

(19) World Intellectual Property Organization
International Bureau



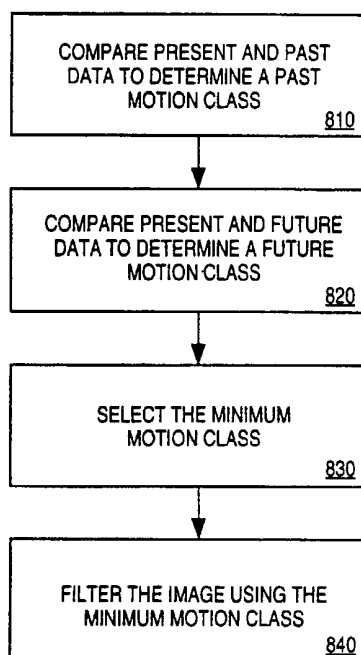
(43) International Publication Date
14 June 2001 (14.06.2001)

PCT

(10) International Publication Number
WO 01/43069 A2

- (51) International Patent Classification⁷: **G06T**
- (21) International Application Number: PCT/US00/42664
- (22) International Filing Date: 6 December 2000 (06.12.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
09/457,419 7 December 1999 (07.12.1999) US
- (71) Applicant: **SONY ELECTRONICS, INC.** [US/US]; 1 Sony Drive, Park Ridge, NJ 07656 (US).
- (72) Inventors: **KONDO, Tetsujiro**; 12-1255 1-14 Tsumadakita, Atsugi-Sji, Kanagawa-prefecture 243 (JP). **CARRIG, James, J.**; 1652 Hicks Avenue, San Jose, CA 95125 (US). **FUJIMORI, Yasuhiro**; 11693 Westshore Court, Cupertino, CA 95014 (US). **CAREY, William, Knox**; 1355 McKendrie Street, San Jose, CA 95126 (US).
- (74) Agents: **SOBRINO, Maria, McCormack et al.**; Blakely, Sokoloff, Taylor & Zafman, 12400 Wilshire Boulevard, 7th floor, Los Angeles, CA 90025-1026 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- Published:**
— Without international search report and to be republished upon receipt of that report.
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD AND APPARATUS FOR PAST AND FUTURE MOTION CLASSIFICATION



(57) Abstract: A method comprising determining a past motion class for target data (810), determining a future motion class for the target data (820), selecting one of the motion classes (830), filtering the target data using the selected motion class is disclosed (840).



WO 01/43069 A2

METHOD AND APPARATUS FOR PAST AND FUTURE MOTION CLASSIFICATION

FIELD OF THE INVENTION

This invention relates generally to the processing of image, sound or other correlated signals, and more particularly, to a method, apparatus, and article of manufacture for past and future motion classification.

BACKGROUND OF THE INVENTION

Conventionally, error recovery has been achieved by correlation evaluation. For example, some recovery choices have been implemented using a conventional error pixel recovery method. Using neighboring data, spatial inclinations of the target data are detected. For example, the inclinations regarding four directions are evaluated according to the predetermined formulae which use the neighboring data. An interpolation filter is chosen where the inclination value, E_i , is the smallest among the four values calculated. In addition to the spatial inclination, a motion factor is also evaluated for error recovery. In the case of the motion area, a selected spatial filter is used for error recovery. On the other hand, the previous frame data at the same location as the target data typically are used for error recovery.

The conventional error recovery process discussed above may cause many serious degradations on changing data, especially on object edges. Actual signal distribution typically varies widely, so these problems are likely to occur. Therefore, there is a need for a way to restore a deteriorated signal to an undeteriorated signal which minimizes degradations on changing data.

SUMMARY OF THE INVENTION

A method comprising determining a past motion class for target data (810), determining a future motion class for the target data (820), selecting one of the

motion classes (830), and filtering the target data using the selected motion class is disclosed (840).

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is illustrated by way of example and may be better understood by referring to the following description in conjunction with the accompanying drawings, in which like references indicate similar elements and in which:

FIGS. 1A and 1B illustrate one embodiment of a classified adaptive error class;

FIG. 2 illustrates one example of motion class tap structures;

FIGS. 3A and 3B show an example of an interlaced video;

FIGS. 4A, 4B, and 4C show an example of three consecutive fields of video;

FIG. 5 shows one example of motion degradation for appearing pixels;

FIG. 6 shows one example of motion degradation for disappearing pixels;

FIGS. 7A and 7B show examples of a filter for past and future data;

FIG. 8 shows one embodiment of a method for minimum motion classification; and

FIG. 9 shows one embodiment of an apparatus for performing minimum motion filtering.

DETAILED DESCRIPTION

In the following description, reference is made to the accompanying drawings which form a part hereof, and in which is shown by way of illustration a specific embodiment in which the invention may be practiced. It is to be

understood that other embodiments may be utilized and structural changes may be made without departing from the scope of the present invention.

A method and apparatus for past and future motion classification described. In the following description, numerous details are set forth. It will be apparent, however, to one skilled in the art, that the present invention may be practiced without these specific details. In other instances, well-known structures and devices are shown in block diagram form, rather than in detail, in order to avoid obscuring the present invention.

Some portions of the detailed descriptions which follow are presented in terms of algorithms and symbolic representations of operations on data bits within a computer memory. These algorithmic descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the following discussion, it is appreciated that throughout the description, discussions utilizing terms such as "processing" or "computing" or "calculating" or "determining" or "displaying" or the like, refer to the action and processes of a

computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

The present invention also relates to apparatus for performing the operations herein. This apparatus may be a circuit or system specially constructed for the required purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but is not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, and magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, or any type of media suitable for storing electronic instructions, and each coupled to a computer system bus.

The algorithms and displays presented herein are not inherently related to any particular computer or other apparatus. Various general purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct more specialized apparatus to perform the required method steps. The required structure for a variety of these systems will appear from the description below. In addition, the present invention is not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the invention as described herein.

Classified adaptive error recovery is a technology which utilizes classified adaptive filter processing. A proper classification with respect to the deteriorated

input signal is performed according to the input signal characteristics. An adaptive filter is prepared for each class prior to error recovery processing.

More than one classification method may optionally be used to generate the plurality of classes. Generated classes may include a motion class, an error class, a spatial activity class or a spatial class. An adaptive class tap structure may optionally be used to generate the plurality of classes. An adaptive filter tap structure may optionally be used according to the class which is detected in each deteriorated input signal. The adaptive filter tap structure may optionally be expanded based upon multiple taps. The number of filter coefficients that must be stored can be reduced by allocating the same coefficient to multiple taps. This process is referred to as filter tap expansion. The deteriorated input signal may optionally be modified by preprocessing peripheral erroneous data. A spatial class may optionally be eliminated according to a spatial class elimination formula.

The present invention can be applied to any form of temporally correlated data, including without limitation, video or other two-dimensional moving images, and three-dimensional moving images, and audio such as stereo. In the description, the term value, in one embodiment, may refer to a component within a set of received or generated data. Furthermore, a data point may be a position, place, instance, location or range within data.

For the sake of clarity, some of the description herein focuses on video data comprising a pixel stream. However, it will be recognized that the present invention may be used with other types of data other than video data and that the terms and phrases used herein to describe the present invention cover a broad range of applications and data types. For example, an adaptive class tap structure is an adaptive structure for class tap definition used in multiple classification. A spatial class, a motion class and an error class may be used to define the structure.

An adaptive filter tap structure is an adaptive structure for filter tap definition based upon a corresponding class.

A class may be defined based on one or more characteristics of the target data. For example, a class may also be defined based on one or more characteristics of the group containing the target data. A class ID is a specific value within the class that is used to describe and differentiate the target data from other data with respect to a particular characteristic. A class ID may be represented by a number, a symbol, or a code within a defined range. A parameter may be used as a predetermined or variable quantity that is used in evaluating, estimating, or classifying the data. For example, the particular motion class ID of a target data can be determined by comparing the level of motion quantity in the block containing the target data against a parameter which can be a pre-determined threshold.

A motion class is a collection of specific values used to describe the motion characteristic of the target data. In one embodiment, the motion class may be defined based on the different levels of motion of the block containing the target data, for example, no motion in the block, little motion in the block, or large motion in the block. A motion class ID is a specific value within the motion class used to indicate a particular level of motion quantity of the target data. For example, motion class ID of "0" may be defined to indicate no motion, motion class ID of "3" may be defined to indicate large motion.

The present invention provides a method and apparatus for adaptive processing that generates data corresponding to a set of one or more data classes. This process is known as "classification." Classification can be achieved by various attributes of signal distribution. For example, Adaptive Dynamic Range Coding (ADRC) may be used for generation of each class as a spatial class, but it will be recognized by one of ordinary skill in the art that other classes, including a motion

class, an error class, and a spatial activity class may be used with the present invention without loss of generality.

For each class, a suitable filter for signal restoration is prepared for the adaptive processing. In one embodiment, each filter may be represented by a matrix of filter coefficients which are applied to the data. The filter coefficients can be generated by a training process, an example of which is described subsequently, that occurs as a preparation process prior to filtering.

In FIG. 1A, an example is shown where the number of class taps is four. In the case of 1-bit ADRC, 16 class IDs are available as given by [formula 3], shown below. ADRC is realized by [formula 2], shown below. Detecting a local dynamic range (DR) is given by [formula 1], shown below,

$$DR = MAX - MIN + 1 \quad [\text{formula 1}]$$

$$q_i = \left\lfloor \frac{(x_i - MIN + 0.5) \cdot 2^Q}{DR} \right\rfloor \quad [\text{formula 2}]$$

$$c = \sum_{i=1}^4 2^{i-1} \cdot q_i \quad [\text{formula 3}]$$

where c corresponds to an ADRC class ID, DR represents the dynamic range of the four data area, MAX represents the maximum level of the four data, MIN represents the minimum level of the four data, q_i is the ADRC encoded data, also referred to as a Q code, and Q is the number of quantization bits. The $\lfloor \cdot \rfloor$ operator represents a truncation operation.

In 1-bit ADRC with four class taps, c may have a value from 0 to 15 with $Q = 1$. This process is one type of spatial classification, but it will be recognized by one

of ordinary skill in the art that other examples of spatial classification, including Differential PCM, Vector Quantization and Discrete Cosine Transform may be used with the present invention without loss of generality. Any method may be used if it can classify a target data distribution.

In the example shown in **FIG. 2B**, each adaptive filter has 12 taps. Output data is generated according to the linear combination operation given by [formula 4], shown below,

$$y = \sum_{i=1}^{12} w_i \cdot x_i \quad \text{[formula 4]}$$

where x_i is input data, w_i corresponds to each filter coefficient, and y is the output data after error recovery. Filter coefficients can be generated for each class ID by a training process that occurs prior to the error recovery process.

As noted above, filter coefficients can be generated by a training process. For example, training may be achieved according to the following criterion:

$$\min_w \|X \cdot W - Y\|_2 \quad \text{[formula 5]}$$

where X , W , and Y are, for example, the following matrices: X is the input data matrix defined by [formula 6], W is the coefficient matrix defined by [formula 7], and Y corresponds to the target data matrix defined by [formula 8].

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix} \quad [\text{formula 6}]$$

$$W = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \quad [\text{formula 7}]$$

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \quad [\text{formula 8}]$$

The coefficient w_i can be obtained according to [formula 5], so that estimation errors against target data are minimized.

In another embodiment of the present invention, motion classification, in addition to spatial classification, may also be used to provide compact definition of temporal characteristics. Further, multiple classification may be added to the classified adaptive error recovery method. As noted above, there are various types of classes, such as a motion class, an error class, a spatial activity class and a spatial class. The combination of one or more of these different classification methods can also improve classification quality.

FIG. 2 shows an example of motion class tap structures. The example shows eight taps in neighborhood of the target error data. In this example, the eight tap accumulated temporal difference can be evaluated according to [formula 9], shown

below, and is classified to four kinds of motion classes by thresholding based on [formula 10], shown below. In one embodiment of the present invention, th0 is equal to 3, th1 is equal to 8, and th2 is equal to 24.

$$fd = \sum_{i=1}^8 |x_i - x'_i| \quad [\text{formula 9}]$$

$$mc = \begin{cases} 0 & (0 \leq fd < th0) \\ 1 & (th0 \leq fd < th1) \\ 2 & (th1 \leq fd < th2) \\ 3 & (th2 \leq fd) \end{cases} \quad [\text{formula 10}]$$

In the above formulas, fd represents an accumulated temporal difference, x_i represents motion class tap data of the current frame, x'_i represents the previous frame tap data corresponding to the current frame, and mc represents a motion class ID. Three thresholds, th0, th1, th2, can be used for this motion classification.

Adaptive filtering systems may use motion classification to select the appropriate filtering technique. This motion classification may include comparing past and present input data to determine an estimate of an object's motion around the pixels of interest. By adding a processing delay to an image filtering system, future data may also be compared with present data and used for motion classification, detection and filtering.

Having both past motion classification and future motion classification may improve the filtering results, particularly in areas of appearing stationary pixels where there was past motion but no future motion. The filtering technique based

on past and future motion classification may be used for up conversion or pixel recovery.

Interlacing two fields to create one frame of an image has an effect on the vertical resolution of an image. In interlaced video systems, only every other line of data is from the current field, as shown in **FIGS. 3A** and **3B**. The highest resolution processing modes rely on past or future data to fill in remaining lines.

When filtering the pixels of a fast moving object only the current field information is safe to use, unless an expensive motion compensation method is provided. Because past or future data for a given location may not be accurate when filtering the fast moving object, the highest resolution processing modes are not available. This may reduce the vertical resolution of the fast moving object by as much as fifty percent.

If the object moves fast enough, the loss of vertical resolution may not be perceived by the human eye. However, a degradation may be noticed when appearing or disappearing pixels are treated as motion pixels.

For purposes of discussion, on herein, an appearing pixel may be defined as a pixel belonging to a background object that appears as the result of being no longer occluded by a foreground object. Conversely, a disappearing pixel may be defined as a pixel belonging to a background object that is occluded by a foreground object.

An example of the visual effects of treating appearing and disappearing pixels as motion pixels is provided with respect to **FIGS. 4A, 4B, and 4C**, which show three consecutive fields of video. The sequence shows a cartoon fly quickly crossing a wooden background from left to right.

For example, the motion of the fly from the past position of **FIG. 4A** to the present position of **FIG. 4B** may create a low resolution cloud that follows the fly, as

shown in **FIG. 5**. The black and gray shaded areas of **FIG. 5** shown the pixels that are classified as motion when the motion class is determined by comparing past and present data for the given area.

The black areas show the moving foreground object, which may be processed by using only one field. The gray area shows the appearing pixels which are highly correlated with present data, but not with past data. In future frames, these pixels may be processed using two field data.

If the pixels in the gray region are presently classified as motion, they will be processed using only one field. As a result, the viewer may notice a change in resolution with respect to the appearing pixels. The degraded resolution of the appearing pixels will form a low resolution cloud that follows the fly.

A similar problem may occur if only the present and future data of **FIGS. 4B** and **4C** are used, as shown in **FIG. 6**. The black area of **FIG. 6** is the moving foreground object, which may be processed using only one field. The gray area shows the disappearing pixels, which are highly correlated with future data but not with present data. If the disappearing pixels are classified as motion pixels, they will be processed using one field data. The result will be a decrease in the resolution of the disappearing pixels, which form a low resolution cloud that precedes the fly.

The past and future data of **FIGS. 4A** and **4C** can be used along the present data of **FIG. 4B** to classify image pixels as stationary, appearing, disappearing, or motion. Within the motion class, different sub-classifications, such as speed or direction, may be included. By comparing past and future data to present data, the image can be filtered using the data which is more closely correlated to the present data, as shown in **FIG. 7A**. Stationary and disappearing pixels may be filtered using past and present field data and appearing pixels may be filtered using present

and future field data, as shown in **FIG. 7B**. Other motion subcategories may be processed using one field or multiple field filters.

FIG. 8 shows an embodiment of a method for minimum motion classification. Present data is compared to past data to determine a past motion class, 810. Present data is compared to future data to determine a future motion class, 820. The minimum motion class is determined by selecting the motion class that is highly correlated to the present data, 830. The image is filtered using the minimum motion class 840.

For example, if the present motion class is less than the future motion class then only past and present data are used in the filter. Alternatively, if the future motion class is less than the past motion class, then only present and future data are used in the filter.

For example, if motion is classified into four speeds, then the forward motion class (FMC) may be expressed as

$$\text{FMC} \in \{0,1,2,3\}$$

The backward motion class (BMC) may be expressed as

$$\text{BMC} \in \{0,1,2,3\}$$

The minimum motion class (MMC) is

$$\text{MMC} = \min \{\text{FMC}, \text{BMC}\}$$

The minimum motion direction detected is

$$\{\text{Forward if } \text{FMC} < \text{BMC}, \text{ else backwards}\}$$

An embodiment of the data used to filter an image for given values of BMC and FMC is provided in Table 1.

Backwards Motion Class	Forwards Motion Class	Use past data	Use present data	Use future data

0	0	✓	✓	
0	1	✓	✓	
0	2	✓	✓	
0	3	✓	✓	
1	0		✓	✓
1	1	✓	✓	
1	2	✓	✓	
1	3	✓	✓	
2	0		✓	✓
2	1		✓	✓
2	2	✓	✓	
2	3	✓	✓	
3	0		✓	✓
3	1		✓	✓
3	2		✓	✓
3	3		✓	

Table 1

An apparatus for performing the method of minimum motion filtering is shown in FIG. 9. Image data including past, present and future data are input to field delays 910 and 920. The present and future data is input into the future motion class detector 930. The present and future data are compared by the detector 930 to create a future motion class. The present and past data are input to a past motion class detector 940, and a past motion class is created. The past and future motion classes are input to minimize motion logic 950, and one of the classes is selected as the minimum motion class. The selected motion class is input to

coefficient memory 960. An appropriate set of coefficients are selected and a filter is selected 970 and provided to filter 980, where the range is filtered.

In one embodiment, the method and apparatus for past and future motion classification may be used to perform pixel error recovery. In an alternative embodiment, the method and apparatus for past and future motion classification may be used to perform up-conversion.

While the invention is described in terms of embodiments in a specific system environment, those of ordinary skill in the art will recognize that the invention can be practiced, with modification, in other and different hardware and software environments within the spirit and scope of the appended claims.

CLAIMS

What is claimed is:

1. A method comprising:
determining a past motion class for target data (810);
determining a future motion class for the target data (820);
selecting one of the motion classes (830); and
filtering the target data using a filter associated with the selected motion class (840).
2. The method of claim 1 wherein determining the past motion class comprises comparing present data with past data.
3. The method of claim 1 wherein determining the future motion class comprises comparing present data with future data.
4. The method of claim 1 wherein selecting comprises selecting a minimum of the future class and past class.
5. The method of claim 1 wherein if the past motion class is less than or equal to the future motion class, then the past motion class is the selected motion class.
6. The method of claim 1 wherein if the future motion class is less than the past motion class, then the future motion class is the selected motion class.

7. The method of claim 1 wherein the filters associated with the future motion class are temporally symmetric to those associated with the past motion class.
8. A computer readable medium containing instructions which, when executed by a processing system, cause the system to perform:
 - determining a past motion class for target data (810);
 - determining a future motion class for the target data (820);
 - selecting one of the motion classes (830); and
 - filtering the target data using a filter associated with the selected motion class (840).
9. The medium of claim 8 wherein determining the past motion class comprises comparing present data with past data.
10. The medium of claim 8 wherein determining the future motion class comprises comparing present data with future data.
11. The medium of claim 8 wherein selecting comprises selecting a minimum of the future class and past class.
12. An apparatus comprising:
 - means for determining a past motion class for target data (940);
 - means for determining a future motion class for the target data (930);
 - means for selecting one of the motion classes (950); and

means for filtering the target data using a filter associated with the selected motion class (980).

13. The apparatus of claim 12 wherein means for determining the past motion class comprises means for comparing present data with past data.

14. The apparatus of claim 12 wherein means for determining the future motion class comprises means for comparing present data with future data.

15. The apparatus of claim 12 wherein means for selecting comprises means for selecting a minimum of the future class and past class.

16. A system comprising:

past motion class logic configured to determine a past motion class for target data;

future motion class logic configured to determine a future motion class for the target data;

minimum motion class logic configured to select one of the motion classes;
and

filtering logic configured to filter the target data using a filter associated with the selected motion class.

17. The apparatus of claim 16 wherein the past motion class logic configured to determine the past motion class comprises comparing logic configured to compare present data with past data.

18. The apparatus of claim 16 wherein the future motion class logic configured to determine the future motion class comprises comparing logic configured to compare present data with future data.
19. The apparatus of claim 16 wherein the minimum motion class logic comprises logic configured to select a minimum of the future class and past class.

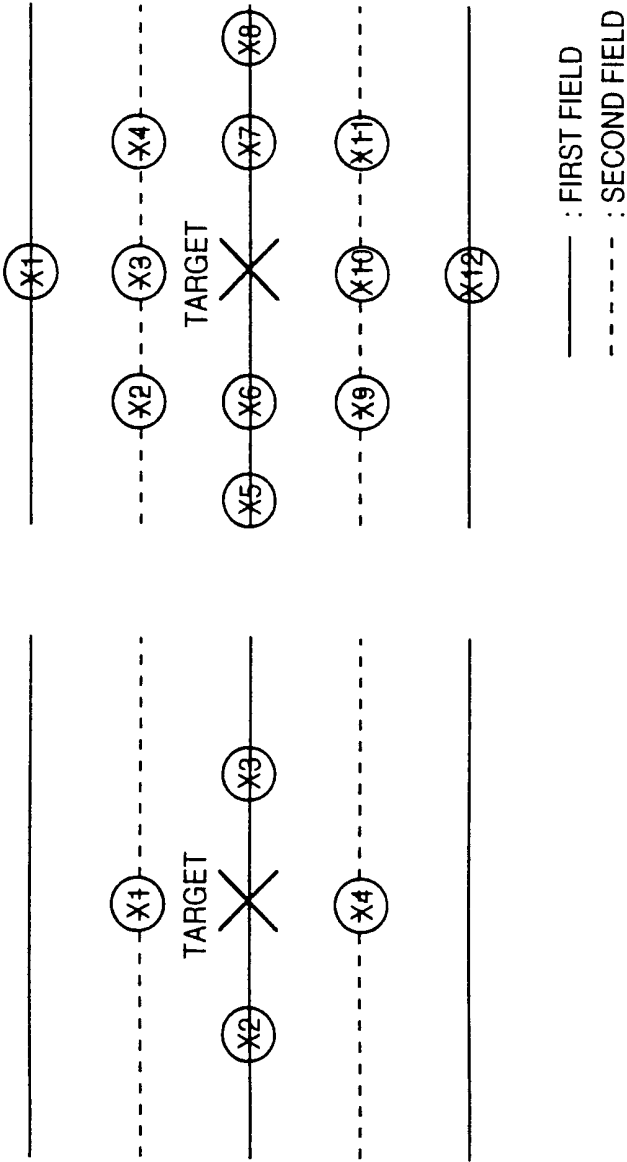


FIG. 1A

FIG. 1B

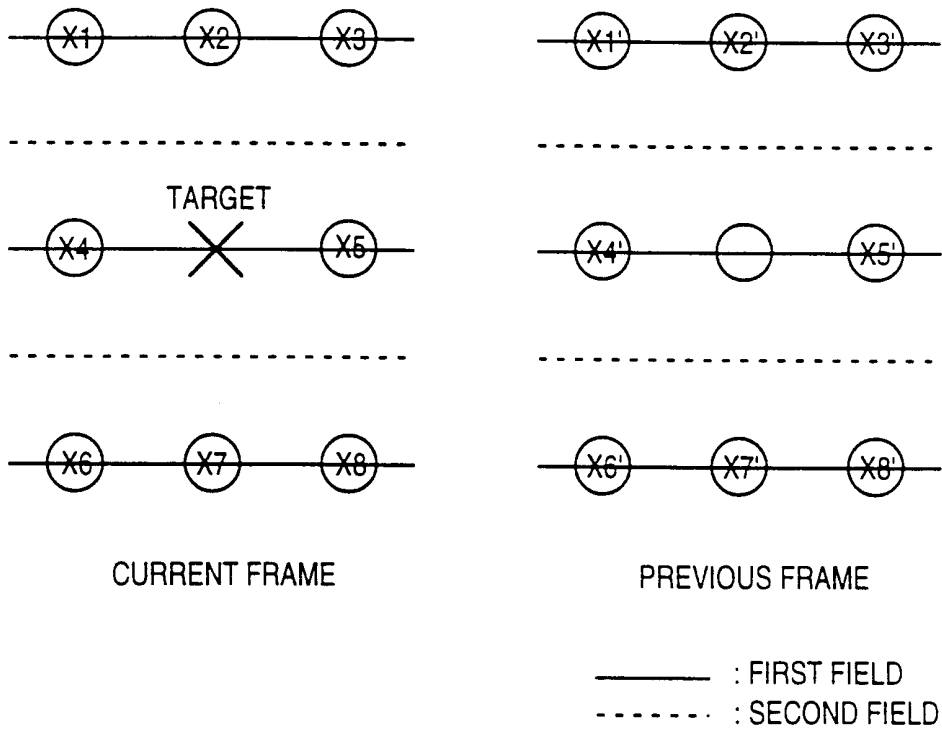


FIG. 2

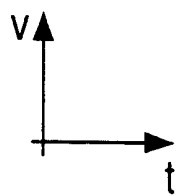
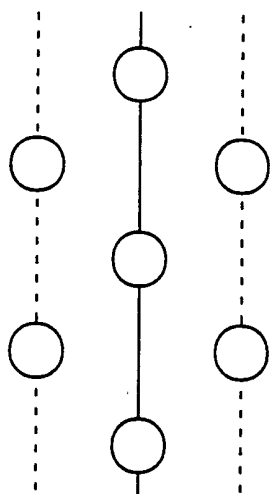


FIG. 3A

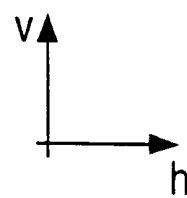
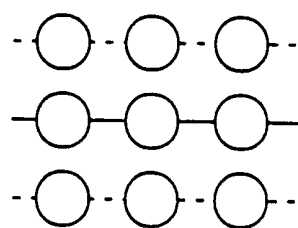


FIG. 3B

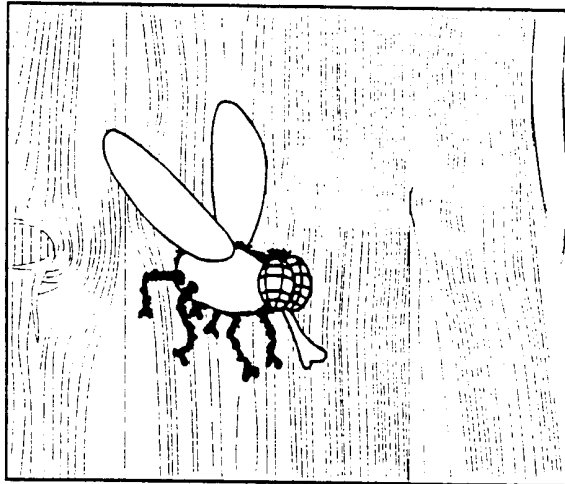


FIG. 4A

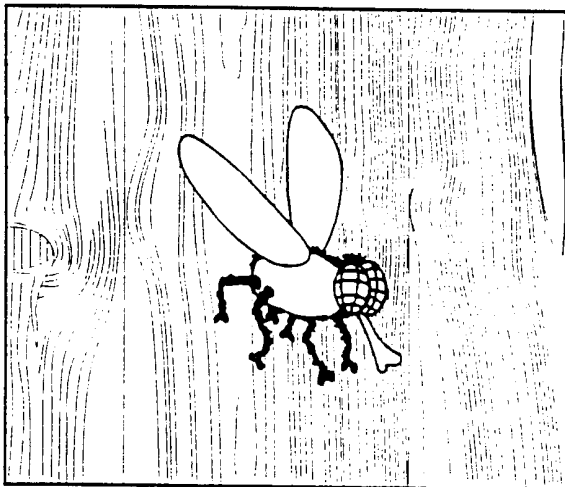


FIG. 4B

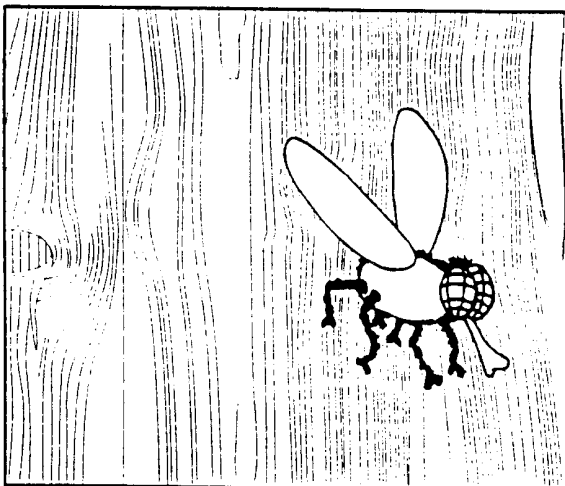


FIG. 4C

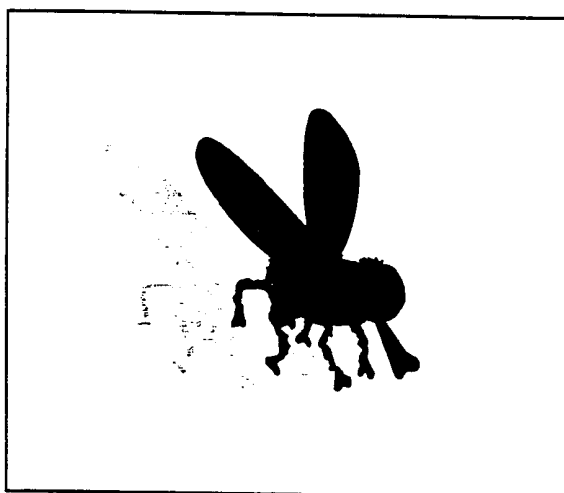


FIG. 5

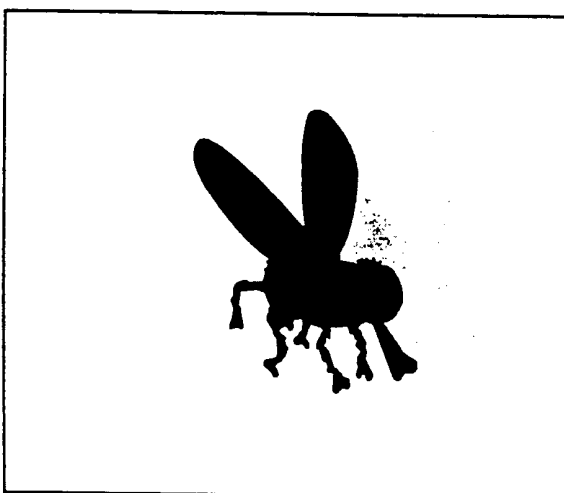


FIG. 6

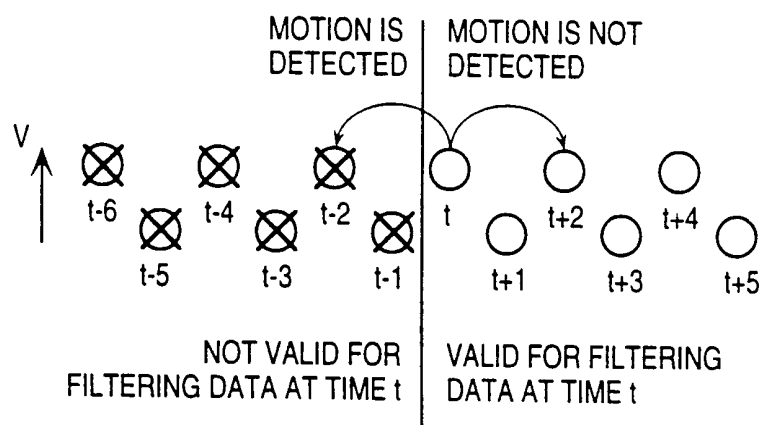


FIG. 7A

7/9

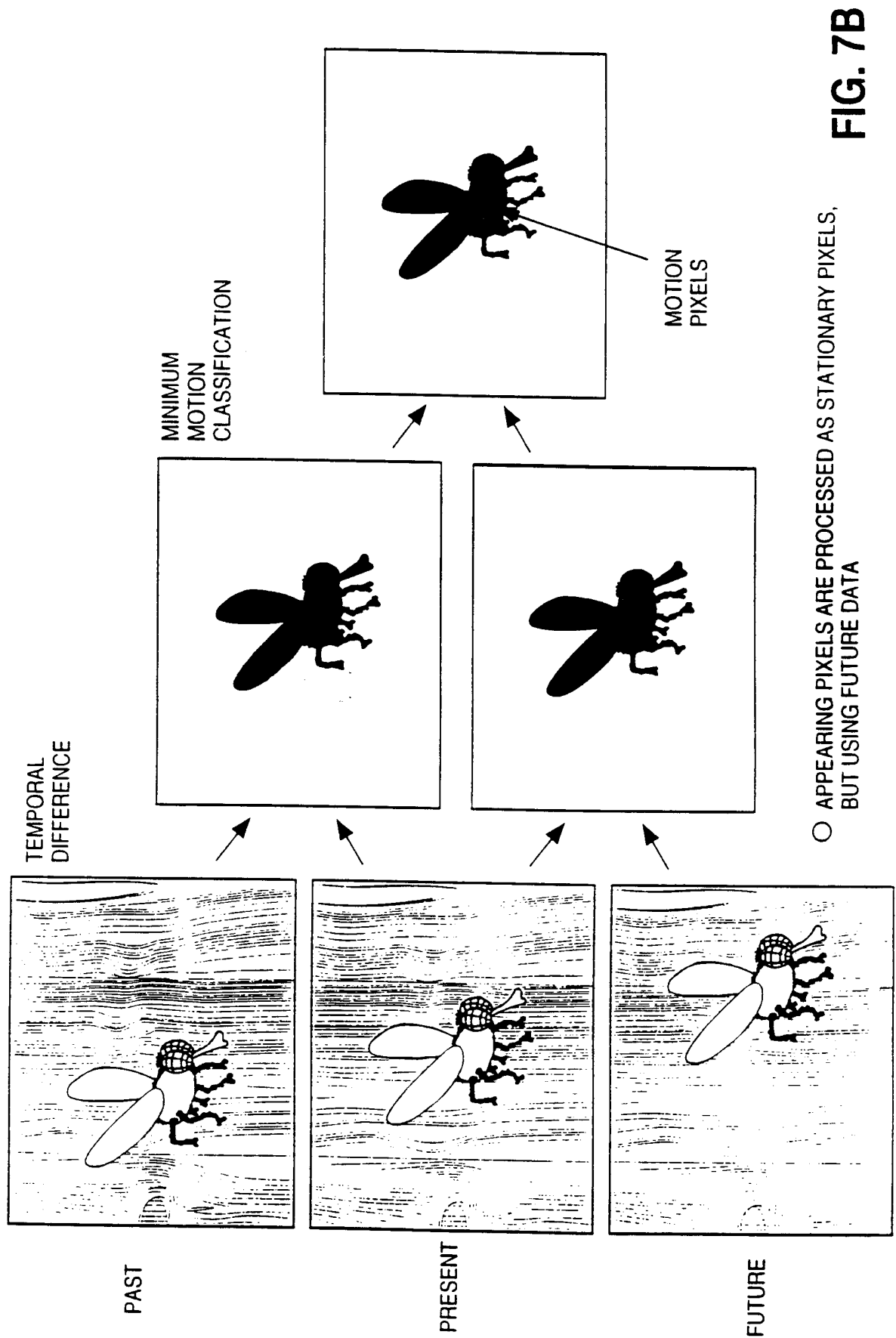
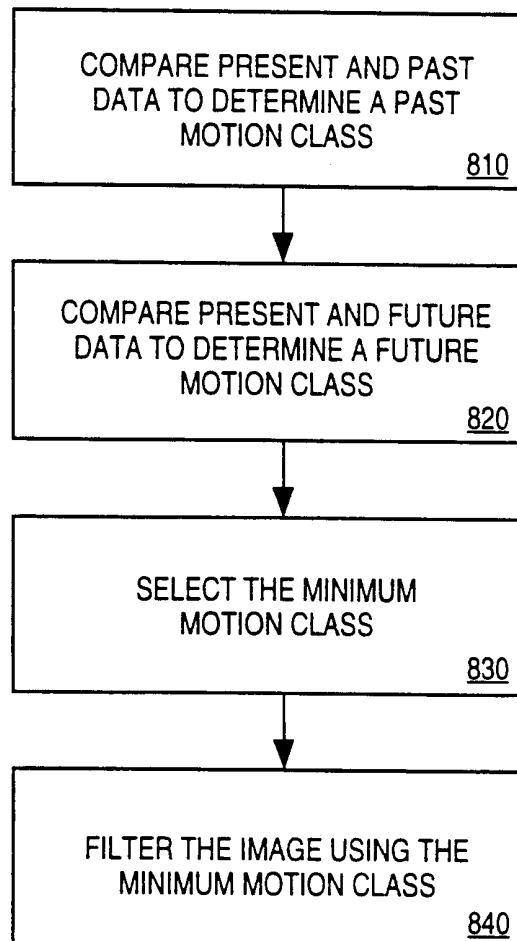
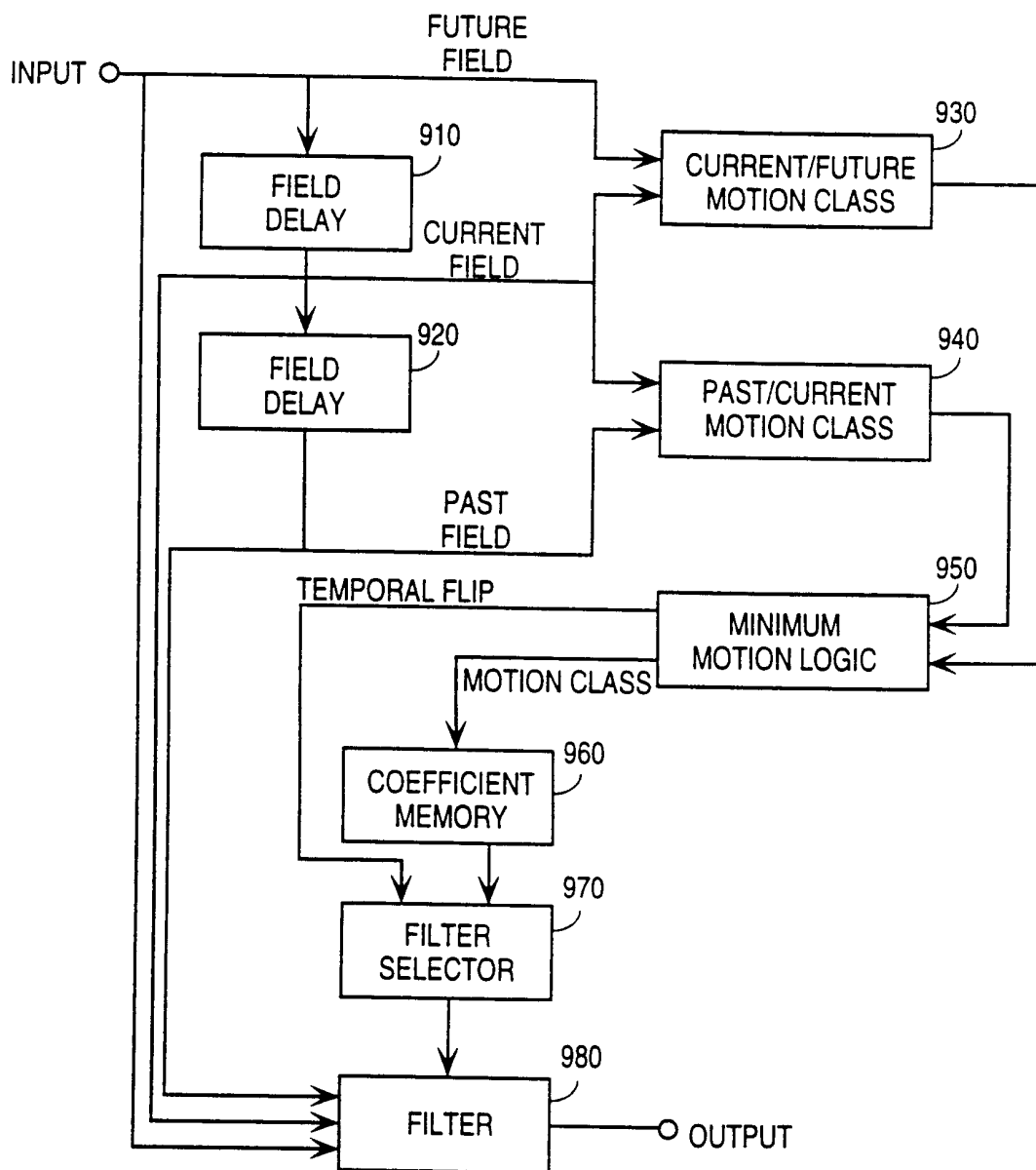


FIG. 7B

**FIG. 8**

**FIG. 9**