then motion is again searched in a finer resolution image of 176x122 pixels which enlarges the outskirts of the searched motion to a greater degree. Thus, if motion is searched in the size of 4 pixels in the least finer image, it has an effect of searching the motion in the size of 16 pixels in the full resolution image. Then, the motion can be estimated by searching in the size of 2 pixels in the next finer
15 resolution image which has the effect of searching motion in the size of 4 pixels in the full resolution image. As will be understood by those skilled in the art, these steps minimize the amount of calculation necessary to estimate motion in the full resolution image. Unfortunately, the hierarchical search method has a disadvantage in that if the motion is estimated in error in the coarser resolution image, the motion search may be
20 collected to a local minimum point ("lm").

In the predictive search method, motion in the reference block to be searched is predicted from the motion of circumferential blocks and the outskirts around a prediction value are searched. However, the predictive search method has a disadvantage in that if the prediction value is not accurate, the motion search may be
25 collected to a local minimum point ("lm").

In the above described methods, a motion vector field is used as a prediction model and a 2-D spatial model and 3-D spatial model based on a coarser resolution image are also used. In the 3-D spatial model, the motion in the current reference block is effectively predicted by effectively combining information about the motion
30 vector field in the coarser resolution image and information about the motion vector

field in the circumferential blocks. The 3-D spatial model also includes multiple elementary models to effectively treat motion discontinuity. The 3-D spatial model is more fully described in a doctoral dissertation entitled "3-D Kalman Filter for Video and Motion Estimation", by coinventor Jae-Min Kim, Rensselaer Polytechnic Institute (1994), the disclosure of which is hereby incorporated herein by reference. However, the 3-D spatial model has disadvantages in that the model parameter set is very complex and because it is difficult to implement the motion estimation since the algorithm is complex as well and the amount of necessary calculation is large when a Kalman filter is used to estimate motion.

Thus, notwithstanding the above described methods of estimating motion in image data, there still continues to be a need for methods and apparatus for efficiently processing full motion image data in compressed format with reduced error.

Summary of the Invention

It is therefore an object of the present invention to provide improved methods and apparatus for processing full motion image data.

It is further object of the present invention to provide methods and apparatus for processing full motion image data in compressed format with reduced error.

It is another object of the present invention to provide methods and apparatus for processing and communicating full motion image data using reduced bandwidth.

These and other objects, features and advantages of the present invention are provided by methods and apparatus for estimating motion in video data so that efficient communication of video data in real time with high fidelity is feasible. Using this method, full motion image data is processed in 3-dimensions using multiple models and by adjusting a search region based on prediction errors from the models to more accurately estimate motion. In particular, according to a preferred embodiment of the present invention, a step is performed to produce a first frame of image data from a received video signal containing information describing an image data frame of predetermined resolution (e.g., 352x244 pixels). However, the first frame of image data is of coarser resolution (e.g., 176x122 pixels) relative to the image data frame of predetermined resolution. In addition, a second image data frame of coarser resolution

(e.g., 88x61 pixels) relative to the first image data frame is also produced so that a more efficient routine for detecting motion can be achieved using 3-dimensional processing.

Once the images of reduced resolution have been produced, an estimate of motion in the second image data frame is then produced and used in the estimation of motion in the first image data frame of higher resolution. The estimation of motion in the first image data frame is then used to obtain an estimation of motion in the image data frame of predetermined resolution. In particular, an estimate of motion in the second image data frame, for example, can be determined by reading a plurality of motion vectors (e.g. nine for a search size of ± 1 pixel) from an image block in a previous version of the second image data frame and reading image data in a reference block in a current version of the second image data frame. An optimum motion vector is also determined by predicting a plurality of motion vectors based on a corresponding plurality of motion models and selecting the best model and vector which yields the smallest sum of absolute difference error. Once an optimum motion vector has been predicted, the plurality of motion vectors previously read are adjusted relative to the optimum motion vector and then a determination is made as to which vector yields the smallest relative motion error so that this vector can be used as an estimation of motion from the previous version of the second image data frame to the current version of the second image data frame. These steps are then performed again for the first image data frame of higher resolution, using the estimation obtained from the second image data frame, and then performed again so that a highly accurate estimation of motion in the image data frame of predetermined resolution can be achieved. The present invention also includes apparatus for performing the above described operations.

Brief Description of the Drawings

FIG. 1 illustrates a pictorial representation of a prior art 4x4 block matching method of motion estimation.

FIG. 2 illustrates a pictorial representation of a prior art full search method of motion estimation.

This is the replacement sheet

Date: - 6 DEC 1996

FIG. 3 illustrates a sum of absolute difference graph for a full search method of motion estimation.

FIG. 4 illustrates a pictorial representation of a prior art 3-step search method of motion estimation.

5 FIG. 5 illustrates a sum of absolute difference graph including local minimums ("lm") for a M-step search method of motion estimation.

FIG. 6 is a block diagram illustrating an apparatus for performing motion estimation according to the present invention.

10 FIG. 7 is a flow chart illustrating operations for estimating motion in image data according to an embodiment of the present invention.

FIG. 8 is a flow chart illustrating operations for estimating motion in an image data frame of full or relatively coarse resolution.

FIG. 9 is a flow chart illustrating operations for predicting an optimum motion vector based on a plurality of models.

15 FIG. 10 illustrates a pictorial representation showing the full, first, second and third resolution frames used to determine a motion estimation according to the present invention.

Description of Preferred Embodiments

20 The present invention will now be described more fully hereinafter with reference to the accompanying drawings, in which preferred embodiments of the invention are shown. This invention may, however, be embodied in different forms and should not be construed as limited to the embodiments set forth herein. Rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art. Like
25 number refer to like elements throughout.

Referring now to FIGS. 6-10, the present invention will be described. In particular, FIG. 6 illustrates a preferred hardware configuration for an apparatus for estimating motion in image data according to the present invention. The apparatus may include an input/output unit 10 for receiving an external video signal containing image
30 data and transmitting an estimated motion signal determined by the operations described

herein. A search memory 20 is also provided for storing the received external video signal and results obtained during intermediate stages of the preferred operations for performing motion estimation according to the present invention. A microcontroller 30 is also provided for producing, among other things, image data frames of multiple resolution relative to an image data frame of predetermined full resolution and a signal representing an estimate of motion in the received image data. A block matching circuit 50, which is responsive to the microcontroller 30, is provided for performing block matching of data preferably stored in the search memory 20. In addition, an interpolation circuit 40 for interpolating a value of the data stored in the search memory 20 in response to a signal from the microcontroller 30 and producing an interpolated signal, is provided.

Information of a general nature relating to other apparatus for performing motion estimation can be found in U.S. Patent Nos. 5,365,281, entitled *Motion Signal Detecting Circuit*; 5,392,073 entitled *Motion Estimating Device For Variable Length Encoded Data*; 5,467,086 entitled *Apparatus and Method of Coding/Decoding Video Data*; 5,489,949 entitled *Method and Apparatus for Motion Estimation*; 5,497,153 entitled *System for Variable-Length-Coding and Variable-Length-Decoding Digital Data for Compressing Transmission Data*; 5,493,344 entitled *Motion Vector Estimation Method and Apparatus Having Half-Pixel Accuracy*; 5,508,745 entitled *Apparatus for Controlling A Quantization Level to be Modified by a Motion Vector*; and 5,510,857 entitled *Motion Estimation Coprocessor*, all of which are assigned to the assignee of the present application and hereby incorporated herein by reference.

Referring now to operations S100-S900 of FIG. 7, the microcontroller 30 reads the external video signal received by the input/output unit 10, operation S100. This signal contains information to construct an image data frame of predetermined full resolution (e.g., 352x244 pixels). Referring now to operation S200, the microcontroller 30 performs the operation of producing a first image data frame from the image data frame of predetermined resolution, so that the first image data frame is of coarser resolution (e.g., 176x122 pixels), as illustrated best by FIG. 10. This

This is the replacement sheet

Date: - 6 DEC 1996

operation is then followed by a step of storing information relating to the first image data frame in the search memory 20. In operation S300, a second image data frame is produced so that the second image data frame is of coarser resolution (e.g., 88x61 pixels) relative to the first image data frame and then this frame is stored in search memory 20. Next, in operation S400, a third image data frame is produced so that the third image data frame is of coarser resolution (e.g., 44x30.5 pixels) relative to the second image data frame and then this frame is also stored in search memory 20.

Referring still to FIG. 7, the microcontroller 30 then estimates motion after a predetermined amount of time from the third image data frame to obtain a first estimation, as illustrated by operation S500. This first estimation is then stored in search memory 20. Next, the microcontroller 30 estimates motion from the second image data frame on the basis of the motion estimated in operation S500, to obtain a second estimation which is then stored in search memory 20, as illustrated by operation S600. Following this step, the microcontroller 30 estimates motion from the first image data frame on the basis of the motion estimated in operation S600, to obtain a third estimation which is stored, as illustrated by operation S700. Finally, the microcontroller 30 estimates motion in the image data frame of predetermined full resolution on the basis of the motion estimated in operation S700 and then this estimate is stored, as illustrated by operation S800. The microcontroller 30 then causes the production of a video signal containing the motion estimation corresponding to the full resolution image data frame for transmission by the input/output unit 10, operation S900.

Referring now to FIG. 8, the operation S500 for estimating motion in the third data frame to obtain an estimation of motion will be described, however the steps performed by this preferred operation are also applicable to operations S600-S800. In operation S500, the microcontroller 30 reads the motion vectors (V_s) of the blocks from a just previous image data frame (S510) and reads image data in a reference block in a current image data frame (S520). During operation S500, a optimum motion vector is also predicted from a motion model (S530) and then image data in a search

region within the current image data frame ¹¹ and around the predicted motion vector, is read (S540). This reading step (S540) is then followed by the step S550 of adjusting or changing the motion vectors obtained from (S510), around or relative to the predicted motion vector obtained from (S530).

5 As illustrated by steps S560-S599, a sum of absolute difference (SAD) between the reference block and block of image data at an adjusted motion vector is determined for each of the adjusted motion vectors obtained from (S550) and the lowest determined SAD and the adjusted motion vector yielding the lowest determined SAD are stored in memory, for use during subsequent operations.

10 Referring now to FIG. 9, the operation S530 for predicting an optimum motion vector according to the present invention will be described. Using this operation S530, the search region can be preferably adjusted based on prediction errors derived by analyzing multiple motion models and minimizing these errors by using a feedback loop to determine the best model to most accurately estimate motion using the steps of FIG.
15 8. In particular, at step S531, a motion vector is predicted using a first motion model and then at step S532, a sum of absolute difference is determined between a reference block and a prediction block when the reference block of a current frame is predicted from a previous video frame using the predicted motion vector based on the first motion model. However, rather than using only one model, a plurality of different
20 models are used to reduce the likelihood of occurrence of a local minimum ("lm") in the difference graph of FIG. 5, which can be problematic using the methods of the prior art. For example, steps S533-S535 illustrate that the model yielding the motion vector corresponding to the lowest SAD, when step S532 is performed, will provide an optimum motion vector to be used in step S530.

25 As illustrated below with respect to Table 1, when compared to the full search method and hierarchical search method described in the background section of the present application, the present invention is preferable to both methods based on the average value of the mean square error (MSE) of temporal interpolation in thirty(30) frames, when a flower garden sequence and mobile sequence are used.

30

This is the replacement sheet

Date: - 6 DEC 1996

	<u>Flower Garden Sequence</u>	<u>Mobile Sequence</u>
	Full Search 278	361
5	Hierarchical Search 282	273
	Present Invention 232	239

10

Table 1

In addition, Table 2 illustrates an average value of the mean square error (MSE) of motion compensated prediction in thirty (30) frames, using the above search methods. Based on the MSE of motion compensated prediction, the full search method is the most effective, but the method according to the present invention is almost as good as the full search method.

15

	<u>Flower Garden Sequence</u>	<u>Mobile Sequence</u>
	Full Search 226	296
20	Hierarchical Search 338	372
	Present Invention 245	303

25

Table 2

Referring now to Table 3, the times needed to calculate the final motion estimation by each method is illustrated. Based on time, the hierarchical search method is the most effective, but the method according to the present invention is almost as efficient based on time.

30

	<u>Flower Garden Sequence</u>	<u>Mobile Sequence</u>
	Full Search 105 sec.	105 sec.
35	Hierarchical Search 6.3 sec.	6.3 sec.

This is the replacement sheet

Date: 26 DEC 1998

5

Table 3

10

15

20

As described above, the microcontroller 30, in conjunction with the other structural elements of FIG. 6, is preferably designed to perform the above operations of FIGS. 7-9. Accordingly, an apparatus for compressing motion video data by motion estimation according to the present invention includes means for receiving a video signal containing information to construct an image data frame of predetermined resolution; means, coupled to said receiving means, for producing information to construct a first image data frame of coarser resolution relative to the image data frame of predetermined resolution and information to construct a second image data frame of coarser resolution relative to the first image data frame; means, coupled to said information producing means, for estimating motion in the second image data frame to obtain a first estimation and for estimating motion in the first image data frame, based on the first estimation, to obtain a second estimation; and means, responsive to said estimating means, for transmitting a video signal containing an estimate of motion occurring in the image data frame of predetermined resolution, based on the second estimation.

25

30

According to the present invention, the apparatus for estimating motion preferably includes means for reading a plurality of motion vectors and reading image data in a reference block in the second image data frame; means for predicting an optimum motion vector based on a motion model; means, responsive to said reading means and predicting means, for reading image data in a first search region in the second image data frame, adjacent the predicted optimum motion vector; means for adjusting one of said plurality of motion vectors relative to the predicted optimum motion vector, to obtain a first motion vector; and means, responsive to said adjusting means, for determining a first sum of absolute difference between the reference block and a block of image data at the first motion vector.

Finally, the apparatus for predicting an optimum motion vector preferably includes means for predicting a motion vector using a first model; means for determining a sum of absolute difference between a reference block and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the first model; means for predicting a motion vector using a second model; means for determining a sum of absolute difference between a reference block and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the second model; and means, responsive to said predicting and determining means, for determining the predicted motion vector based on the first model as an optimum motion vector, if the sum of absolute difference determined using the predicted motion vector based on the first model is less than the sum of absolute difference determined using the predicted motion vector based on the second model.

In the drawings and specification, there have been disclosed typical preferred embodiments of the invention and, although specific terms are employed, they are used in a generic and descriptive sense only and not for purposes of limitation, the scope of the invention being set forth in the following claims.

This is the replacement sheet

Date: 6 DEC 1996

THAT WHICH IS CLAIMED IS:

1. A method of estimating motion in image data, comprising the steps of:
producing a first image data frame from an image data frame of predetermined resolution, so that the first image data frame is of coarser resolution relative to the image data frame of predetermined resolution;
producing a second image data frame from the image data frame of predetermined resolution or the first image data frame, so that the second image data frame is of coarser resolution relative to the first image data frame;
estimating motion in the second image data frame to obtain a first estimation;
estimating motion in the first image data frame based on the first estimation, to obtain a second estimation; and
estimating motion in the image data frame of predetermined resolution based on the second estimation.
2. The method of Claim 1, wherein said step of estimating motion in the second image data frame comprises the steps of:
predicting a motion vector based on a motion model;
reading image data in a first search region within the second image data frame, based on the predicted motion vector;
determining a first motion vector based on the predicted motion vector; and
determining a sum of absolute difference based on the first motion vector.
3. The method of Claim 1, wherein said step of estimating motion in the second image data frame comprises the steps of:
reading a plurality of motion vectors;
reading image data in a reference block in the second image data frame;
predicting an optimum motion vector based on a motion model;
reading image data in a first search region adjacent the predicted optimum motion vector;



This is the replacement sheet

Date: 6 DEC 1996

adjusting one of said plurality of motion vectors relative to the predicted optimum motion vector, to obtain a first motion vector; and

determining a first sum of absolute difference between the reference block and a block of image data at the first motion vector.

4. The method of Claim 3, wherein said step of reading image data in a first search region adjacent the predicted optimum motion vector comprises reading image data in a first search region surrounding the predicted optimum motion vector.

5. The method of Claim 3, further comprising the steps of:

adjusting another of said plurality of motion vectors relative to the predicted optimum motion vector, to obtain a second motion vector;

determining a second sum of absolute difference between the reference block and a block of image data at the second motion vector; and

5 determining the first motion vector as an estimate of motion in the second image data frame instead of the second motion vector, if the first sum of absolute difference is less than the second sum of absolute difference.

6. The method of Claim 3, wherein said step of predicting an optimum motion
10 vector based on a motion model comprises the steps of:

predicting a motion vector using a model; and

determining a sum of absolute difference between a reference block and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector.

15

7. The method of Claim 6, wherein said step of predicting a motion vector using a model comprises predicting a plurality of motion vectors using a corresponding plurality of models.



* C L H S *

This is the replacement sheet

Date: - 6 DEC 1996

8. The method of Claim 7, further comprising the steps of determining a plurality of sums of absolute difference between the reference block and a plurality of prediction blocks in the second image data frame which are determined from the reference block and corresponding ones of the plurality of motion vectors.0

5

9. The method of Claim 8, further comprising the step of determining the lowest value of the plurality sums of absolute difference.

10. The method of Claim 3, wherein said step of predicting an optimum motion vector based on a motion model comprises the steps of:

predicting a motion vector using a first model;

determining a sum of absolute difference between a reference block and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the first model;

predicting a motion vector using a second model;

determining a sum of absolute difference between a reference block and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the second model; and

determining the predicted motion vector based on the first model as an optimum motion vector, if the sum of absolute difference determined using the predicted motion vector based on the first model is less than the sum of absolute difference determined using the predicted motion vector based on the second model.

25

11. A method of estimating motion in image data, comprising the steps of:

producing a first image data frame from an image data frame of predetermined resolution, so that the first image data frame is of coarser resolution relative to the image data frame of predetermined resolution;



This is the replacement sheet

Date: - 6 DEC 1996

producing a second image data frame from the image data frame of predetermined resolution or the first image data frame, so that the second image data frame is of coarser resolution relative to the first image data frame;

producing a third image data frame from the image data frame of predetermined resolution, the first image data frame or the second image data frame, so that the third image data frame is of coarser resolution relative to the second image data frame;

estimating motion in the third image data frame to obtain a first estimation;

estimating motion in the second image data frame based on the first estimation, to obtain a second estimation;

estimating motion in the first image data frame based on the second estimation, to obtain a third estimation; and

estimating motion in the image data frame of predetermined resolution based on the third estimation.

12. The method of Claim 11, wherein said step of estimating motion in the third image data frame comprises the steps of:

predicting a motion vector based on a motion model;

reading image data in a first search region within the third image data frame, based on the predicted motion vector;

determining a first motion vector based on the predicted motion vector; and determining a sum of absolute difference based on the first motion vector.

13. The method of Claim 11, wherein said step of estimating motion in the third image data frame comprises the steps of:

reading a plurality of motion vectors;

reading image data in a reference block in the third image data frame;

predicting an optimum motion vector based on a motion model;

reading image data in a first search region adjacent the predicted optimum motion vector;

adjusting one of said plurality of motion vectors relative to the predicted optimum motion vector, to obtain a first motion vector; and

This is the replacement sheet



Date: 6 DEC 1996

determining a first sum of absolute difference between the reference block and a block of image data at the first motion vector.

14. The method of Claim 13, wherein said step of reading image data in a first
5 search region adjacent the predicted optimum motion vector comprises reading image data in a first search region surrounding the predicted optimum motion vector.

15. The method of Claim 13, further comprising the steps of:
adjusting another of said plurality of motion vectors relative to the predicted
10 optimum motion vector, to obtain a second motion vector;
determining a second sum of absolute difference between the reference block and a block of image data at the second motion vector; and
determining the first motion vector as an estimate of motion in the third image data frame instead of the second motion vector, if the first sum of absolute difference is
15 less than the second sum of absolute difference.

16. The method of Claim 13, wherein said step of predicting an optimum motion vector based on a motion model comprises the steps of:
predicting a motion vector using a model; and
20 determining a sum of absolute difference between a reference block and a prediction block in the third image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector.

17. The method of Claim 16, wherein said step of predicting a motion vector
25 using a model comprises predicting a plurality of motion vectors using a corresponding plurality of models.

18. The method of Claim 17, further comprising the steps of determining a plurality of sums of absolute difference between the reference block and a plurality of



* C L N S *

This is the replacement sheet

Date: - 6 DEC 1996

prediction blocks in the third image data frame which are determined from the reference block and corresponding ones of the plurality of motion vectors.

19. The method of Claim 18, further comprising the step of determining the
5 lowest value of the plurality sums of absolute difference.

20. The method of Claim 13, wherein said step of predicting an optimum motion vector based on a motion model comprises the steps of:

predicting a motion vector using a first model;
10 determining a sum of absolute difference between a reference block and a prediction block in the third image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the first model;

predicting a motion vector using a second model;
15 determining a sum of absolute difference between a reference block and a prediction block in the third image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the second model; and

determining the predicted motion vector based on the first model as an optimum
20 motion vector, if the sum of absolute difference determined using the predicted motion vector based on the first model is less than the sum of absolute difference determined using the predicted motion vector based on the second model.

21. An apparatus for compressing motion video data by motion estimation,
25 comprising:

means for receiving a video signal containing information to construct an image data frame of predetermined resolution;

means, coupled to said receiving means, for producing information to construct a first image data frame of coarser resolution relative to the image data frame of



* C L M S *

This is the replacement sheet

Date: - 6 DEC 1996

predetermined resolution and information to construct a second image data frame of coarser resolution relative to the first image data frame;

means, coupled to said information producing means, for estimating motion in the second image data frame to obtain a first estimation and for estimating motion in
 5 the first image data frame, based on the first estimation, to obtain a second estimation; and

means, responsive to said estimating means, for transmitting a video signal containing an estimate of motion occurring in the image data frame of predetermined resolution, based on the second estimation.

10

22. The apparatus of Claim 21, wherein said means for estimating motion comprises:

means for reading a plurality of motion vectors and reading image data in a reference block in the second image data frame;

15 means for predicting an optimum motion vector based on a motion model;

means, responsive to said reading means and predicting means, for reading image data in a first search region in the second image data frame, adjacent the predicted optimum motion vector;

means for adjusting one of said plurality of motion vectors relative to the
 20 predicted optimum motion vector, to obtain a first motion vector; and

means, responsive to said adjusting means, for determining a first sum of absolute difference between the reference block and a block of image data at the first motion vector.

25 23. The apparatus of Claim 22, wherein said means for predicting an optimum motion vector comprises:

means for predicting a motion vector using a model; and

means for determining a sum of absolute difference between a reference block and a prediction block in the second image data frame when the reference block is
 30 predicted from a previous image data frame using the predicted motion vector.



* C L M S *

This is the replacement sheet

Date: - 6 DEC 1996

24. The apparatus of Claim 22, wherein said means for predicting an optimum motion vector comprises:

means for predicting a motion vector using a first model;

means for determining a sum of absolute difference between a reference block

5 and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the first model;

means for predicting a motion vector using a second model;

means for determining a sum of absolute difference between a reference block

10 and a prediction block in the second image data frame when the reference block is predicted from a previous image data frame using the predicted motion vector based on the second model; and

means, responsive to said predicting and determining means, for determining the predicted motion vector based on the first model as an optimum motion vector, if the
15 sum of absolute difference determined using the predicted motion vector based on the first model is less than the sum of absolute difference determined using the predicted motion vector based on the second model.



* C L H S *

This is the replacement sheet

Date: - 6 DEC 1996

METHODS OF ESTIMATING MOTION IN IMAGE DATA
AND APPARATUS FOR PERFORMING SAME

Abstract of the Disclosure

5 Methods of estimating motion in image data include the steps of producing a first frame of image data from a received video signal containing information of an image data frame of higher predetermined resolution (e.g., 352x244 pixels) and producing a second image data frame of coarser resolution (e.g., 88x61 pixels) relative to the first image data frame. An estimate of motion in the second image data frame is then produced and used to determine a estimate of motion in the first image data frame of higher resolution. The estimate of motion in the first image data frame is then used to obtain an estimate of motion in the received image data frame of predetermined full resolution. The estimate of motion in the second image data frame can be determined by reading a plurality of motion vectors from an image block in a previous version of the second image data frame and reading image data in a reference block in a current version of the second image data frame. To obtain this estimate, an optimum motion vector is determined by predicting a plurality of motion vectors based on a corresponding plurality of motion models and selecting the best model and vector which yields the smallest difference signal. Once an optimum motion vector has been predicted, the plurality of motion vectors previously read are adjusted relative to the optimum motion vector. Then a determination is made as to which one of the plurality of vectors yields the smallest relative motion error so that this vector can be used as an estimation of motion from the previous version of the second image data frame to the current version of the second image data frame. These steps are then performed again for the first image data frame of higher resolution, using the estimation obtained from the second image data frame, and then performed again so that a highly accurate estimation of motion in the image data frame of predetermined resolution can be achieved.

#24613

30

This is the replacement sheet

Date: - 6 DEC 1996

FIG. 1 ¹/₁ Prior Art

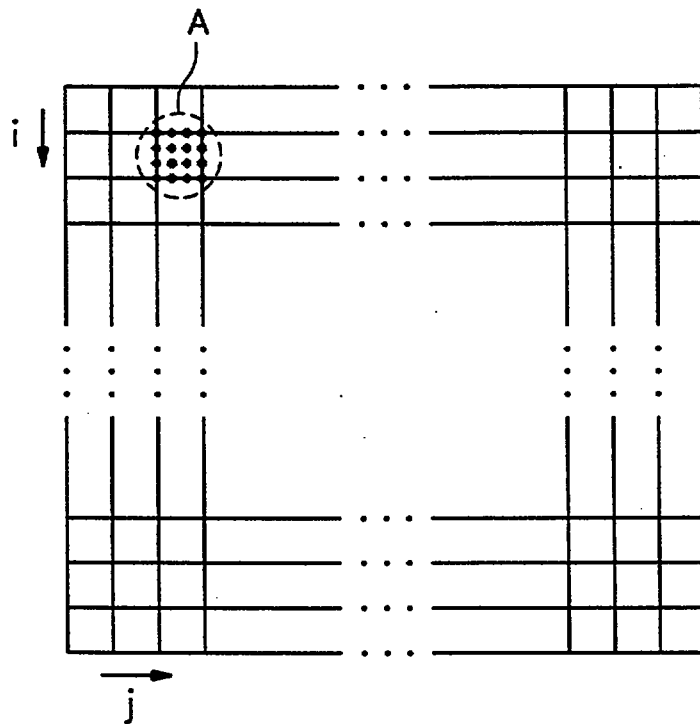
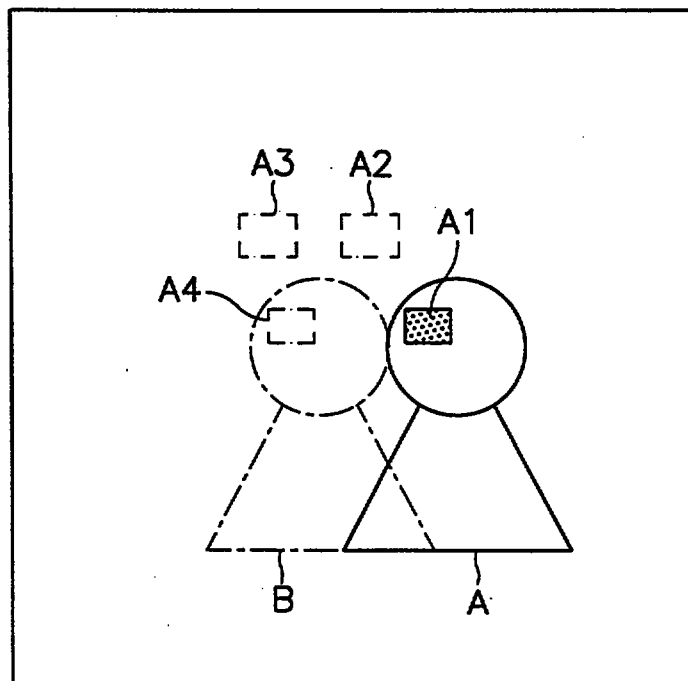


FIG. 2 Prior Art



This is the replacement sheet

Date: - 6 DEC 1996

FIG.3 Prior Art

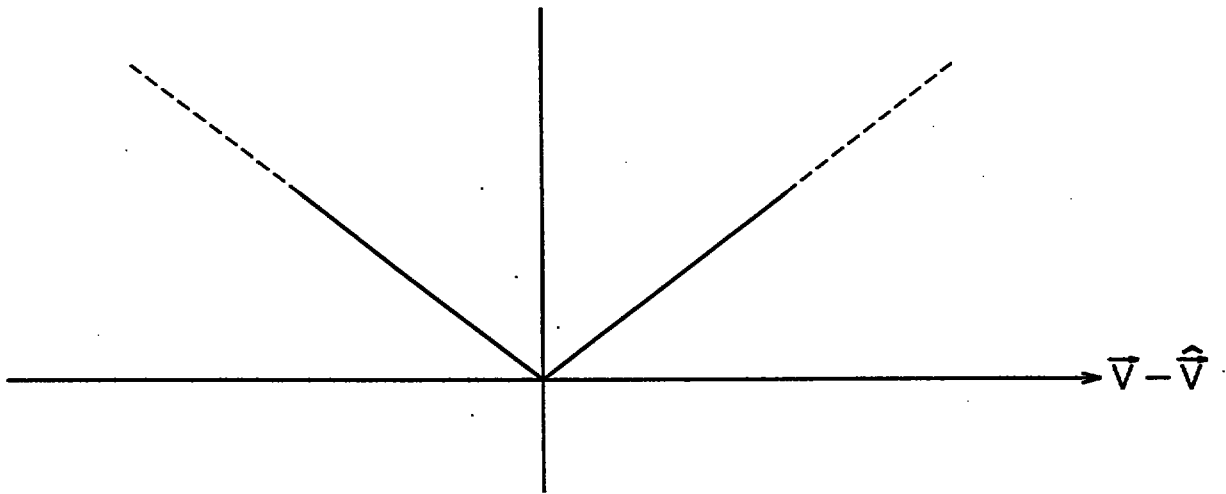
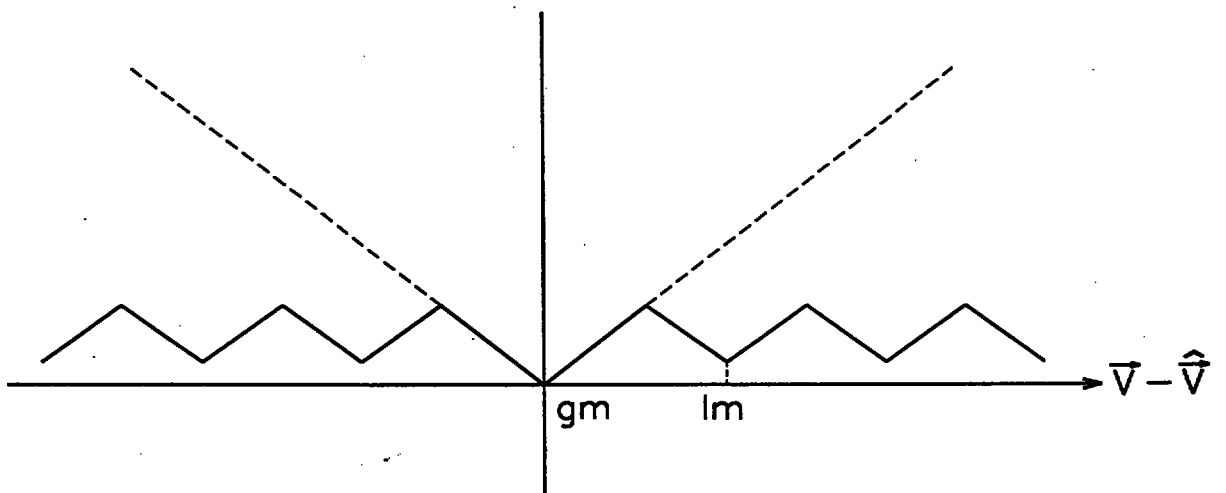


FIG.5 Prior Art



This is the replacement sheet

Date: - 6 DEC 1996

FIG.4A Prior Art

8pixels

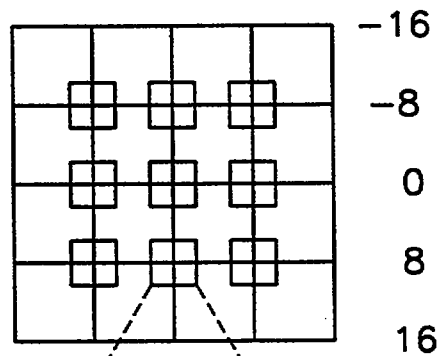


FIG.4B Prior Art

4pixels

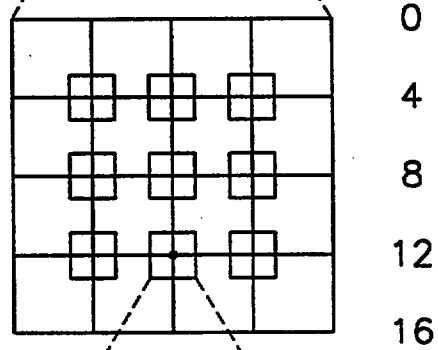
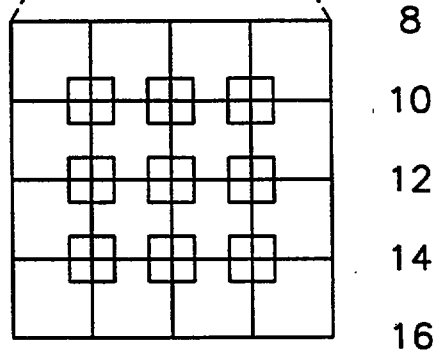


FIG.4C Prior Art

2pixels



This is the replacement sheet

Date: - 6 DEC 1996

FIG. 6

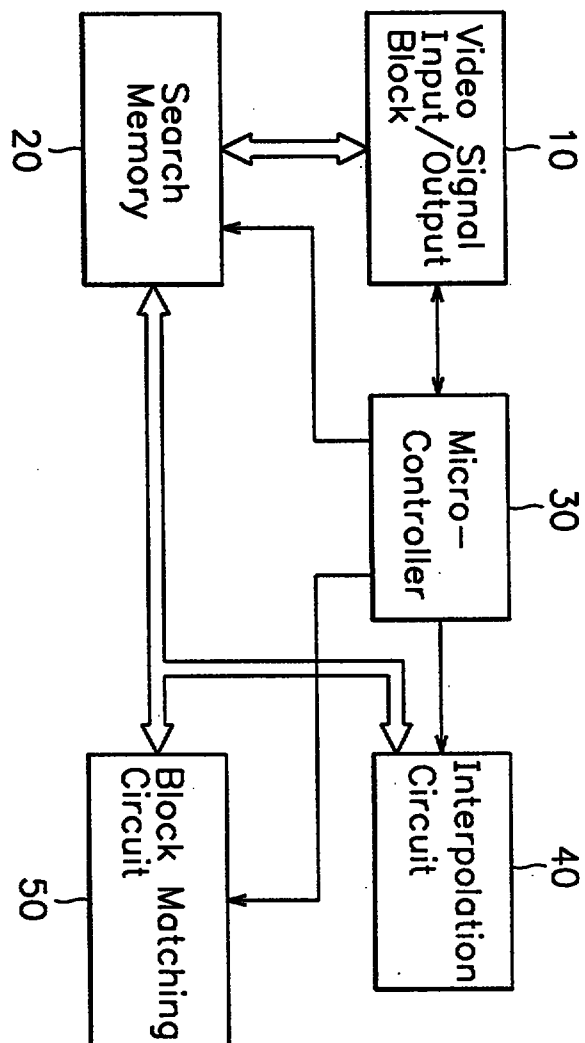


FIG.7

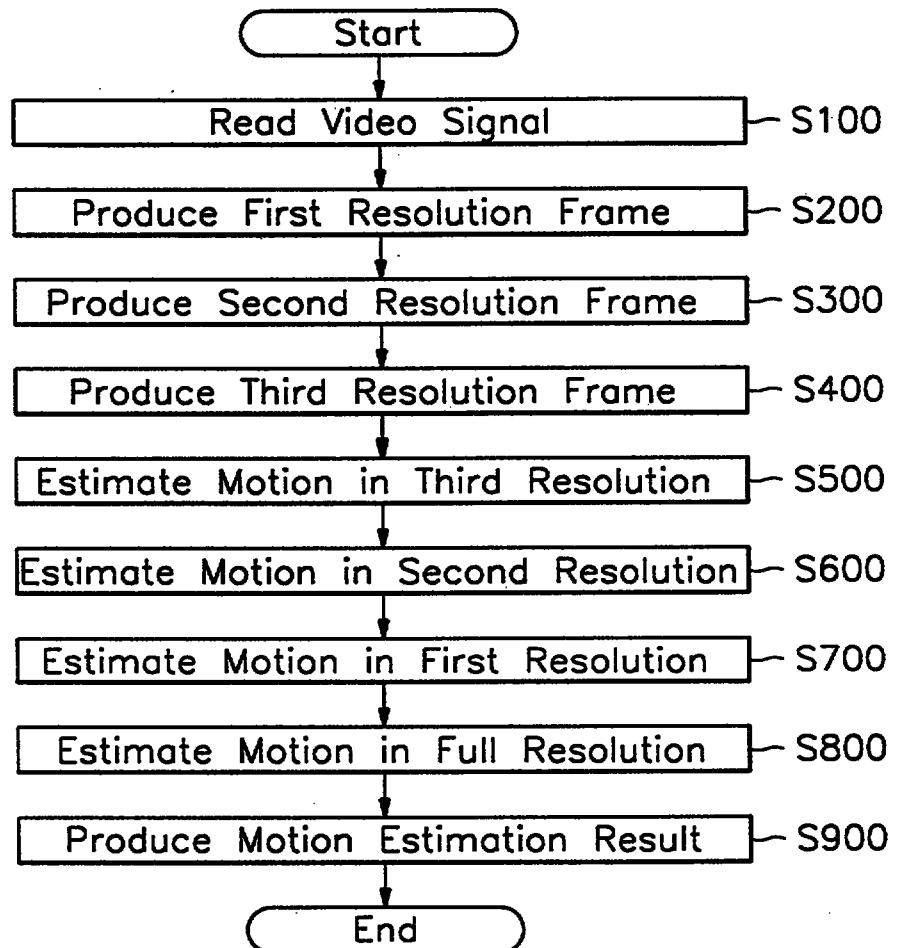


FIG.8

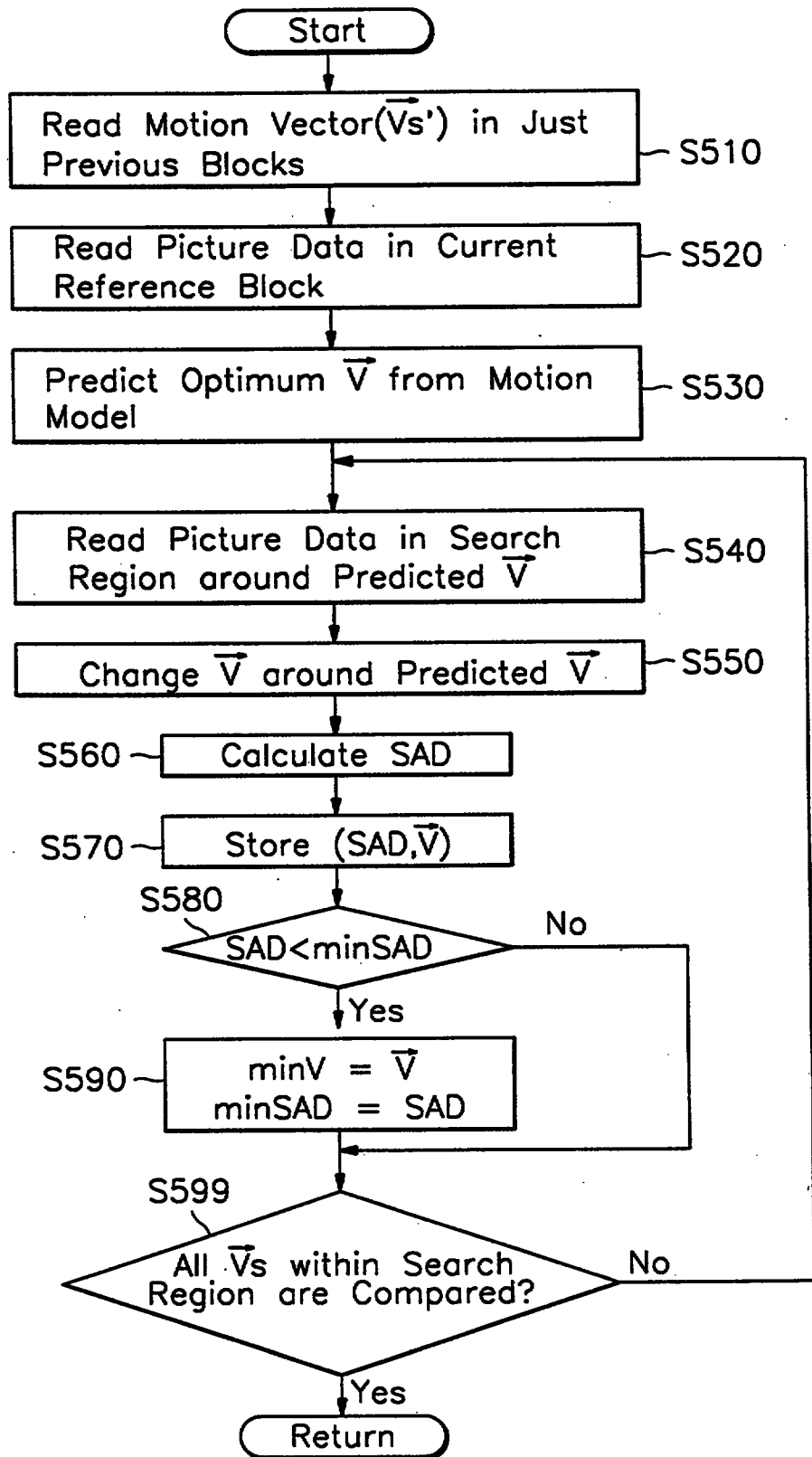
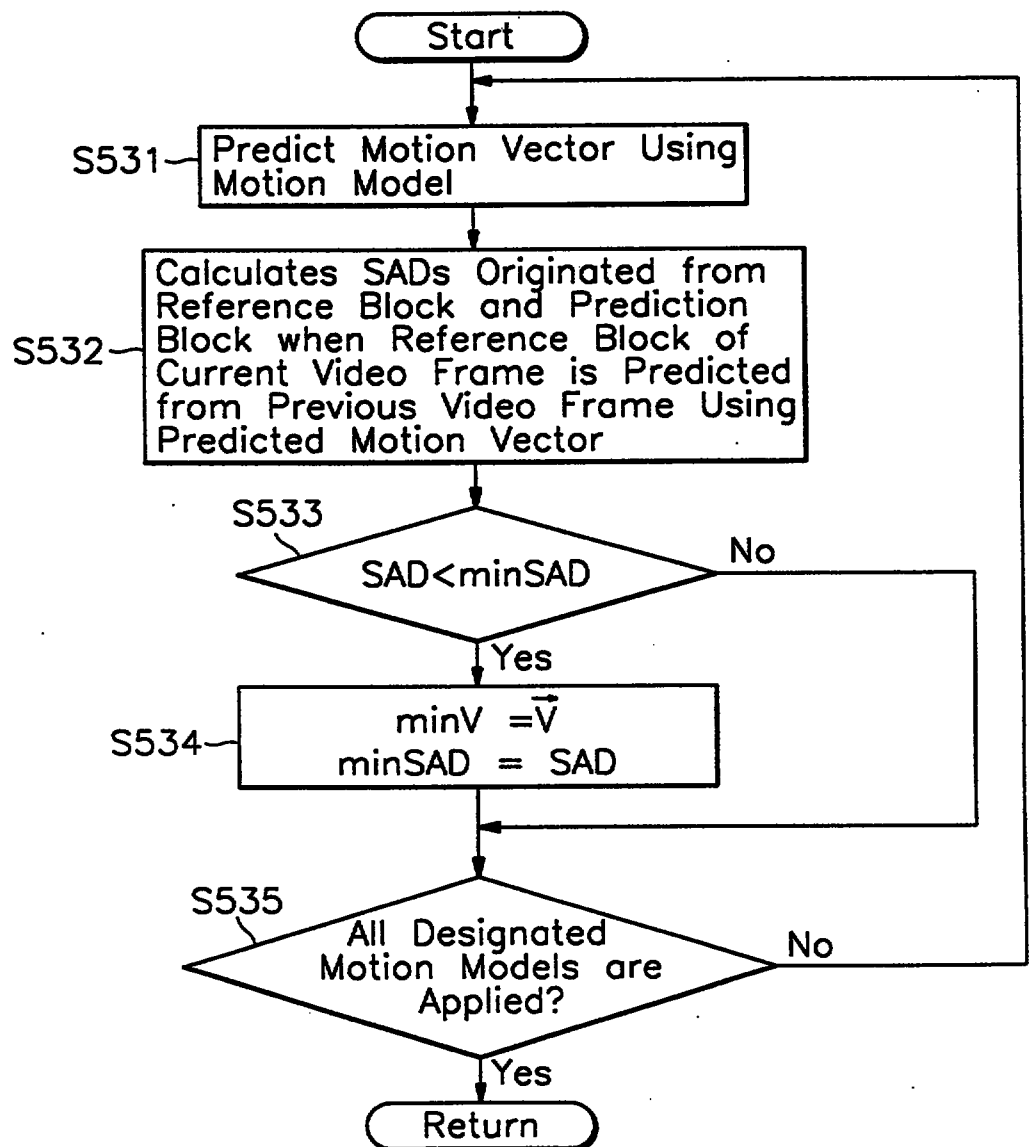
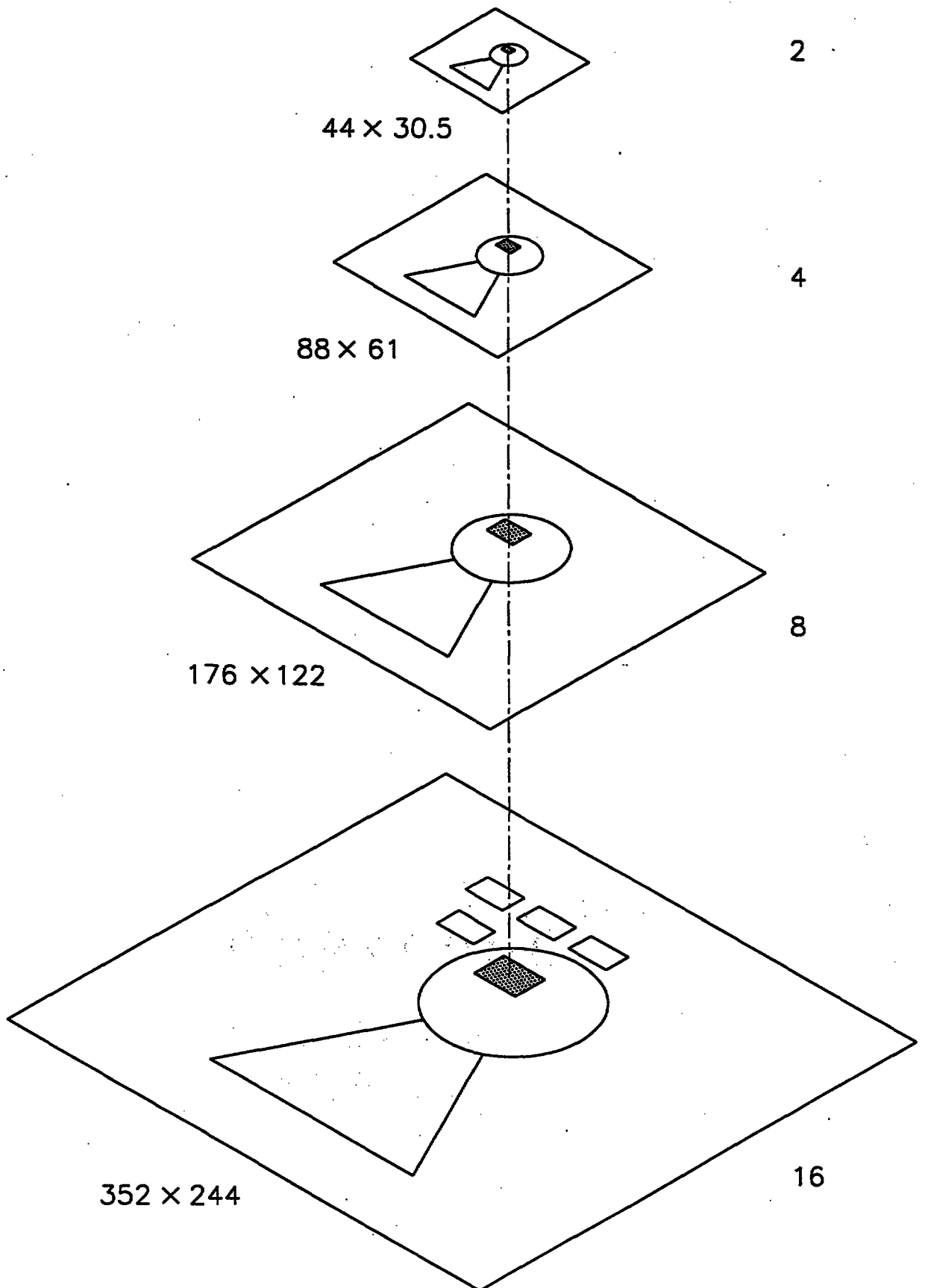


FIG.9



8 FIG.10



This is the replacement sheet

Date: - 6 DEC 1996