(12) **PATENT APPLICATION** (11) Application No. **AU 200037921 A1**
(19) **AUSTRALIAN PATENT OFFICE**

(54) Title
Method and apparatus for improved video data motion estimation

(51)[7] International Patent Classification(s)
H04N 007/28

(21) Application No: **200037921** (22) Application Date: **2000.06.05**

(30) Priority Data

(31) Number	(32) Date	(33) Country
PQ0901	**1999.06.10**	**AU**

(43) Publication Date : **2000.12.14**
(43) Publication Journal Date : **2000.12.14**

(71) Applicant(s)
Canon Kabushiki Kaisha

(72) Inventor(s)
Delphine Anh Dao Le

(74) Agent/Attorney
SPRUSON and FERGUSON,GPO Box 3898,SYDNEY NSW 2001

Abstract

METHOD AND APPARATUS FOR IMPROVED VIDEO DATA MOTION ESTIMATION

A method of generating a field of motion vectors based on motion compensated video data, the field of motion comprising a plurality of blocks of one or more pixels, where each block has an associated motion vector. The method initially provides (201) a first frame and a second frame of video data and corresponding macroblock vectors. Each macroblock of the first frame has at least one associated prediction motion vector referenced to a corresponding macroblock of the second frame. The method assigns these prediction motion vectors as initial vectors (202) of said field of motion vectors associated with the first frame, wherein the initial motion vectors are constant within the macroblocks. The method then modifies the field of initial motion vectors by performing a relaxation procedure (204) thereon to minimise a predetermined energy function.

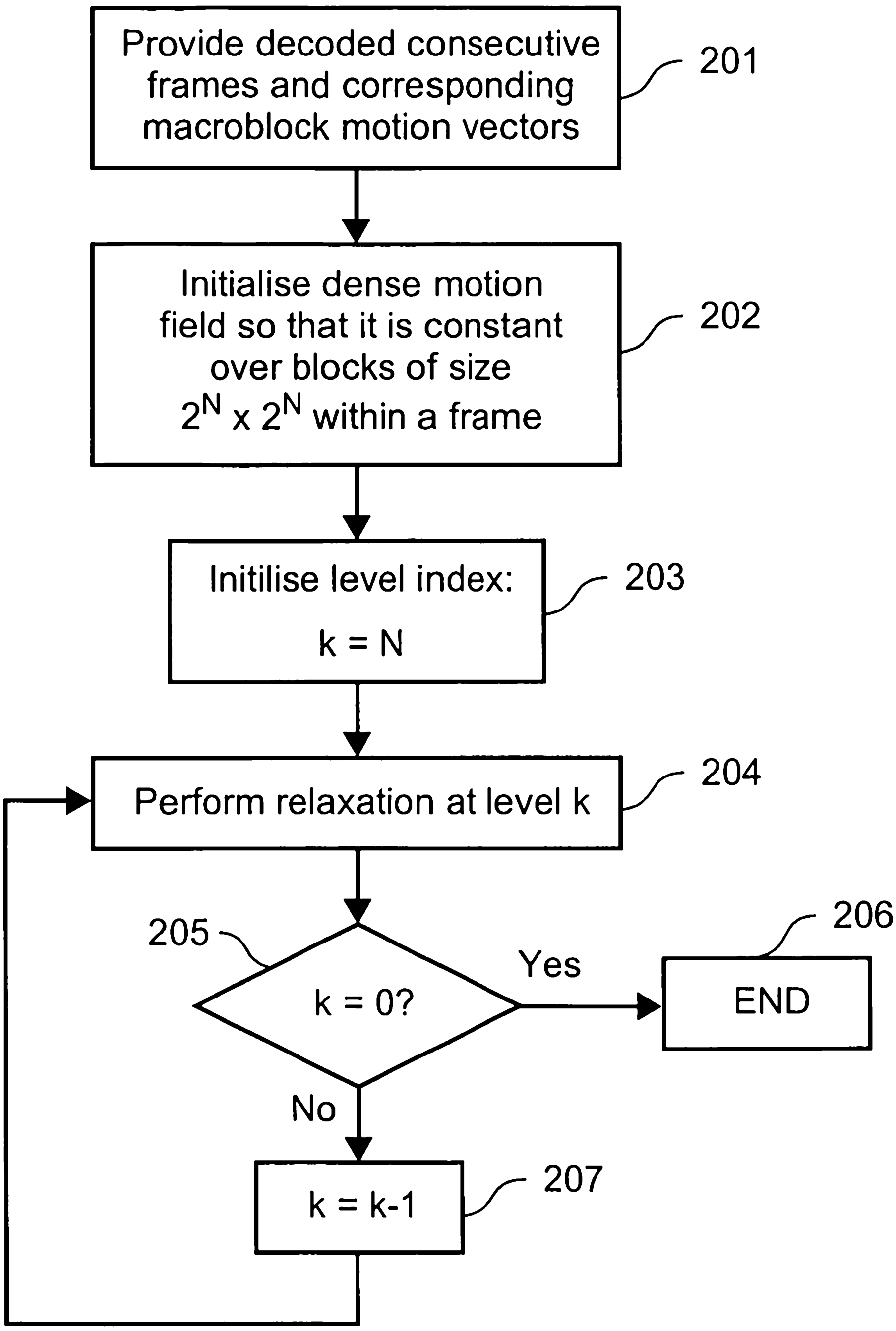

Provide decoded consecutive frames and corresponding macroblock motion vectors
201
Initialise dense motion field so that it is constant over blocks of size 2^N x 2^N within a frame
202
Initilise level index:
k = N
203
Perform relaxation at level k
204
205
k = 0?
Yes
END
206
No
k = k-1
207

FIG. 2

S&F Ref: 504626

AUSTRALIA

PATENTS ACT 1990

COMPLETE SPECIFICATION

FOR A STANDARD PATENT

ORIGINAL

Name and Address of Applicant :	Canon Kabushiki Kaisha 3-30-2, Shimomaruko Ohta-ku Tokyo 146 Japan
Actual Inventor(s):	Delphine Anh Dao Le
Address for Service:	Spruson & Ferguson St Martins Tower 31 Market Street Sydney NSW 2000
Invention Title:	Method and Apparatus for Improved Video Data Motion Estimation

ASSOCIATED PROVISIONAL APPLICATION DETAILS

[33] Country	[31] Applic. No(s)	[32] Application Date
AU	PQ0901	10 Jun 1999

The following statement is a full description of this invention, including the best method of performing it known to me/us:-

METHOD AND APPARATUS FOR IMPROVED VIDEO DATA MOTION ESTIMATION

Field of Invention

The present invention relates to motion estimation in compressed video data and, in particular, to motion-compensated encoded data.

The invention has been developed primarily for producing a field of motion vectors based on motion compensated encoded data, and will be described herein with reference to this specific application. However, it will be appreciated that the invention is not limited to use in this field.

Background

Image motion plays an important role in computer vision and scene understanding. Motion vector analysis has been applied to many fields over the last few decades, including object tracking, autonomous navigation, surveillance and virtual reality. More recently, motion information has played an important role in video indexing, contributing to video segmentation and shot classification.

Some of these applications require motion information to be estimated as a motion field, where a per-pixel vector field density is required. It is difficult to find a typical motion estimation method that is both efficient and accurate at the same time.

Summary of the Invention

It is an object of the present invention to substantially overcome, or at least ameliorate, one or more disadvantages of existing arrangements.

According to one aspect of the invention, there is provided a method of generating a field of motion vectors based on motion compensated video data, the field comprising a plurality of first blocks of one or more pixels, where each said first block has an associated motion vector, said method comprising the steps of: (a) providing a first frame and a second frame of video data, wherein the first and second frames each comprise a plurality of macroblocks each having a plurality of said first blocks and wherein each said macroblocks of the first frame has at least one associated prediction motion vector referenced to a corresponding macroblock of the second frame; (b) assigning said prediction motion vectors as initial vectors of said field of motion vectors associated with the first frame, wherein the initial motion vectors are constant within the macroblocks; and (c) modifying said field of initial motion vectors by performing a relaxation procedure thereon to minimise a predetermined energy function.

According to another aspect of the invention, there is provided a method of producing a field of motion vectors from a first and second frame of video data, wherein the first and second frame comprise a plurality of first blocks of one or more pixels, said method comprising the steps of: generating a field of motion vectors by performing an iterative relaxation procedure to optimise a predetermined energy function, wherein each first block of the first frame has an associated said motion vector; wherein said generating step comprises: providing a said first frame and second frame of video data, wherein the first and second frames each comprise a plurality of macroblocks each having a plurality of said first blocks and wherein each said macroblocks of the first frame has at least one associated prediction motion vector referenced to a corresponding macroblock of the second frame; and assigning said prediction motion vectors as initial vectors of said field of motion vectors associated with the first frame, wherein the initial motion vectors are constant within the macroblocks.

According to another aspect of the invention, there is provided apparatus for generating a field of motion vectors based on motion compensated video data, the field comprising a plurality of first blocks of one or more pixels, where each said first block has an associated motion vector, said apparatus comprising: means for providing a first frame and a second frame of video data, wherein the first and second frames each comprise a plurality of macroblocks each having a plurality of said first blocks and wherein each said macroblocks of the first frame has at least one associated prediction motion vector referenced to a corresponding macroblock of the second frame; means for assigning said prediction motion vectors as initial vectors of said field of motion vectors associated with the first frame, wherein the initial motion vectors are constant within the macroblocks; and means for modifying said field of initial motion vectors by performing a relaxation procedure thereon to minimise a predetermined energy function.

According to another aspect of the invention, there is provided apparatus for producing a field of motion vectors from a first and second frame of video data, wherein the first and second frame comprise a plurality of first blocks of one or more pixels, said apparatus comprising: means for generating a field of motion vectors by performing an iterative relaxation procedure to optimise a predetermined energy function, wherein each first block of the first frame has an associated said motion vector; wherein said generating means comprises: means for providing a said first frame and second frame of video data, wherein the first and second frames each comprise a plurality of macroblocks each having a plurality of said first blocks and wherein each said macroblocks of the first frame has at least one associated prediction motion vector referenced to a corresponding macroblock of the second frame; and means for assigning said prediction motion vectors as initial vectors of said field of motion vectors associated with the first frame, wherein the initial motion vectors are constant within the macroblocks.

According to another aspect of the invention, there is provided a computer readable medium comprising a computer program for generating a field of motion vectors based on motion compensated video data, the field comprising a plurality of first blocks of one or more pixels, where each said first block has an associated motion vector, said computer program comprising: means for providing a first frame and a second frame of video data, wherein the first and second frames each comprise a plurality of macroblocks each having a plurality of said first blocks and wherein each said macroblocks of the first frame has at least one associated prediction motion vector referenced to a corresponding macroblock of the second frame; means for assigning said prediction motion vectors as initial vectors of said field of motion vectors associated with the first frame, wherein the initial motion vectors are constant within the macroblocks; and means for modifying said field of initial motion vectors by performing a relaxation procedure thereon to minimise a predetermined energy function.

According to another aspect of the invention, there is provided a computer readable medium comprising a computer program for producing a field of motion vectors from a first and second frame of video data, wherein the first and second frame comprise a plurality of first blocks of one or more pixels, said computer program comprising: means for generating a field of motion vectors by performing an iterative relaxation procedure to optimise a predetermined energy function, wherein each first block of the first frame has an associated said motion vector; wherein said generating means comprises: means for providing a said first frame and second frame of video data, wherein the first and second frames each comprise a plurality of macroblocks each having a plurality of said first blocks and wherein each said macroblocks of the first frame has at least one associated prediction motion vector referenced to a corresponding macroblock of the second frame; and means for assigning said prediction motion vectors as initial vectors of said field of motion vectors associated with the first frame, wherein the initial motion vectors are constant within the macroblocks.

Other aspects of the invention are also disclosed.

Brief Description of Drawings

Preferred embodiments of the invention will now be described, by way of example only, with reference to the accompanying drawings, in which:

Fig. 1 shows a general-purpose computer for implementing preferred embodiments of the invention;

Fig. 2 shows a flow chart setting out steps involved in implementing the preferred embodiment of the invention;

Fig. 3 shows a frame divided into a plurality of macroblocks;

Fig. 4 shows a macroblock from the frame in Fig. 3, divided into sub-blocks;

Fig. 5 shows the sub-blocks from Fig. 4 with energy-minimised vectors applied to each of the sub-blocks; and

Fig. 6 shows a sub-block from Fig. 5 subdivided into further sub-blocks.

Detailed Description

Recent compression standards for moving images, like H.261 and MPEG, have adopted a macroblock approach for motion compensation. A target macroblock in a frame to be encoded is matched with a most similar displaced macroblock in a previous (or consecutive) frame, called a reference image. The position of the best matching macroblock, or prediction macroblock, is indicated by a motion vector that describes a displacement from the target macroblock to the prediction macroblock. The motion vector information is encoded and transmitted along with compressed image frames. In the H.261 and MPEG standards, the macroblock size is chosen to be 16x16 pixels, representing a tradeoff between motion-compensated compression and the cost of transmitting the motion vectors.

Using the preferred embodiment of the invention, it is possible to take advantage of the motion information embedded in the compressed data in order to compute a relatively dense and more accurate motion field.

Preferred embodiments of the invention are implemented on a conventional general-purpose (host) computer system, such as the computer system 40 shown in Fig. 1. The application program discussed above (and thereby the preferred embodiments of the invention described with reference to the other drawings) is implemented as software executed on the computer system 40.

The computer system 40 includes a computer module 41, input devices such as a keyboard 42 and mouse 43, and output devices including a printer 56 and a display device 57. A modulator-demodulator ("modem") transceiver device 52 is used by the computer module 41 for communication via a telecommunications network 61, such as a telephone line, Local Area Network ("LAN") or Wide Area Network ("WAN"). The computer module 41 comprises an I/O Interface 10 for coupling the computer module 41 to the computer Network 61. The modem 52 can also be used to access the Internet or other telecommunications networks.

The computer module 41 includes a number of functional components, including (in the embodiment shown):

a processor unit 45;

a memory unit 46 including semiconductor Random Access Memory ("RAM") and Read Only Memory ("ROM"); and

input/output ("I/O") interfaces including a video interface 47;

an I/O interface 48 for the keyboard 42 and mouse 43;

a storage device 49, including a hard disk drive 53 and a floppy disk drive 54; and

a CD-ROM drive 55 for use as a non-volatile source of data.

The components 45 to 49 and 53 to 55 of the computer module 41 communicate via an interconnecting bus 50 in a manner that results in conventional operation of the computer system 40, as known to those in the relevant art. Examples of computers on which the embodiments can be practised include IBM-PC's and compatibles, and Sun Sparcstations.

Typically, the application program of the preferred embodiment is resident on a hard disk drive 53 and read and controlled using the processor 45. Intermediate storage of the program and the print list and any data fetched from the network may be accomplished using the semiconductor memory 46, possibly in concert with the hard disk drive 53. In some instances, the application program may be supplied to the user encoded on a CD-ROM or floppy disk, or alternatively could be read by the user from the network via the modem device 52.

Depending upon the implementation, it may be desirable to use a hardware-based graphics accelerator to improve performance of the system as a whole.

A flowchart 200 in Fig. 2 sets out the steps involved in implementing the preferred embodiment of the invention. In an initial step 201, decoded frames from an MPEG stream are provided. At least a first frame 300 (Fig. 3) and a second frame (not shown) are provided. Each frame 300 consists of macroblocks 301, which, in the case of frames originally MPEG encoded, are 16x16 pixels in size.

Each frame is based on either an I frame having DCT encoded macroblocks, or an intermediate B or P frame. Where the frame 300 is based on a B or P frame, each macroblock is based on a best match macroblock from a preceding or subsequent frame in the originally encoded series of video frames. The best matching macroblock vector is referred to by means of a motion vector 302. The motion vector 302 represents a relative coordinate offset, the resultant position of which represents the best match macroblock in a preceding or subsequent frame (the second frame referred to in this specification).

In step 202, a motion field is generated covering the whole of the frame 300, as shown in Fig. 4. For this step, each of the macroblocks 301 is considered a sub-block associated with the first frame. Each sub-block (macroblock) 301 is defined by a plurality of pixels 303, which represent, in this case, the desired resolution of the motion field. Within each sub-block (macroblock) 301, each pixel 303 is initialised with the vector 302 having a value equal to that of the corresponding sub-block (macroblock) 301, as shown in Fig. 3.

In step 203, a level index k is initialised with a value N, the purpose of which will become apparent from the following paragraphs. It will be appreciated that the precise point in the method at which k is initialised is not critical. For example, k could be initialised at the very beginning of the procedure prior to step 201.

A relaxation at level k is then performed in step 204. Relaxation in the present context refers to the iterative implementation of an energy function E to each of the pixels 303, until E is minimised. The implementation of the energy function is discussed in greater detail below.

In step 205 an assessment is made as to whether k=0. If the answer is affirmative, the method ends at step 206. If not, the method proceeds to step 207, in which k is decremented. Decrementing k amounts to dividing the sub-block 301 into further sub-blocks 304 (Fig. 6).

The method then returns to step 204, and repeats the relaxation for the new value of k. In effect, changing the value of k reduces the size of the sub-block within which the motion vectors are kept the same as each other. Steps 204, 205 and 207 are then repeated until k=0 and the process ends. At that point, each pixel has associated with it an individual motion vector that is not constrained to be the same as the motion vectors associated with any of its neighbouring pixels.

As discussed above, encoded video data have a motion vector per macroblock, those vectors being defined hereinafter as Macroblock Motion Vectors (MMV's). Scene understanding or computer vision tasks which use motion information often require a dense and accurate motion vector field, sometimes to the level of one motion vector per pixel; motion-based segmentation is one such case.

When using an iterative method, using macroblock motion vectors to provide an initial guess, in accordance with the preferred embodiment, can improve the speed and accuracy of convergence. This is particularly appropriate if the refined motion estimation method is a hierarchical approach which performs a coarse-to-fine motion estimation, for example, the multiscale relaxation algorithm proposed by Heitz et al. [F. Heitz, P. Perez and P. Bouthemy, "Multiscale minimisation of global energy functions in some visual recovery problems", CVGIP: Image Understanding, vol.59, pp.125-134, 1994].

In implementation, the preferred embodiment minimizes a global energy function, in which the energy function models desirable properties of the motion field. Such properties can include, for example, intensity invariance under motion, or smoothness of the motion field. These two exemplary constraints can be applied by formulating the energy as the weighted sum of two terms. An energy function E can then be written as:

$$E(U) = \sum_{s \in S} [I_t(s) - I_{t+1}(s + u(s))]^2 + \alpha \sum_{<s,p> \in C} \| u(s) - u(p) \|^2$$

wherein:

S is the set of all pixels s in a frame;

U is the set of motion vectors u(s);

I is the image intensity at time t;

I_{t+1} is the image intensity of the subsequent frame;

α is a weighting factor;

<s,p> is a pair of neighbouring pixels; and

C is the set of all neighbouring pixels.

It will be appreciated by those skilled in the art that any other suitable constraint or constraints can also be utilised in an energy function. Multiple constraints are usually dealt with by applying a weighting factor to one or more of the constraints within the energy function. This allows manipulation of the relative contribution of each constraint to the value E of the energy function.

Preferably the weighting factor α is optimised for each image. Otherwise it can be arbitrarily set. For instance, it can be set to a high value, e.g. 100 which tends to result in a smooth motion vector field.

The energy function (E) is minimised over nested subspaces of the original space of possible solutions, these subspaces consisting of solutions constrained at different scales. For example, scale 0 corresponds to the original configuration space, whilst scale k corresponds to configurations which are constant over blocks of size $2^k x 2^k$.

Letting N be the predetermined number of scales, and following a coarse-to-fine strategy, the preferred energy minimisation method starts at level N. A relaxation is performed in order to find a motion field that is constant over blocks of size $2^N x 2^N$ and minimises the global energy function E. At level N-1, each parent block of level N is divided into 4 child blocks (sub-blocks), and the motion field is then constrained to be constant over the child blocks (sub-blocks) of size $2^{N-1} x 2^{N-1}$. The estimate obtained at level N defines a crude solution which can be used as initial guess at level N-1. The energy minimisation is performed again at level N-1, and so on, until level 0 is reached.

In the preferred embodiment, at each level, the energy minimisation is performed using an iterative deterministic relaxation process known as Iterated Conditional Modes ("ICM"). This optimisation method converges relatively quickly but towards the closest local minimum. For this reason, the initial "guess" provided by the motion vectors associated with the encoded macroblock data is an important aspect of this embodiment in that it leads to fast and reliable results.

If the coarsest chosen scale corresponds directly to the block size used for motion-compensated coding (eg N=4 for MPEG), then the MMV can directly be used as the initial values for the multiscale motion estimation algorithm. It will be appreciated, on the other hand, that the scale chosen need not correspond directly with such a block size. Further, it is not strictly necessary that square blocks be used, nor that the block division process be exponential in nature.

Further, when bidirectional temporal prediction is used as in MPEG, two MMVs may be available for one block. That being the case, the two available MMVs can be combined or averaged to obtain a unique motion vector for each block.

Another potential use of the MMV is to provide an indicator of the motion magnitude, thereby enabling a reduction of the search space required for the refined motion estimation. In this aspect, for a given macroblock, only a subset of the motion vector space centered around the compressed motion vector needs to be explored. It will be appreciated by those skilled in the art that the range of the search space is an important factor in the time complexity of the algorithm, especially for relaxation methods like the ICM algorithm. For example, the boundaries of the search space can be defined depending on the range spanned by the MMVs. For instance, the search space can be of the order of the largest possible magnitude of the MMVs. This can be done on a global or local basis. Thus the complexity of the calculation of the minimisation of the energy function can be reduced.

Both aspects of the preferred embodiment (ie, the use of an initial "guess" and restriction of the search space) can considerably reduce the processing time of the consequent motion estimation procedure, by accelerating the convergence of the motion estimation algorithm and yielding a better final estimate.

The foregoing describes only one embodiment/embodiments of the present invention, and modifications and/or changes can be made thereto without departing from the spirit or scope of the invention. The present embodiments are, therefore, to be considered in all respects to be illustrative and not restrictive.

In the context of this specification, the word “comprising” means “including principally but not necessarily solely” or “having” or “including” and not “consisting only of”. Variations of the word comprising, such as “comprise” and “comprises” have corresponding meanings.

The claims defining the invention are as follows:

1. A method of generating a field of motion vectors based on motion compensated video data, the field comprising a plurality of first blocks of one or more pixels, where each said first block has an associated motion vector, said method comprising the steps of:

(a) providing a first frame and a second frame of video data, wherein the first and second frames each comprise a plurality of macroblocks each having a plurality of said first blocks and wherein each said macroblocks of the first frame has at least one associated prediction motion vector referenced to a corresponding macroblock of the second frame;

(b) assigning said prediction motion vectors as initial vectors of said field of motion vectors associated with the first frame, wherein the initial motion vectors are constant within the macroblocks; and

(c) modifying said field of initial motion vectors by performing a relaxation procedure thereon to minimise a predetermined energy function.

2. A method as claimed in claim 1, wherein the modifying step is performed iteratively.

3. A method as claimed in claim 2, wherein the modifying step comprises the sub-steps of:

(c)(1) performing the relaxation procedure over the macroblocks to determine a said field of motion vectors constant over the corresponding macroblocks;

(c)(2) dividing said macroblocks into sub-blocks;

(c)(3) performing the relaxation procedure over the sub-blocks to determine a said field of motion vectors constant over the corresponding sub-blocks;

(c)(4) dividing the sub-blocks into further sub-blocks;

(c)(5) repeating steps (c)(3) and (c)(4) until the sub-blocks reach the size of the first blocks; and

(c)(6) performing the relaxation procedure over the first blocks to determine a final said field of motion vectors.

4. A method according to claim 1, wherein the relaxation procedure includes the step of minimising a global energy function E(U) based on one or more variable features associated with the video data.

5. A method according to claim 4, wherein:

$$E(U) = \sum_{s \in S} [I_t(s) - I_{t+1}(s + u(s))]^2 + \alpha \sum_{<s,p> \in C} \| u(s) - u(p) \|^2$$

and:

S is the set of all pixels s in a frame;

U is the set of motion vectors u(s);

I is the image intensity at time t;

I_{t+1} is the image intensity of the subsequent frame;

α is a weighting factor;

<s,p> is a pair of neighbouring pixels; and

C is the set of all neighbouring pixels.

6. A method according to claim 3, wherein the macroblocks (and thereby the initial sub-blocks) are of size $2^N x 2^N$, and steps c(2) and c(4) include the step of dividing the macroblocks or sub-blocks into four further sub-blocks of equal size.

7. A method according to claim 1, wherein the first blocks are a single pixel in size.

8. A method as claimed in claim 1, wherein the relaxation procedure has a search space, the range of which is of the order of the largest magnitude of the prediction vectors on a local or global basis.

9. A method of producing a field of motion vectors from a first and second frame of video data, wherein the first and second frame comprise a plurality of first blocks of one or more pixels, said method comprising the steps of:

generating a field of motion vectors by performing an iterative relaxation procedure to optimise a predetermined energy function, wherein each first block of the first frame has an associated said motion vector; wherein said generating step comprises:

providing a said first frame and second frame of video data, wherein the first and second frames each comprise a plurality of macroblocks each having a plurality of said first blocks and wherein each said macroblocks of the first frame has at least one associated prediction motion vector referenced to a corresponding macroblock of the second frame; and

assigning said prediction motion vectors as initial vectors of said field of motion vectors associated with the first frame, wherein the initial motion vectors are constant within the macroblocks.

10. Apparatus for generating a field of motion vectors based on motion compensated video data, the field comprising a plurality of first blocks of one or more pixels, where each said first block has an associated motion vector, said apparatus comprising:

means for providing a first frame and a second frame of video data, wherein the first and second frames each comprise a plurality of macroblocks each having a plurality of said first blocks and wherein each said macroblocks of the first frame has at least one associated prediction motion vector referenced to a corresponding macroblock of the second frame;

means for assigning said prediction motion vectors as initial vectors of said field of motion vectors associated with the first frame, wherein the initial motion vectors are constant within the macroblocks; and

means for modifying said field of initial motion vectors by performing a relaxation procedure thereon to minimise a predetermined energy function.

11. Apparatus as claimed in claim 10, wherein the modifying means is performed iteratively.

12. Apparatus as claimed in claim 11, wherein the modifying means comprises:

means for performing the relaxation procedure over the macroblocks to determine a said field of motion vectors constant over the corresponding macroblocks;

means for dividing said macroblocks into sub-blocks;

means for performing the relaxation procedure over the sub-blocks to determine a said field of motion vectors constant over the corresponding sub-blocks;

means for dividing the sub-blocks into further sub-blocks;

means for repeating the operations of the means for performing the relaxation procedure over the sub-blocks and the means for dividing the sub-blocks until the sub-blocks reach the size of the first blocks; and

means for performing the relaxation procedure over the first blocks to determine a final said field of motion vectors.

13. Apparatus according to claim 1, wherein the relaxation procedure includes means for minimising a global energy function E(U) based on one or more variable features associated with the video data.

14. Apparatus according to claim 13, wherein:

$$E(U) = \sum_{s \in S} [I_t(s) - I_{t+1}(s + u(s))]^2 + \alpha \sum_{<s,p> \in C} \| u(s) - u(p) \|^2$$

and:

S is the set of all pixels s in a frame;

U is the set of motion vectors u(s);

I is the image intensity at time t;

I_{t+1} is the image intensity of the subsequent frame;

α is a weighting factor;

<s,p> is a pair of neighbouring pixels; and

C is the set of all neighbouring pixels.

15. Apparatus according claim 12, wherein the macroblocks (and thereby the initial sub-blocks) are of size $2^N x 2^N$, and the means for dividing said macroblocks and the means for dividing the sub-blocks include means for dividing the macroblocks or sub-blocks into four further sub-blocks of equal size respectively.

16. Apparatus according to claim 10, wherein the first blocks are a single pixel in size.

17. Apparatus according to claim 10, wherein the relaxation procedure has a search space, the range of which is of the order of the largest magnitude of the prediction vectors on a local or global basis.

18. Apparatus for producing a field of motion vectors from a first and second frame of video data, wherein the first and second frame comprise a plurality of first blocks of one or more pixels, said apparatus comprising:

means for generating a field of motion vectors by performing an iterative relaxation procedure to optimise a predetermined energy function, wherein each first block of the first frame has an associated said motion vector; wherein said generating means comprises:

means for providing a said first frame and second frame of video data, wherein the first and second frames each comprise a plurality of macroblocks each having a plurality of said first blocks and wherein each said macroblocks of the first frame has at least one associated prediction motion vector referenced to a corresponding macroblock of the second frame; and

means for assigning said prediction motion vectors as initial vectors of said field of motion vectors associated with the first frame, wherein the initial motion vectors are constant within the macroblocks.

19. A computer readable medium comprising a computer program for generating a field of motion vectors based on motion compensated video data, the field comprising a plurality of first blocks of one or more pixels, where each said first block has an associated motion vector, said computer program comprising:

means for providing a first frame and a second frame of video data, wherein the first and second frames each comprise a plurality of macroblocks each having a plurality of said first blocks and wherein each said macroblocks of the first frame has at least one associated prediction motion vector referenced to a corresponding macroblock of the second frame;

means for assigning said prediction motion vectors as initial vectors of said field of motion vectors associated with the first frame, wherein the initial motion vectors are constant within the macroblocks; and

means for modifying said field of initial motion vectors by performing a relaxation procedure thereon to minimise a predetermined energy function.

20. A computer readable medium comprising a computer program for producing a field of motion vectors from a first and second frame of video data, wherein the first and second frame comprise a plurality of first blocks of one or more pixels, said computer program comprising:

means for generating a field of motion vectors by performing an iterative relaxation procedure to optimise a predetermined energy function, wherein each first block of the first frame has an associated said motion vector; wherein said generating means comprises:

means for providing a said first frame and second frame of video data, wherein the first and second frames each comprise a plurality of macroblocks each having a plurality of said first blocks and wherein each said macroblocks of the first frame has at least one associated prediction motion vector referenced to a corresponding macroblock of the second frame; and

means for assigning said prediction motion vectors as initial vectors of said field of motion vectors associated with the first frame, wherein the initial motion vectors are constant within the macroblocks.

21. A method of producing a field of motion vectors, the method substantially as described herein with reference to the accompanying drawings.

22. Apparatus for producing a field of motion vectors, the apparatus substantially as described herein with reference to the accompanying drawings.

23. A computer readable medium comprising a computer program for producing a field of motion vectors, the computer program substantially as described herein with reference to the accompanying drawings.

DATED this FIFTH Day of JUNE 2000

Canon Kabushiki Kaisha

Patent Attorneys for the Applicant

SPRUSON & FERGUSON

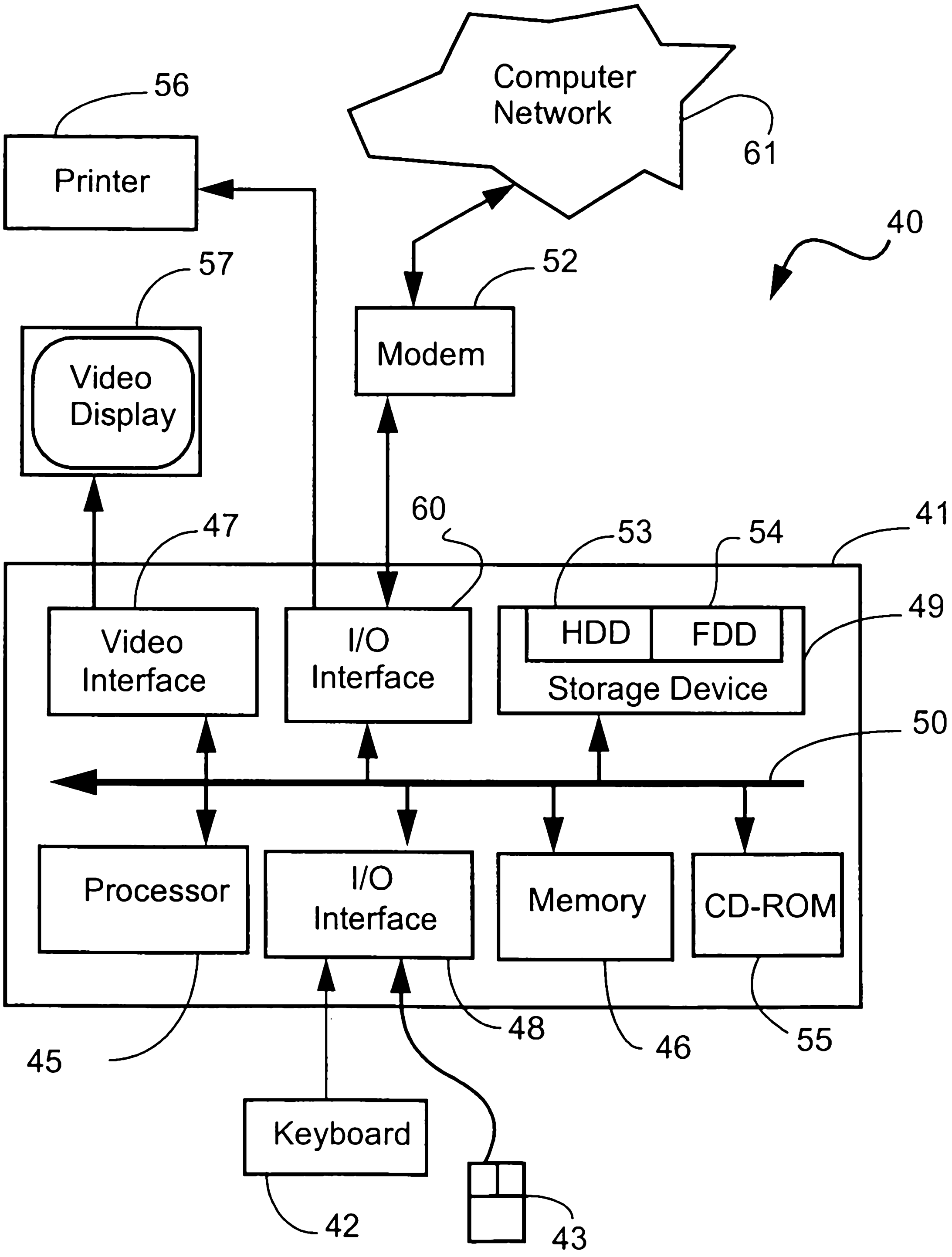
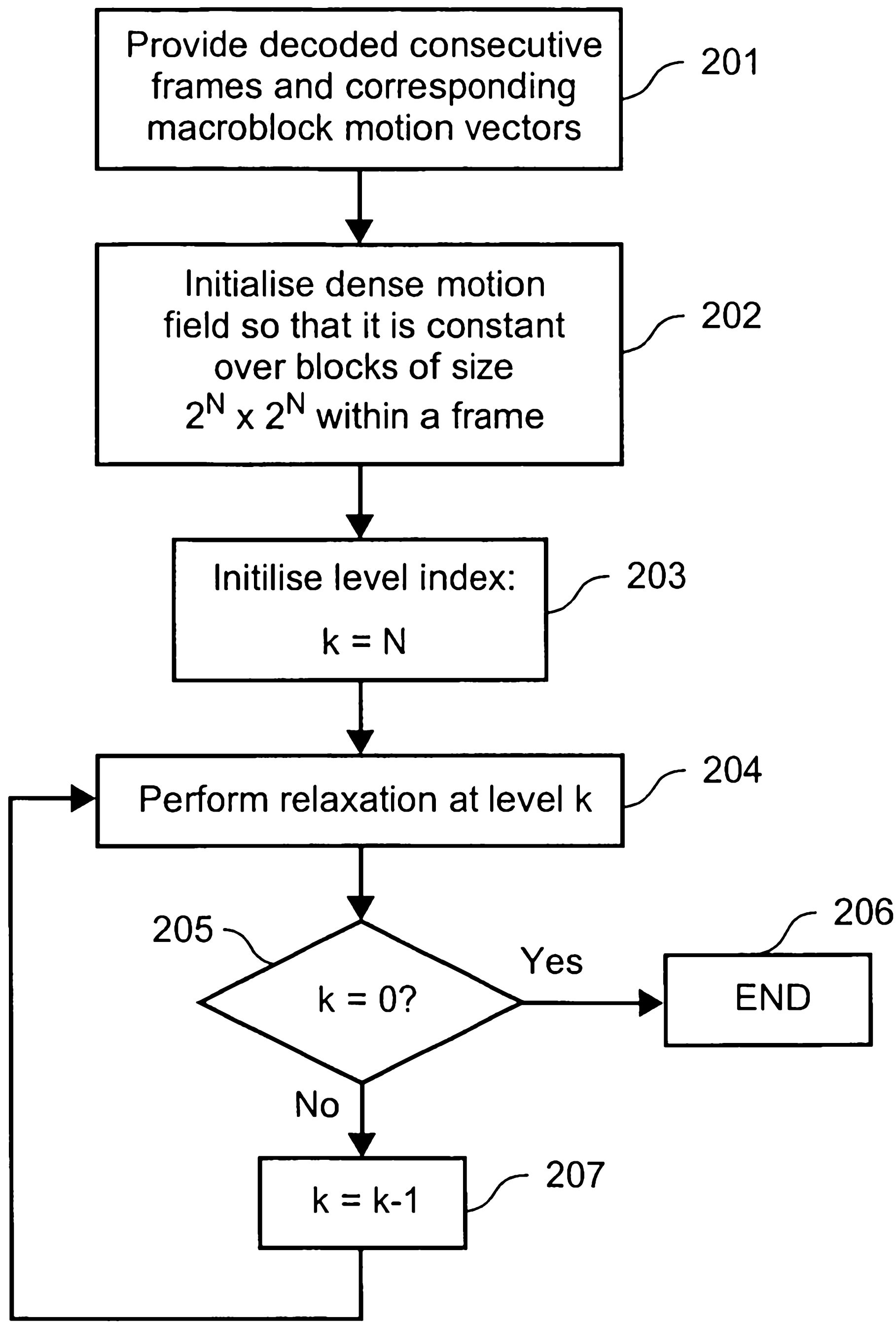

Computer
Network
61
56
Printer
40
57
52
Modem
Video
Display
60
47
53
54
41
Video
Interface
I/O
Interface
HDD
FDD
Storage Device
49
50
Processor
I/O
Interface
Memory
CD-ROM
45
48
46
55
Keyboard
42
43

FIG. 1



Provide decoded consecutive frames and corresponding macroblock motion vectors
201
Initialise dense motion field so that it is constant over blocks of size $2^N \times 2^N$ within a frame
202
Initilise level index:
k = N
203
Perform relaxation at level k
204
205
k = 0?
Yes
END
206
No
k = k-1
207

FIG. 2

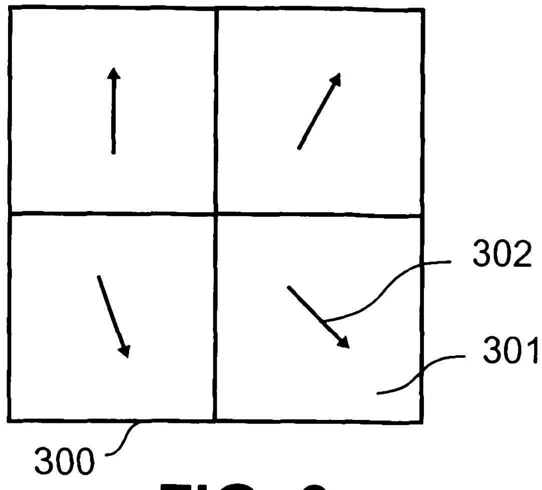
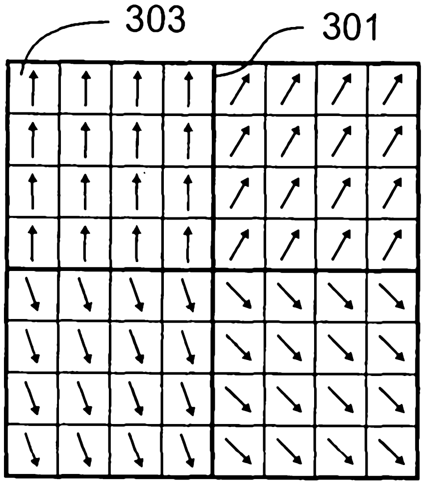


FIG. 3



Level N

FIG. 4

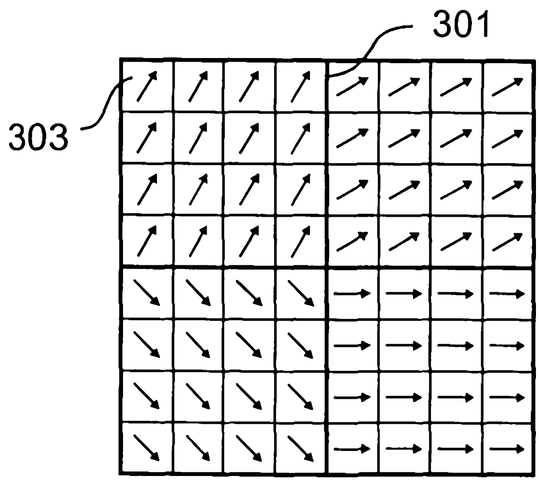


FIG. 5

Level N-1

FIG. 6