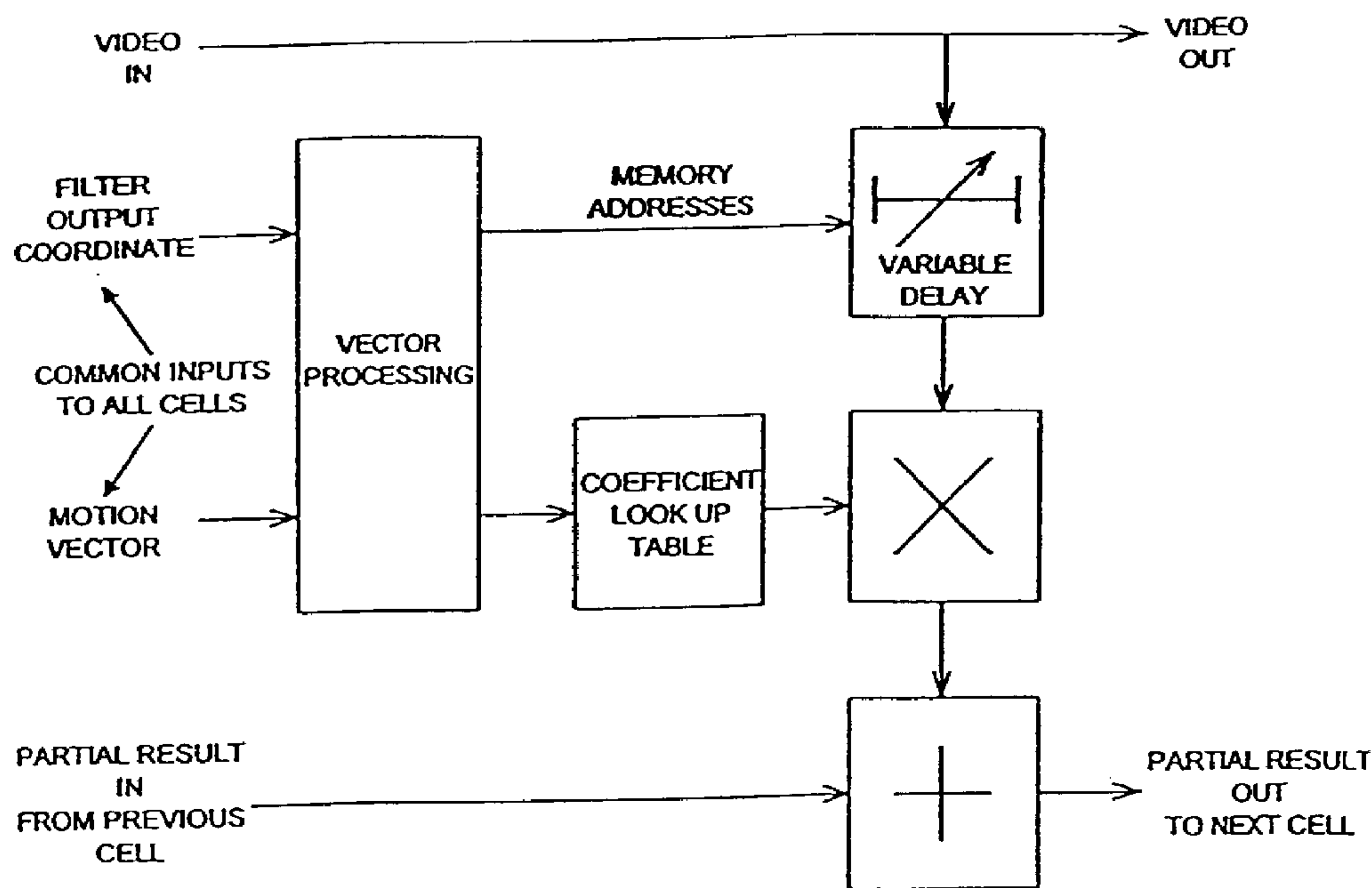




(72) Borer, Timothy John, GB
(73) INNOVISION LIMITED, GB

(51) Int.Cl.⁶ H04N 7/00

(54) **FILTRAGE POUR COMPENSATION DE MOUVEMENT**
(54) **MOTION COMPENSATED FILTERING**



(57) Méthode et système de traitement vidéo permettant le filtrage à compensation de mouvement de signaux de télévision entrelacés. Comprend un dispositif d'estimation de mouvement pour produire des vecteurs de mouvement correspondant à chaque emplacement d'échantillonnage de sortie, un processeur de vecteurs générant un ensemble de coordonnées d'entrée pour chaque emplacement d'échantillonnage de sortie en fonction de son vecteur de mouvement correspondant, une mémoire de données pour stocker des valeurs de pixels d'entrée et produire un retard variable, un certain nombre de multiplicateurs et de coefficients stockés

(57) A video processing method and system for motion compensated filtering of interlaced television signals comprising a motion estimation device for producing motion vectors corresponding to each output sampling site, a vector processor generating a set of input coordinates for each output sampling site depending on its corresponding motion vector, a data store in memory for storing input pixel values and providing a variable delay, a plurality of multipliers and corresponding coefficient stores, and an adder to sum the partial result from each multiplier, the multipliers multiply an input pixel value, selected from the data store depending on a



(11) (21) (C) **2,188,984**
(22) 1996/10/28
(43) 1998/04/28
(45) 2000/10/24

correspondants, et un sommateur pour mettre en sommation le résultat partiel provenant de chaque multiplicateur. Les multiplicateurs multiplient une valeur de pixel d'entrée, sélectionnée dans la mémoire de données en fonction d'une partie de la coordonnée d'entrée, par un coefficient sélectionné parmi les coefficients stockés, en fonction d'une autre partie de la coordonnée d'entrée. L'ouverture du filtre (ensemble des coefficients du filtre) est sélectionnée en fonction de la composante de vitesse verticale du vecteur de mouvement produit par le dispositif d'estimation de mouvement.

part of the input coordinate, by a coefficient selected from the coefficient store depending on another part of the input coordinate, the filter aperture (set of filter coefficients) is selected depending on the vertical velocity component of the motion vector supplied by the motion estimation device.

Abstract

A video processing method and system for motion compensated filtering of interlaced television signals comprising a motion estimation device for producing motion vectors corresponding to each output sampling site, a vector processor generating a set of input coordinates for each output sampling site depending on its corresponding motion vector, a data store in memory for storing input pixel values and providing a variable delay, a plurality of multipliers and corresponding coefficient stores, and an adder to sum the partial result from each multiplier, the multipliers multiply an input pixel value, selected from the data store depending on a part of the input coordinate, by a coefficient selected from the coefficient store depending on another part of the input coordinate, the filter aperture (set of filter coefficients) is selected depending on the vertical velocity component of the motion vector supplied by the motion estimation device.

This invention relates to improved motion compensated filtering or interpolation of interlaced signals and in particular video sequences such as television pictures, with application in all areas where interlaced video sequences are processed using motion vectors, including (but not limited to):

- i. video compression
- ii. video to film conversion
- iii. motion compensated colour decoding
- iv. video motion estimation
- v. television standards conversion
- vi. slow motion display

In this application the term video sequence refers to an image sequence sampled in an interlaced fashion. An image may be a multi-dimensional representation of some property (typically, but not necessarily brightness) of a scene, thus a moving image is a temporal and spatial representation of a property of a scene and an image sequence is a moving image sampled at specific instants. The instant at which a point in the moving image is sampled may depend on the spatial location of that point. A pixel is a sample of an image having a value corresponding to some property of the scene which the image represents and corresponding to a specific image co-ordinate. The image co-ordinate, a specification of the location of a point in an image, need not correspond to a point at which a sampled image has been sampled. Hence image co-ordinates (input and output) are not necessarily quantised. A sampling site is the image co-ordinate of a pixel in a sampled image.

Conventional motion compensated filtering or interpolation is described in detail in references 2, 3, 4, 5, 10, 12, 13, 14, 16, 17, 18 & 20 listed in appendix 1. This process will henceforth be referred to as MCF (or conventional MCF) which will be assumed to include the process of interpolation as well as filtering.

The hardware implementation of a conventional motion compensated filter/interpolator is described in references 3, 4, 18 & 20. One simple implementation is as a cascade of filter interpolation cells as illustrated in figure 1.

Figure 1 shows a conventional MCF filter cell having a vector processor

which converts the motion vectors for each output sampling site in an output image into the input coordinates of an input image used in the filtering or interpolation operation.

The vector processor receives the output sampling site, generated by counters as described in the references, and the associated motion vector from motion estimator hardware (described in the references) and searches through consecutive fields to find the correct input co-ordinate corresponding to the output sampling site concerned.

The motion estimator is a device for estimating the motion vectors corresponding to a moving image. The estimated motion vectors can be sampled on either the input or output sampling lattices for an image processing system. For MCF input and output sampling lattices are (usually) identical. For motion compensated interpolation, the estimated motion vectors should be sampled on the output sampling lattice.

The input coordinate depends on the output sampling site, the motion vector, and the location of the filter tap within the filter aperture. A variable delay supplies a multiplier with the correct input pixel value (from an input image to an image processing system) for the current output sampling site which is determined by the integer part of the current output sampling site. The multiplier multiplies the value of the pixel by a pre-calculated coefficient held in a lookup table stored in ROM. The coefficient is determined by the required filter characteristics and the sub-pixel, or fractional, part of the input coordinate motion vector supplied by the vector processor as shown in figure 1.

A typical motion compensated transversal filter interpolator, such as that used for television standards conversion, might have an aperture spanning four fields, 4 field lines on each field and 2 pixels on each field line. Field line spacing is twice the picture line spacing which is the distance between vertically consecutive lines in a vertically sampled image. Vertically consecutive picture lines need not be sampled at the same temporal instant. For a (interlaced) video sequence, vertically consecutive picture lines will be on temporally adjacent fields. Such a filter would, therefore,

require 32 interpolation/filter cells cascaded together. In one possible implementation the variable delay could be implemented using dual port RAM (random access memory). By making the dual port RAM able to store 4 complete fields, any point in the filter aperture is accessible from the variable delay. A delay is implemented by providing offset read and write addresses as described in reference 3, page 179. The specific delay implemented for a given output pixel depends on the integer part of the motion vector for that pixel. The multiplier coefficients used are stored, as a lookup table, in ROM (read only memory). The specific coefficient to be used is selected from the complete filter aperture by the fractional part of the motion vector. In this example each coefficient lookup table would contain 1/32nd part of the complete filter/interpolation aperture. In other implementations the coefficient lookup tables might contain the complete aperture (set of coefficients) (reference 3). For convenience the coefficient lookup table could be combined with the multiplier in a look-up table (implemented in ROM) as shown in reference 4.

The size of the filter/interpolation coefficient lookup table, for a conventional motion compensated filter, is determined by the size of the filter aperture and the precision of motion vectors supplied to the filter. Consider, for example, the standards conversion interpolator described above assuming a vector precision of 1/16 picture line spacing or horizontal pixel spacing per field. The size of the coefficient lookup tables is calculated as shown below. The number of spatial phases corresponds to the spatial precision of the motion vectors. The number of temporal phases determines the temporal precision of the interpolation (see references 3,6 & 7), typically there are 16 temporal phases for a commercial standards converter.

Horizontally	2 taps 16 phases	= 5 bits
Vertically	4 field lines and 16 phases = 8 picture lines and 16 phases	=7 bits
Temporally	4 fields and 16 phases	=6 bits
TOTAL		= 18 bits

In this example the total coefficient lookup table size has an address space of 18 bits. The lookup table is partitioned into 32 separate smaller lookup tables, one for each interpolation cell, hence the size of each lookup table is only 13 bits. In other implementations the coefficient lookup table might have to contain the complete aperture. The difference between a motion compensated filter and interpolator is simply that a filter only has a single temporal phase whereas an interpolator has multiple temporal phases. Typically a filter will generate an output temporally co-sited with the input fields. Multiple phase, or polyphase filters are described in detail in reference 8.

It is highly desirable to prevent aliasing in video processing as it is visually very disturbing. In conventional (spatio-temporal) MCF a single interpolation aperture (set of filter coefficients) is used. This process is unable to remove undesirable signal aliases due to the interlaced nature of the video signal. These unwanted aliases, present in the output of a conventional MCF, can significantly degrade the performance of a video processing system using MCF. The reason for this unwanted aliasing is that a single interpolation aperture is used for all motion speeds. Furthermore, the filter aperture is generally designed considering only the magnitude of the (spatio-temporal) signal spectrum ignoring phase.

Accordingly the invention provides a filter or interpolation device for processing video sequences comprising a plurality of filter cells having between them a filter aperture which varies depending on the vertical velocity components of motion vectors.

The invention also provides a system for motion compensated filtering of interlaced television signals comprising a motion estimation device for producing motion vectors corresponding to each output sampling site, a vector processor generating a set of input coordinates for each output sampling site depending on its corresponding motion vector, a data store in memory for storing input pixel values and providing a variable delay, a plurality of multipliers and corresponding coefficient stores, the multipliers multiply an input pixel value selected from the data store depending on a part of the input coordinate by a coefficient selected from the coefficient store depending on another part of the input coordinate, the filter aperture (set of filter

coefficients) is selected depending on the vertical velocity component of the motion vector supplied by the motion estimation device, and an adder to sum the partial result from each multiplier.

The filter or interpolator may comprise a coefficient look-up table stored in memory, wherein the total number of different coefficient sets available is given by 2^{p+1} , where p is the vertical vector precision in bits. The vertical vectors being measured in units of input picture lines/field period.

The invention also consists in a method for processing interlaced video signals which adapts the filter/interpolator aperture according to the phases of the signal and alias components of the spatial spectra of the input fields. The method is achieved by using a filter aperture which depends on the vertical velocity associated with each output sampling site. Different filter apertures are used for different velocities and the filter aperture used can vary between adjacent output sampling sites.

The method may comprise assigning a motion vector to each output sampling site, producing a set of input coordinates for each output sampling site, selecting a filter aperture (set of filter coefficients) from multiple pre-calculated filter apertures depending on the vertical part of each motion vector, selecting a coefficient from the filter aperture depending on the fractional part of each input coordinate, selecting an input pixel value depending on the integer part of each input coordinate (where input co-ordinates are measured in input picture lines per field period), and multiplying each input pixel value by the correspondingly selected coefficient to produce a partial result for each output sampling site and summing the partial results.

The number of filter apertures required is dependent on the precision of the vertical component of the motion vector.

Improved filtering is achieved by designing the filter's frequency response, based on the phase and magnitudes of the signal and alias components in the spatial spectra of input fields, in order to cancel the unwanted alias components leaving only the desired signal components.

This invention, by designing filter apertures as described above and

detailed below and adapting the filter/interpolation aperture to the motion vector, enables a significant reduction of unwanted aliasing at the output of a motion compensated filter/interpolator. By considering the phase and magnitude of signal and alias components of consecutive spatial spectra it is possible, as described below, to design a filter aperture that substantially cancels aliasing which would otherwise be present in the output of a conventional MCF. To achieve this a different aperture is required for each different vertical motion speed. In practical terms this is achieved by selecting a different filter/interpolation aperture depending on vertical motion speed. Quantisation of the vertical motion vectors supplied to the motion compensated filter ensures that only a finite number of different filter apertures are required. This invention might, concisely, be described as vertical velocity adaptive motion compensation filtering (VAMCF).

Brief Description of the Drawings

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

Figure 1 is a block diagram of a motion compensation filter/interpolation cell according to a prior art implementation of motion compensated filtering;

Figure 2 is a block diagram of a VAMCF/ interpolation cell according to a first embodiment;

Figure 3 is a block diagram of an alternative VAMCF/ interpolation cell according to another embodiment;

Figure 4 is a block diagram of a method of interlace to progressive conversion;

Figure 5 is a block diagram of a composite video decoder; and

Figure 6 is a schematic of field timing for calculating field difference signals;

Detailed Description of the Invention

In the method of the invention one of many different interpolation apertures is used for filtering depending on the vertical velocity component of a motion vector. With appropriate filter design (as described in appendix 2) the aliasing present in the interlaced input can be cancelled to produce an output substantially free from unwanted alias. In order to do this different interpolation apertures are required

for different vertical motion speeds. The number of different filter apertures required depends on the vertical vector precision. Assuming the vertical motion speed is measured in picture lines/field period then the number of different coefficient sets required is given by: $Sets = 2^{p+1}$ (where p is the vertical vector precision in bits).

So for vertical vectors measured to a precision of 1/16 picture line/field period $2^{4+1} = 32$ different coefficient sets are required. The $P+1$ (5 in this example) bits required to address the coefficient sets are the fractional vertical vector bits plus the least significant integer vertical vector bit. VAMCF can therefore be implemented according to Figure 2.

The filter cell shown in Figure 2 differs from that shown in figure 1 (described above) in that the coefficient table contains multiple filter apertures (sets of filter coefficients) and is, therefore, considerably larger and the filter aperture is selected by the vertical velocity component of the motion vector.

To implement, in this example, VAMCF rather than MCF, the size of the coefficient lookup table would increase from a 13 bit address (see above) to an 18 bit address. In an implementation which requires all the coefficients (from all apertures) to be stored in a single lookup table the address space would increase from 18 to 23 bits. In either case this is a significant increase in the size of lookup table required.

For VAMCF there may be an advantage to combining the vertical parts of the vector processing and the coefficient lookup table. This may reduce the size of lookup table required because the vertical part of the coefficient address and the aperture (coefficient set) address contain redundant information. Given the vertical motion speed and the temporal coefficient address, both the vertical coefficient address and the aperture (coefficient set) address can be found. This results in the implementation of figure 3.

The best implementation (figure 2 or figure 3) depends on the vertical aperture size and the maximum vertical velocity. The size of the lookup table (implemented in ROM) for figure 3, assuming a maximum vertical vector of +/- 32 pixels (ie. picture lines)/field period, is given by (in this example):

Horizontal	2 taps	16 phases	= 5 bits
Temporal	4 fields	16 phases	= 7 bits
Vertical	+ - 32 picture lines/field, precision 1/16=9 bits		
TOTAL	=21 bits		

which is 2 bits less than the implementation of figure 2.

The filter cells shown in Figures 2 and 3 contain modifications to the conventional motion compensated filter, shown in figure 1, which result in motion compensated filters/interpolators that adapt to vertical motion speed to cancel aliasing. The use of different interpolation apertures for different vertical velocities allows aliasing inherent in interlaced television signals to be cancelled, thereby preventing degradation of the output of motion compensated filtering/interpolation systems.

Motion compensated filtering requires a minimum filter aperture spanning at least 2 fields but better results can be achieved, as indicated in appendix 2, by using more than two fields.

Example 1

Interlace to progressive conversion for video compression.

Many video compression systems (e.g. JPEG, MPEG, sub-band coding, wavelet transform schemes) perform better with progressive rather than interlaced pictures. Many video sources, however, produce interlaced rather than progressively scanned video. Hence to get the best out of compression systems requires the conversion of interlaced video to a progressive format. Existing interlace to progressive converters, be they a fixed filter, motion adaptive, or motion compensated, are unable to perform this conversion without leaving considerable aliasing in the progressive output. This aliasing can significantly degrade the performance of the video compression system.

Using the improved motion compensated filtering of this invention, the quality of the interlace to progressive conversion can be significantly improved by reducing the unwanted aliasing in the progressive signal. This, in turn, allows improved quality or a reduced bit rate from the compression system. An interlace to

progressive converter is, conceptually, a filter as illustrated in figure 4.

The interlaced input is first padded with zeros (as is described in many texts on digital signal processing) to create a highly aliased progressive signal. The zero padded signal is then filtered to remove as much aliasing as possible. The filter can be a fixed linear filter, motion adaptive filter, non-linear (for example median) filter, or a conventional motion compensated filter. If an alias cancelling motion compensated filter is used (that is a VAMCF according to the invention) a substantially alias free progressive output can be produced.

The alias cancelling motion compensated filter (VAMCF) used for interlace to progressive conversion requires an aperture spanning at least two fields. However, this provides no immunity from signal noise or inaccuracies in the motion vectors used. A two field filter also has the disadvantage of either a non-zero group delay (producing progressive output frames half-way between input fields) which is inconvenient, or an asymmetric aperture which may produce inferior results.

A better implementation of alias cancelling motion compensated filtering (VAMCF) for interlace to progressive conversion would use a 3 (or more) field aperture. This would provide a symmetric aperture with zero group delay (output frames co-timed with input fields) and some immunity from input signal noise and motion vector errors. Although 3 fields provide more information than is necessarily required for interlace to progressive conversion (mathematically an over-determined problem), the output frames can be determined as at least mean square (LMS) error approximation to all 3 input fields. This LMS approximation (described in appendix 2) provides the robust characteristics described above. In the computation of the filter characteristic to give the "best" (LMS) output frame, it is not necessary to weight the errors from all 3 fields equally. A better solution would probably be to weight the errors from the central field of the aperture twice as heavily as the two flanking fields. This weighted LMS approach to filter design is described in appendix 2.

Obviously it is possible to implement a VAMCF with more than three fields in its aperture. As the number of input fields increases, immunity to input noise and motion vector errors increases, but so does hardware complexity and the

assumption of constant motion (implicit in motion compensation) across many fields is more likely to be violated. In practice a good compromise is probably to use a 3 or 5 field filter.

Example 2

Video to film conversion.

It is sometimes desirable to be able to convert a video signal for recording onto film (tele-recording). This might be done for archive purposes or for inserting a video sequence (perhaps with special effects) into a motion picture film.

A 50 fields/second video sequence could be recorded onto 25 frame/second film simply by recording 2 fields on each frame. Similarly 60 fields/second video could be recorded on to 30 frame/second film. However, recovering the original two fields from the recorded film frame would be very difficult because of registration problems and spatial distortions in both the optical recording and playback (telecine) operations. If the separate fields were confused in the replayed signal (as is likely) highly disturbing impairments would result in moving scenes. Even if the fields could be recovered correctly from the recorded film, motion rendition of paired fields would probably be unacceptable when the film was displayed using an optical projector.

To avoid the problems of recording paired fields it is necessary to convert interlaced signals to a progressive format. Single progressive frames can then be recorded on each film frame. This technique avoids the problems associated with recording paired fields described above. The interlace to progressive conversion process for tele-recording can be performed in exactly the same way as for video compression (example 1 above) using VAMCF to provide improved quality. The only difference with tele-recording is that, for example given a 50Hz interlaced source, only a 25Hz (rather than 50 Hz) progressive output is required for tele-recording.

It is probably also worth noting that video graphic post production systems often perform best with a progressively scanned input. This progressive input can be also obtained using a VAMCF, as described above, to give improved quality.

Example 3

Motion compensated colour decoding.

Many techniques are described in the literature for decoding composite colour television signals (PAL and NTSC). One of these variations is shown in figure 5.

This configuration contains two (different) baseband chrominance pass filters. The chrominance pass filters can be 1 dimensional (horizontal), 2 dimensional (spatial) or 3 dimensional (spatio-temporal) filters. Multi-dimensional linear filters are usually referred to as 'comb filters' in the context of colour decoding.

The comb filters in a colour decoding circuit can be replaced by conventional motion compensated spatio-temporal (comb) filters. This would produce a motion compensated colour decoder. This might provide improved performance over a non-motion compensated comb filter decoder. Unfortunately the output of the motion compensated comb filters would still include alias components which would impair chrominance/luminance separation. If instead a VAMCF was used to replace the comb filters in figure 5, aliasing at the output of filters would be reduced, resulting in improved luminance/chrominance separation. The filter coefficients for such a filter would be designed to take account of both the magnitude and phase of the (demodulated) composite signal, but a VAMCF would still be required to avoid aliasing in the filter output.

Example 4

Motion estimation using interlaced video signals.

In many techniques for video motion estimation (references 1, 2, 3, 9, 11, 13, 4,19 & 21), it is necessary to compare displaced versions of input pictures to test the validity of a possible motion vector. Typically, the current input picture is subtracted from a displaced version of the preceding or subsequent picture of a sequence. Unfortunately, if the input is an interlaced sequence, the displaced field difference signal contains significant energy due to aliasing implicit in the interlaced signal. This alias energy cannot be removed by intrafield filtering prior to motion estimation.

Aliased energy can confuse a motion estimator, causing it to generate erroneous motion vectors with deleterious effects to down stream processing using the motion vectors.

By using alias cancelling motion compensated filtering (VAMCF), substantially alias free displaced field difference signals can be generated for use in motion estimation. This affords improved motion estimation by reducing the chance of the motion estimator being confused by aliasing. In order to generate alias cancelled displaced field difference signals it is necessary to use at least three fields rather than the usual 2.

A conventional forward (or backward) field difference signal is generated by a conventional motion compensated filter with a 2 field aperture. The aperture used is a field difference, that is a temporal high pass filter, and may include spatial (intra-field) interpolation. The motion vector used is the vector under test. This process could, alternatively, be considered as first calculating the mean of the two fields and then subtracting the mean field from each of the two input fields. This gives 2 (identical) error signals. If the fields were progressively scanned, rather than being interlaced, the mean field would be a least mean square (LMS) best fit to the two input fields.

To generate an alias cancelled field difference, conceptually one first calculates a LMS estimate of a progressive frame from 3 or more input fields. Next a displaced field difference signal can be calculated from that progressive frame, for each of the input fields in the aperture.

Typically one might use 3 fields (i.e. fields 0,1 and 2 illustrated in figure 6), although more fields (but not less) are possible. With a 3 field aperture, 3 distinct field difference signals can be calculated rather than the single unique displaced field difference of the conventional approach. Any one of these field differences could be used in place of the conventional displaced field difference. Alternatively the multiple field differences, provided by the alias cancelling approach, could be combined together (for example by taking sum of squares) to provide a more useful measure of the validity of a vector than could be obtained using the single field difference alone.

The calculation of alias cancelled field differences can be implemented using VAMCF as described earlier. To calculate each field difference requires a separate (spatio-temporal) filter. The filter design technique for this application is outlined in the appendix 2.

Example 5

Television standards conversion.

Television standards conversion is the process of converting between different television standards, for example between 625 line 50Hz PAL and 525 line 60Hz NTSC. Assuming colour decoding and re-encoding are performed separately, standards conversion is a 3 dimensional (spatio-temporal) interpolation problem. The theory and techniques of standards conversion are described in detail in references 3, 6 & 7. Motion compensation has been shown to provide improved quality compared to non-motion compensated standards conversion. Conventional MCF, however, still leaves some aliasing in the output pictures with consequent degradation to picture quality. This degradation would be particularly evident for slowly moving, high resolution images such as computer graphics. These sort of pictures could be processed admirably using motion compensation were it not for this residual aliasing.

VAMCF can be directly applied to television standards conversion. In this application removing aliasing results in a direct improvement in picture quality. Standards conversion includes a temporal interpolation process. It is necessary to generate output fields at different times from the input fields. This is in contrast to the filter examples presented above where output fields are co-timed with the input fields. The design of filter coefficients for temporal interpolation is more complex than for filtering and is outlined in appendix 2.

Example 6

Slow motion displays.

It is often desirable to replay a video signal more slowly than the original action. A common example of this is 'action replays' in television sports coverage. One way to achieve this is to simply repeat fields or frames as necessary to provide a

sufficient picture rate. This technique is commonly used in video recorders to achieve slow motion replay. Unfortunately it results in an unnatural and disturbing 'jerky' motion rendition. Motion compensation can be used to provide improved motion rendition as described in reference 18. Conventional motion compensation, however, leaves residual aliasing in the output signal. In a slow motion replay system, aliasing from vertical detail (for example the top edge of a wall) can result in wobbling detail on an otherwise stationary image. This can be visually very disturbing.

VAMCF can be used for motion compensated slow motion replay in the same way as for television standards conversion. The theory is essentially identical but with a greater number of output fields per input field. In this case, the removal of unwanted aliasing is even more important than for standards conversion because the aliasing is more visible on the slowed down signal.

Further details relating to methods for the implementation of motion compensated processing of interlaced video signals for improved picture quality are provided in Schedule "A" following this specification.

Appendix 1

References

1. Aggarwal, J.K. & Nandhakumar, N. 1988. On the computation of motion from sequences of images - a review. Proc. IEEE, vol. 76, pp. 917-935, August 1998.
2. Bierling, M., Thoma, R. 1986. Motion compensating field interpolation using a hierarchically structured displacement estimator. Signal Processing, Volume 11, No. 4, December 1986, pp. 387-404. Elsevier Science publishers.
3. Borer, T.J. , 1992. Television Standards Conversion. Ph.D. Thesis, Dept. Electronic & Electrical Engineering, University of Surrey, Guildford, Surrey, GU2 5XH, UK. October 1992.
4. Borer, T.J., Hulyer, M.G. & Parker, D.W. 1990. Motion compensated display field rate upconversion. International Broadcasting Convention, Brighton, England, September 1990. IEE Publication No. 327. ISBN 0 85296000
5. Chiariglione, L., Corgnier, L. & Guglielmo, M. 1986. Pre- and post- processing in a video terminal using motion vectors. IBC. Brighton 1986.
6. Clarke, C.K.P. & Tanton, N.E. 1984. Digital standards conversion: interpolation theory and aperture synthesis. BBC Research Department Report 1984/20.
7. Clarke, C.K.P., 1990. Interpolation. Digital television, Sandbank, C.P. (Editor) , John Wiley & Sons Ltd. 1990. ISBN 0 471 92360 5, pp. 287-374.
8. Crochiere, R.E. & Rabiner, L.R. 1983. Multirate Digital Signal Processing. Prentice Hall, Inc., Englewood Cliffs, New Jersey 07632, USA. ISBN --13- 605162-6.
9. Dubois, E., Konrad, J. , 1990. Review of techniques for motion estimation and motion compensation. Forth international colloquium on advanced television systems, Ottawa, Canada, June 1990.
10. Girod, B. & Thoma, R. 1985. Motion-compensating field interpolation from

interlaced and non-interlaced grids. 2nd international technical symposium on optical and electro-optical science and engineering: Image processing symposium, Cannes, December 1985.

11. Huange, T.S., Tsai, R.Y., 1981. Image sequence analysis: Motion estimation. Image sequence analysis, T.S. Huange (editor), Spnngge-Verlag, Berlin, Germany, 1981, pp. 1-18.
12. Marcozzi, E. & Tubaro, S. 1987. Motion compensated interpolation. SPIE, Vol. 804, Advances in image processing, pp 61-68.
13. Martinez, D.M. 1987. Model-based motion estimation and its application to restoration and interpolation of motion pictures. RLE Technical Report No.530. June 1987. Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.
14. Netravali, A.N., Robbins, J.D. 1979. Motion compensated television coding, Part 1. Bell Syst. Tech. J., vol. 58, pp 631-670, March 1979.
15. Parks, T.W. & Burrus, C.s. 1987. Digital Filter Design. John Wiley & Sons 1987, ISBN 0-471-82896-3.
16. Reuter, T. 1989. Standards conversion using motion compensation. Signal processing 16,1989, pp 73-82.
17. Thoma, R. & Bierling, M. 1989. Motion compensated interpolation considering covered and uncovered background.
18. Thomas, G.A., & Lau, H.Y.K. 1990. Generation of high quality slow-motion replay using motion compensation. International Broadcasting Convention, Brighton, England. IEE Publication No. 327. ISBN 0 85296000.
19. Thomas, G.A., 1987. Television motion estimation for DAtHl and other applications. BBC Research Department report No. 1987/11. September 1987.

20. Thomas, G.A., 1988. Distorting the time axis: Motion compensated image processing in the studio. International Broadcasting Convention, Brighton, England. September 1988. Published by IEE ISBN 0 85296368 8.
21. Vega-riveros, J.F., Jabbour, K. 1986. Review of motion analysis techniques. IEE Proceedings, Vol. 136, Pt I, No. 6, December 1989.

Appendix 2

For simplicity this appendix only considers the vertical and temporal dimensions. Extension to include the horizontal dimension is straight forward for those skilled in the art.

Consider an object, denoted h , moving with a velocity, v , to produce a moving image, denoted g .

$$g(y,t)=h(y-v.t) \quad \text{Equation 1}$$

In a television system the continuous, spatio-temporal, image g is sampled on an interlaced sampling lattice. The spatial position of the sampling grid alternates between odd and even positions on consecutive fields. Two sampled signals, S , are possible for each field position depending on whether the image, g , is sampled with an even or odd field lattice. This results in two distinct possible signals for each field denoted S_e and S_o . Taking the Fourier transform of an even or odd sampled signal gives the following result.

$$S_e(n,t) = \mathfrak{F}_y(s_e(y,t)) = H(n) \cdot \exp(-j2\pi nvt) + H\left(n - \frac{1}{2Y}\right) \exp\left(-j2\pi\left(n - \frac{1}{2Y}\right)vt\right)$$

$$S_o(n,t) = \mathfrak{F}_y(s_o(y,t)) = H(n) \cdot \exp(-j2\pi nvt) - H\left(n - \frac{1}{2Y}\right) \exp\left(-j2\pi\left(n - \frac{1}{2Y}\right)vt\right)$$

Equation 2

In these equations n represents vertical frequency, script F (subscript y) represents a Fourier transform (with respect to y only), H is the Fourier transform (complex spectrum) of the object, h , and Y is picture line spacing (field line spacing is $2Y$). The equations consider first order aliasing only and are valid for positive frequencies only (similar equations apply for negative frequencies).

Consider the example of interlace to progressive conversion using 2 input fields. Assume the two input fields are at $t = 0$ and $t = T$ (T is the field period) and that the first field is even and the second field odd. Taking the Fourier transforms of the two fields yields the following matrix equation.

$$\begin{bmatrix} S_e(n, t=0) \\ S_o(n, t=T) \end{bmatrix} = \begin{bmatrix} 1 & \\ \exp(-j2\pi n v T) & -\exp\left(-j2\pi\left(n - \frac{1}{2Y}\right)vT\right) \end{bmatrix} \begin{bmatrix} H(n) \\ H\left(n - \frac{1}{2Y}\right) \end{bmatrix}$$

or

$$\mathbf{S} = \mathbf{P} \cdot \mathbf{H}$$

Equation 3

where S is a vector of the spectra of sampled fields, P is a matrix of phase factors (similar to a discrete Fourier transform matrix) and H is a vector of the object spectrum and alias spectrum.

In calculating the progressive frame from the interlaced fields we are trying to find the unsampled image at $t = 0$ from the sampled (interlaced) fields. The unsampled image we require is $h(y)$. By solving equation 3 we can find $H(n)$ (and also $H(n - 1/2Y)$, the aliasing, which we do not require) and then $h(y)$ by an inverse Fourier transform.

$$\mathbf{H} = \mathbf{P}^{-1} \cdot \mathbf{S}$$

or

$$\begin{bmatrix} H(n) \\ H\left(n - \frac{1}{2Y}\right) \end{bmatrix} = \frac{1}{1 + \exp\left(j\frac{\pi v T}{Y}\right)} \begin{bmatrix} \exp\left(j\frac{\pi v T}{Y}\right) & \exp(j2\pi n v T) \\ 1 & -\exp(j2\pi n v T) \end{bmatrix} \begin{bmatrix} S_e(n, t=0) \\ S_o(n, t=T) \end{bmatrix}$$

Equation 4

Considering the expression for $H(n)$ only we have:

$$H(n) = \frac{\exp\left(j\frac{\pi v T}{Y}\right)}{1 + \exp\left(j\frac{\pi v T}{Y}\right)} \cdot S_e(n, t=0) + \frac{\exp(j2\pi n v T)}{1 + \exp\left(j\frac{\pi v T}{Y}\right)} \cdot S_o(n, t=T)$$

Equation 5

Inverse Fourier transforming equation 5 gives:

$$h(y) = \text{filter}_0(y, v) * S_e(y, t=0) + \text{filter}_1(y, v) * S_o(y, t=T) = \text{filter}(y, t, v) * S(y, t) \quad \text{Equation 6}$$

where * represents the convolution operation.

Equation 6 shows that we can find the progressive image, $h(y)$, that we require by filtering each of the interlaced fields separately using intrafield filters and then summing the results. Summing two intrafield filters on consecutive fields can be considered as using a single spatio-temporal filter. The intrafield filters used are given by the inverse Fourier transforms of the coefficients of S_e and S_o in equation 5. Note that the filters in equation 6 are functions of v , the vertical motion speed.

The filters of equation 6 represent an idealised situation. The filters have infinite extent and therefore cannot be realised in practice. However, the coefficients of S_e and S_o in equation 5 represent the frequency domain specifications for the ideal interlace to progressive conversion filters. There are many techniques in the literature (see for example reference 15) for designing practical filters from an ideal frequency domain specification.

The performance of an interlace to progressive converter can be improved by using more than a two field aperture as discussed in the main text. Consider a 3 field aperture with fields at $t = -T, 0$ & $+T$. Equation 3 then becomes:

$$\begin{bmatrix} S_e(n, t = -T) \\ S_e(n, t = 0) \\ S_e(n, t = +T) \end{bmatrix} = \begin{bmatrix} \exp(+j2\pi n v T) & -\exp\left(+j2\pi\left(n - \frac{1}{2Y}\right)vT\right) \\ 1 & 1 \\ \exp(-j2\pi n v T) & -\exp\left(-j2\pi\left(n - \frac{1}{2Y}\right)vT\right) \end{bmatrix} \begin{bmatrix} H(n) \\ H\left(n - \frac{1}{2Y}\right) \end{bmatrix} \quad \text{Equation 7}$$

or

$$S = P.H$$

In this case the phase matrix is rectangular, not square, and it is no longer possible to directly invert equation 7 to find $H(n)$. It is, however, possible to find a least mean square approximation to $H(n)$ by solving the 'normal' equations given by:

$$Ph.S = (Ph.P)\tilde{H} \quad \text{Equation 8}$$

Here the symbol, h , represents the Hermitian conjugate (complex conjugate of the transpose) of the matrix. The above H indicates that H is a least mean square approximation. Solving for H gives:

$$\tilde{H} = (Ph.P)^{-1}.Ph.S \quad \text{Equation 9}$$

The LMS solution for H again gives a frequency domain specification for the aliasing cancelling filters, practical versions of which can then be calculated using many published techniques.

With a 3 field aperture, for interlace to progressive conversion, it may be appropriate to give greater prominence to the centre field in the aperture. For example it might be appropriate to weight the 2 outer odd fields half as much as the central even field. This can be done by multiplying equation 7 by an appropriate weighting matrix for example:

$$\text{where } W = \begin{matrix} \text{W.S.} = \text{W.P.H.} \\ \left[\begin{array}{ccc} 1/4 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/4 \end{array} \right] \end{matrix} \quad \text{Equation 10}$$

Equation 10 can then be solved to find a weighted least mean square solution for H:

$$\tilde{H} = ((Ph.W^2.P)^{-1}.Ph.W^2)S \quad \text{Equation 11}$$

Again practical filters can be designed from the frequency domain specifications embodied in equation 11.

The design of interpolation filters for standards conversion is similar to the design of interlace to progressive conversion filters. The process of standards conversion could be looked on as calculating a sequence of progressive frames from an interlaced input. Typically an interlaced output is required from a standards converter. This is easily achieved, conceptually, by simply spatially subsampling the progressive fields appropriately to give an interlaced signal. In practice this means that only half the number of output points need be calculated with a corresponding saving in the hardware required.

For standards conversion, in contrast to interlace to progressive conversion, output fields are required at different time instants to the input fields. The effect of this is to change the phase terms in the phase matrix P. The elements of the

phase matrix are still calculated using equation 2 as before. For standards conversion the values of t used are no longer integer multiples of a field period. Other than this small change the equations above still apply. Hence filter design for standards conversion is, in essence, very similar to that for interlace to progressive conversion.

In motion estimation it is often required to test the validity of a possible motion vector by comparing consecutive fields. Unfortunately this comparison is, conventionally, complicated by the interlaced nature of video signals. Improved results are possible by correctly allowing for the interlaced nature of the video signal.

Conceptually the comparison of consecutive fields, allowing for interlace, is a two stage process. First a 3 (or more) field aperture is used to calculate a least mean square (LMS) estimate of a progressive frame assuming the test motion vector is correct. Then the input fields are compared to the expected fields given that the progressive frame and motion vector are correct and allowing for interlace. The least mean square estimate of (the complex spectrum of) the progressive frame is given by equation 12.

$$\tilde{H} = (Ph.P)^{-1}.Ph.S \quad \text{Equation 12}$$

If the estimated spectrum, H , is correct then the input fields would be given by:

$$\tilde{S} = P.\tilde{H} \quad \text{Equation 13}$$

The differences between the actual fields, S , and the estimated fields, \tilde{S} , is given by:

$$S - \tilde{S} = (I - P.(Ph.P)^{-1}.Ph)S \quad \text{Equation 14}$$

Equation 14 gives a frequency domain specification, for an alias cancelled field difference signal, for each field in the aperture. Practical filters can then be designed from the frequency domain specifications using published techniques.

I CLAIM:

1. A filter or interpolation device for motion compensated filtering or interpolation of interlaced video signals and comprising a plurality of filter taps and implemented between the filter taps a filter aperture which varies and is selected depending on vertical velocity components of motion vectors used in the filtering or interpolation.
2. A filter or interpolation device as claimed in claim 1, wherein each filter tap comprises a multiplier and a corresponding coefficient store.
3. A filter device as claimed in claim 1 or 2 comprising a coefficient look-up table stored in memory, wherein the total number of different coefficient sets available is given by 2^{p+1} , where p is the vertical vector precision in bits.
4. A video processing system for motion compensated filtering of interlaced video signals comprising a motion estimation device for producing motion vectors corresponding to each of a plurality of output sampling sites, a vector processor generating a set of input coordinates for each output sampling site which input coordinates are selected according to a motion vector corresponding to the set of input coordinates, a data store in memory for storing input pixel values and providing a variable delay, a plurality of multipliers and coefficient stores corresponding to the multipliers, and an adder to sum the partial result from each multiplier, wherein the multipliers multiply an input pixel value selected from the data store depending on a part of the input coordinate by a coefficient selected from the coefficient store depending on another part of the input coordinate, and wherein the filter aperture is selected depending on the vertical velocity component of the motion vector supplied by the motion estimation device.
5. A video processing system as claimed in claim 4, comprising a coefficient look-up table stored in memory, wherein the total number of different coefficient sets available is given by 2^{p+1} , where p is the vertical vector precision in bits.

6. A video processing method for motion compensated filtering or interpolation of interlaced video signals comprising assigning a motion vector to each output sampling site, producing a set of input coordinates and selecting a filter aperture depending on the vertical velocity component of the motion vector associated with each sampling site, such that different filter apertures are used for different vertical velocities.
7. A video processing method for processing interlaced video signals as claimed in claim 6, the filter aperture is selected from multiple pre-calculated filter apertures depending on the vertical velocity component of each motion vector, and the method comprising selecting a coefficient from the filter aperture depending on the fractional part of each input coordinate, selecting an input pixel value depending on the integer part of each input coordinate, and multiplying each input pixel value by the correspondingly selected coefficient to produce a partial result for each output sampling site and summing the partial results.
8. A video processing method for processing interlaced video signals as claimed in claim 6 or 7, wherein the number of precalculated filter apertures is dependent on the precision of the vertical component of the motion vector.

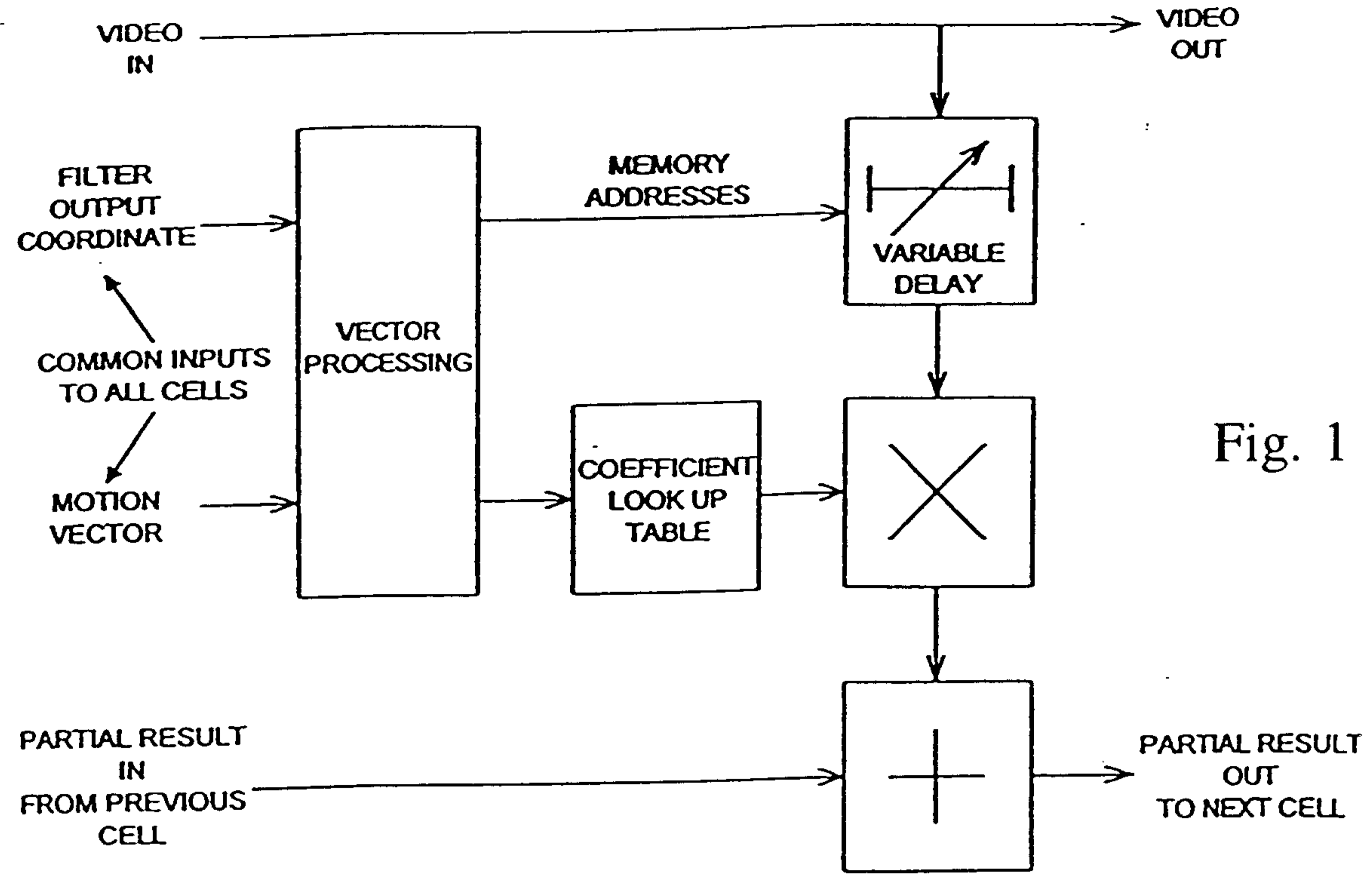


Fig. 1

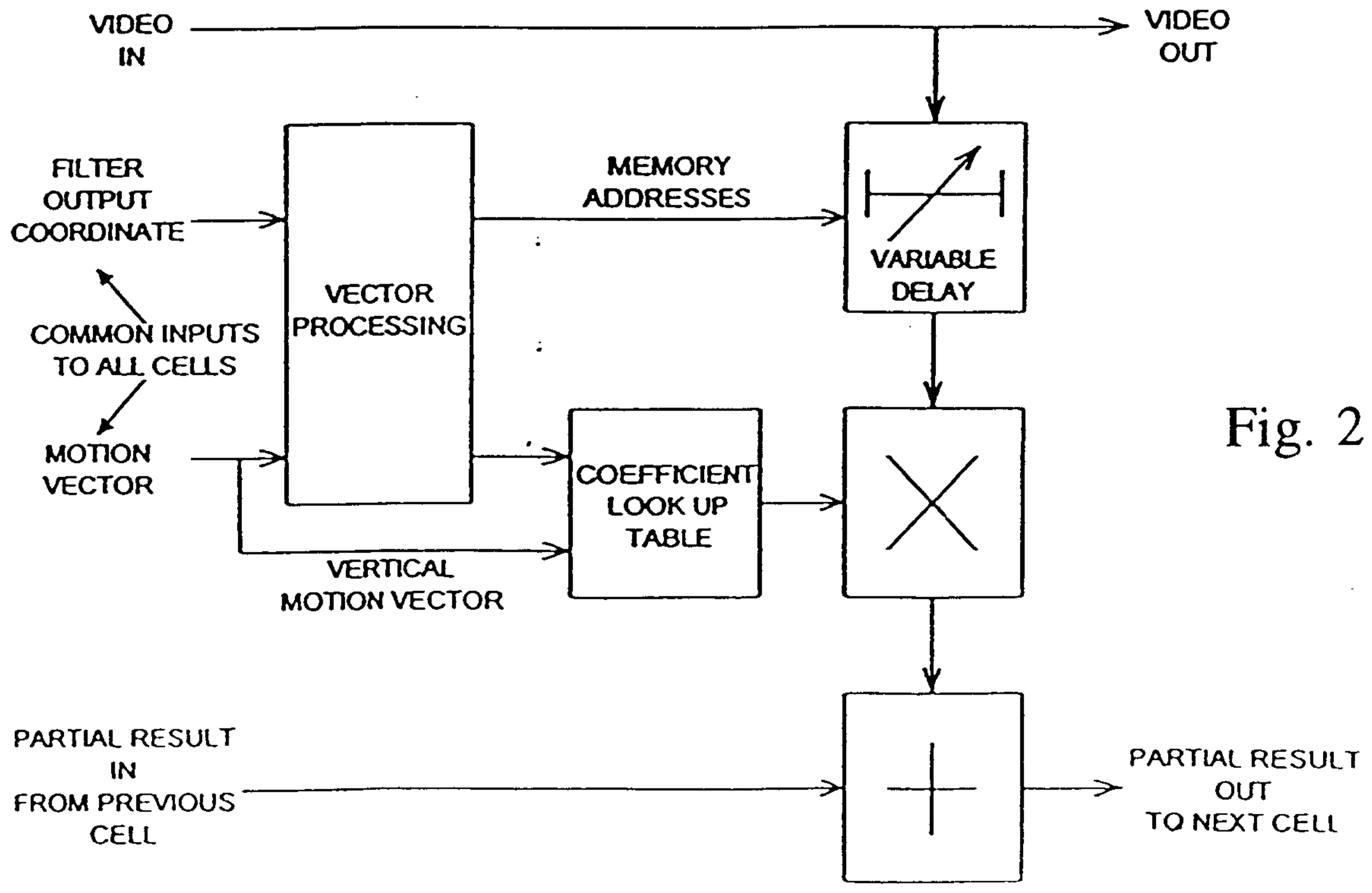


Fig. 2

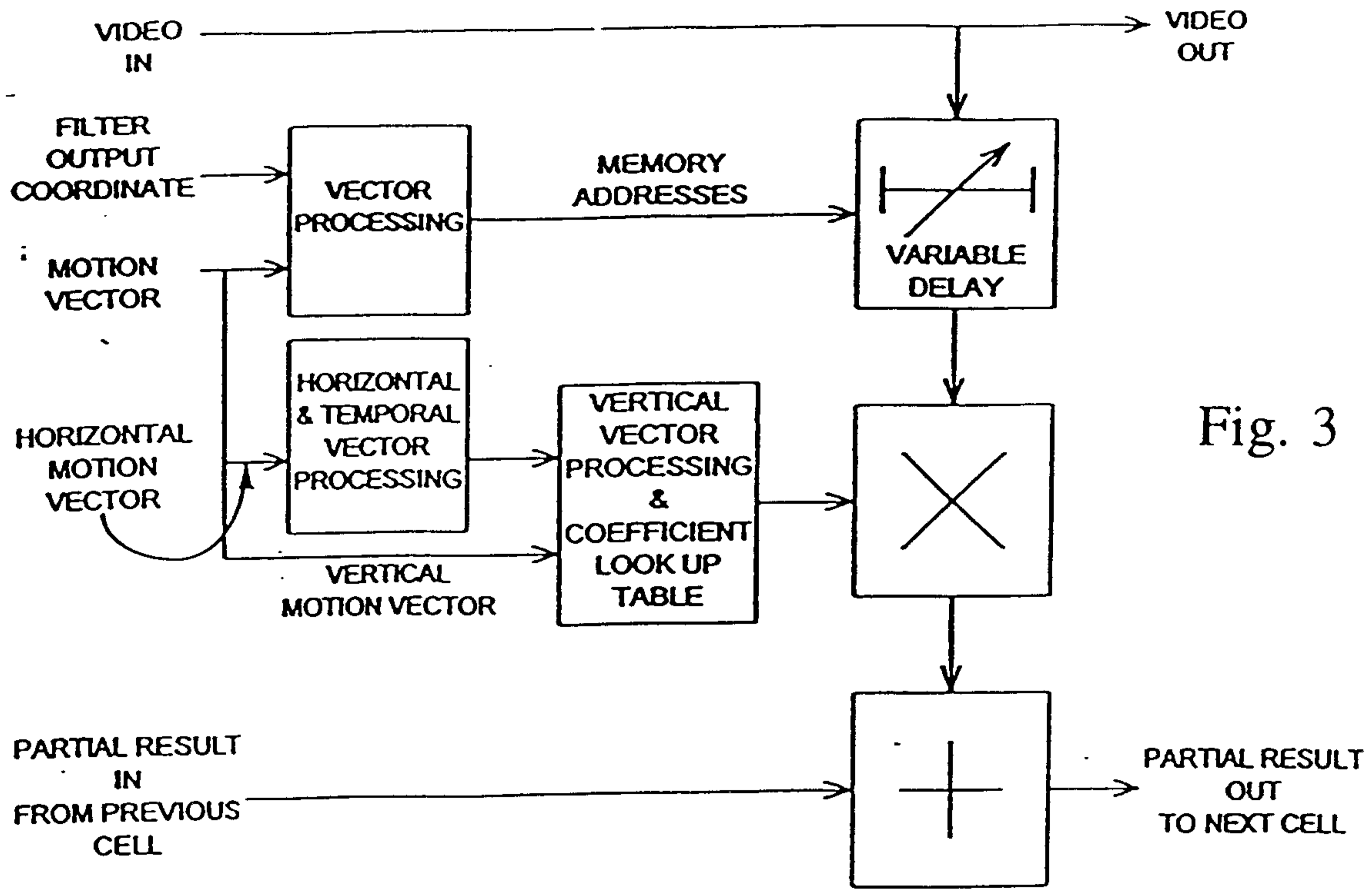


Fig. 3

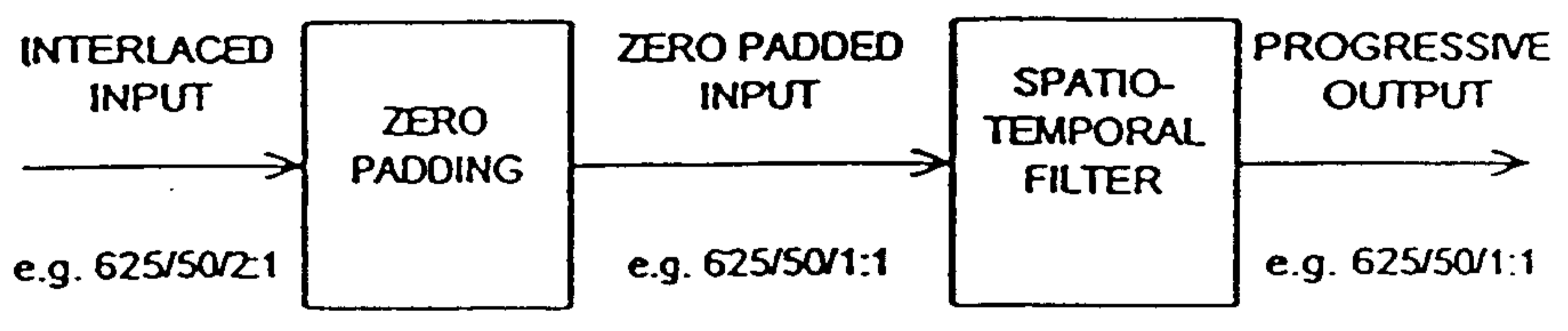


Fig. 4

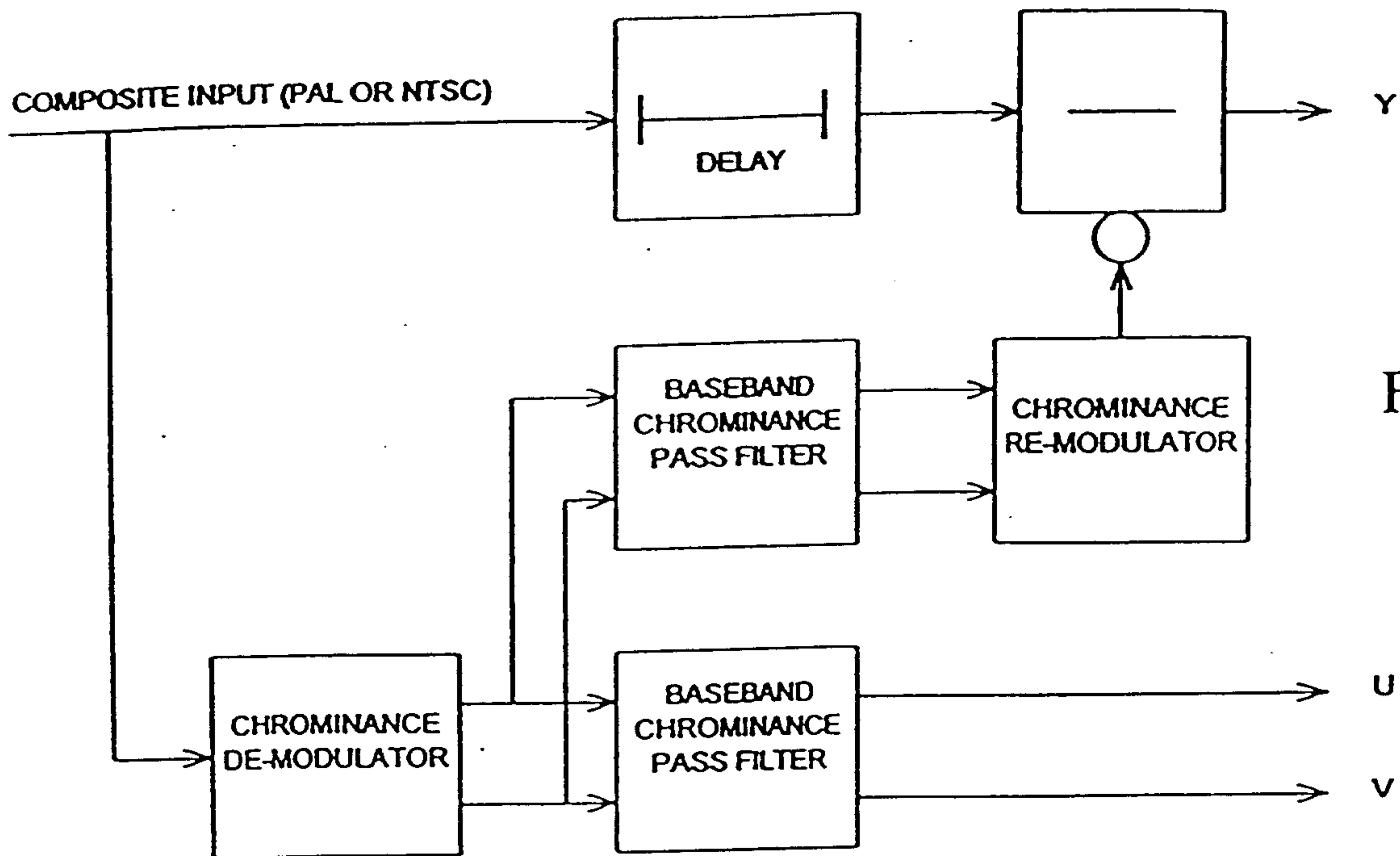


Fig. 5

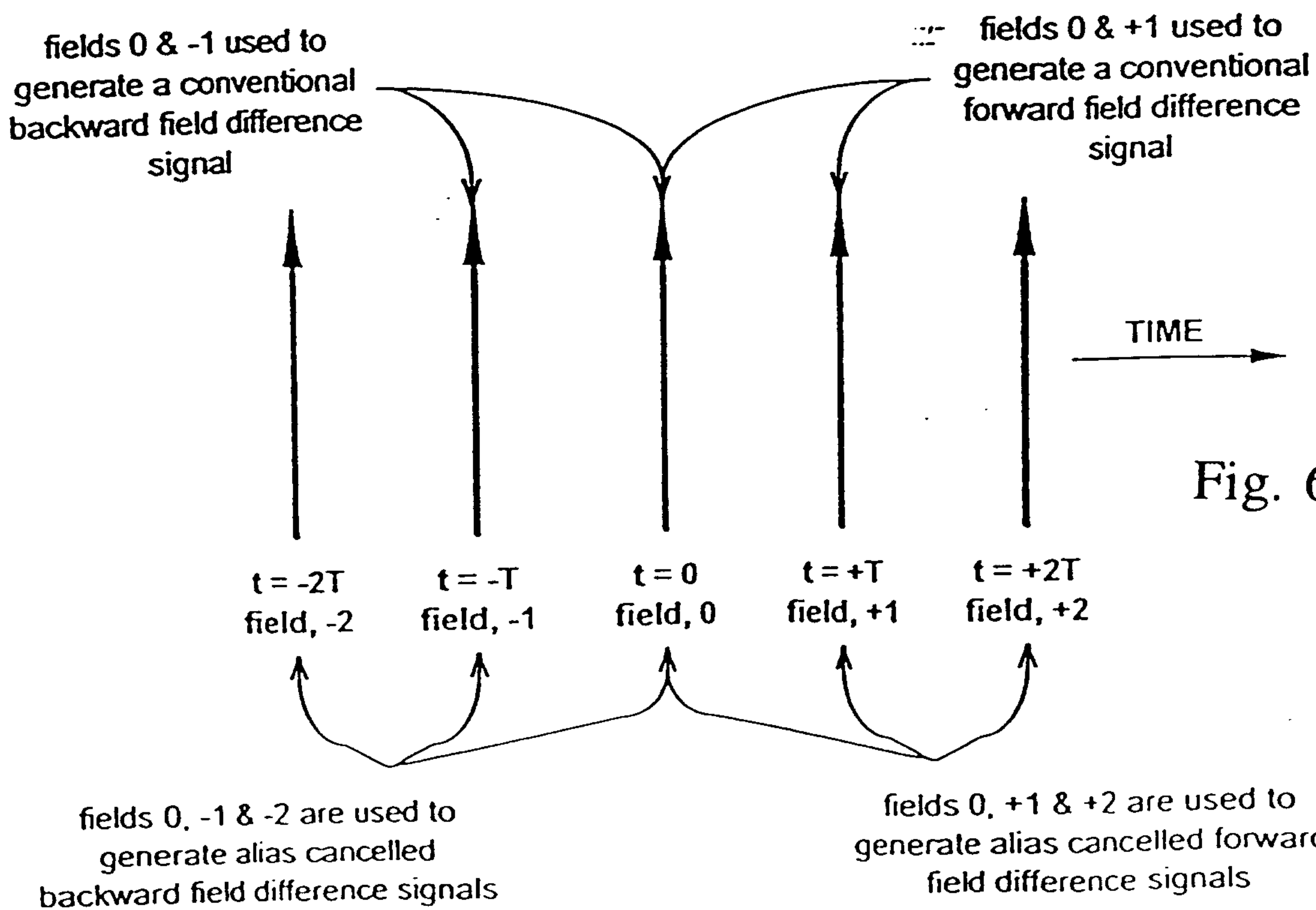


Fig. 6

