

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
19 December 2002 (19.12.2002)

PCT

(10) International Publication Number  
WO 02/102049 A2

(51) International Patent Classification<sup>7</sup>: H04N

(21) International Application Number: PCT/US02/18458

(22) International Filing Date: 11 June 2002 (11.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/296,766 11 June 2001 (11.06.2001) US  
60/296,768 11 June 2001 (11.06.2001) US  
10/170,019 11 June 2002 (11.06.2002) US

(71) Applicant: BROADCOM CORPORATION [US/US];  
16215 Alton Parkway, Irvine, CA 92618 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): MORAD, Amir  
[—/US]; 16215 Alton Parkway, Irvine, CA 92618 (US).  
YAVITS, Leonid [—/US]; 16215 Alton Parkway, Irvine,  
CA 96218 (US). OXMAN, Gadi [—/US]; 16215 Alton

Parkway, Irving, CA 92618 (US). SPEKTOR, Evgeny  
[—/US]; 16215 Alton Parkway, Irving, CA 92618 (US).  
KHRAPKOVSKY, Michael [—/US]; 16215 Alton  
Parkway, Irvine, CA 92618 (US). CHERNOV, Gregory  
[—/US]; 16215 Alton Parkway, Irving, CA 92618 (US).

(74) Agent: WINSLADE, Christopher, C.; McAndrew Held  
& Malloy, Ltd., 500 W. Madison Street, Suite 3400,  
Chicago, IL 60661 (US).

(81) Designated State (national): US.

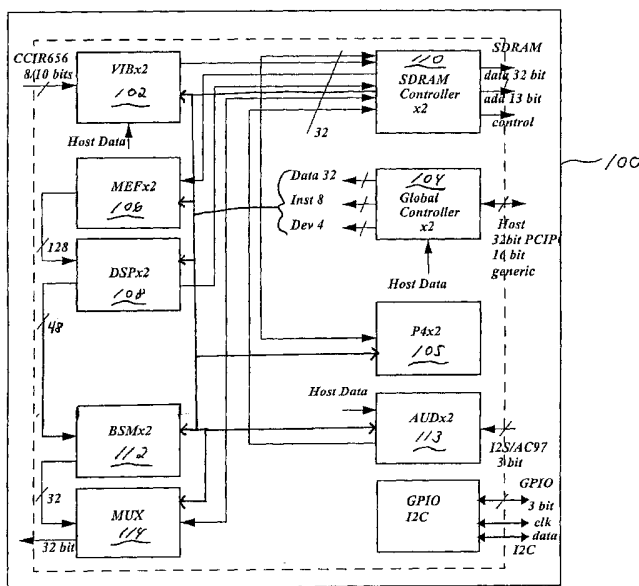
(84) Designated States (regional): European patent (AT, BE,  
CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,  
NL, PT, SE, TR).

Published:

— without international search report and to be republished  
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

(54) Title: SYSTEM AND METHOD FOR MULTI-CHANNEL VIDEO AND AUDIO ENCODING ON A SINGLE CHIP



(57) Abstract: An apparatus is disclosed for performing real time multi-channel video/audio encoding on a single chip. Within the single chip, a multi-channel video encoder generates at least two channels of encoded video data from uncompressed video data and a multi-channel audio encoder generates at least two channels of encoded audio data from uncompressed audio data. A mux processor within the single chip generates at least two multiplexed output streams of data from the at least two channels of encoded video data and the at least two channels of encoded audio data.

WO 02/102049 A2

## SYSTEM AND METHOD FOR MULTI-CHANNEL VIDEO AND AUDIO ENCODING ON A SINGLE CHIP

### RELATED APPLICATIONS

This application is a continuation-in-part of United States Patent Application Serial No. 09/543,904 filed April 6, 2000, which claims the benefit of Israel Application Serial No. 129345 filed April 6, 1999.

This application also makes reference to, claims priority to and claims the benefit of United States Provisional Patent Application Serial No. 60/296,766 filed on June 11, 2001 and United States Provisional Patent Application Serial No. 60/296,768 filed on June 11, 2001.

All of the above-listed patent applications are incorporated herein by reference in their entirety.

### BACKGROUND OF THE INVENTION

Methods for encoding an audio-visual signal are known in the art. According to the methods, a video signal is digitized, analyzed and encoded in a compressed manner. The methods are implemented in computer systems, either in software, hardware or combined software-hardware forms.

Most hardware encoding systems consist of a set of semiconductor circuits arranged on a large circuit board. State of the art encoding systems include a single semiconductor circuit. Such a circuit is typically based on a high-power processor.

Reference is now made to Fig. 1, which is a block diagram illustration of a prior art video encoding circuit 10.

Encoding circuit 10 includes a video input processor 12, a motion estimation processor 14, a digital signal processor 16 and a bitstream processor 18. Processors 12-18, respectively, are generally connected in series.

Video input processor 12 captures and processes a video signal, and transfers it to motion estimation processor 14. Motion estimation processor 14 analyzes the motion of the video signal, and transfers the video signal and its associated motion analysis to digital signal processor 16. According to the data contained within the associated motion analysis, digital signal processor 16 processes and compresses the video signal, and transfers the compressed data to bitstream processor 18. Bitstream processor 18 formats the compressed data and creates therefrom an encoded video bitstream, which is transferred out of encoding circuit 10.

It will be appreciated by those skilled in the art that such an encoding circuit has several disadvantages. For example, one disadvantage of encoding circuit 10 is that bitstream processor 18 transfers the encoded video bitstream, data word by data word, directly to an element external to encoding circuit 10. Accordingly, each time such data word is ready, the encoded video data word is individually transferred to the external element. Transfer of the encoded video in such a fashion greatly increases the data traffic volume and creates communication bottlenecks in communication lines such as computer buses. Additionally, circuit 10 requires a dedicated

storage/bus which is allocated on a full time basis, hence, magnifying these disturbances.

Another disadvantage is that encoding circuit 10 is able to perform the encoding of video signals, only. Usually, moving picture compression applications include multiframe videos and their associated audio paths. While the encoding circuit 10 performs video compression and encoding, the multiplexing of compressed video, audio and user data streams are performed separately. Such an approach increases the data traffic in the compression system and requires increased storage and processing bandwidth requirements, thereby greatly increasing the overall compression system complexity and cost.

Reference is now made to Fig. 2, which is a block diagram of a prior art video input processor 30, as may be typically included in encoding circuit 10. Video input processor 30 includes a video capture unit 32, a video preprocessor 34 and a video storage 36. The elements are generally connected in series.

Video capture unit 32 captures an input video signal and transfers it to video preprocessor 34. Video preprocessor 34 processes the video signal, including noise reduction, image enhancement, etc., and transfers the processed signal to the video storage 36. Video storage 36 buffers the video signal and transfers it to a memory unit (not shown) external to video input processor 30.

It will be appreciated by those skilled in the art that such a video input processor has several disadvantages. For example, one disadvantage of processor 30 is that it does not perform image resolution scaling. Accordingly, only original resolution pictures can be processed and encoded.

Another disadvantage is that processor 30 does not perform statistical analysis of the video signal, since in order to perform comprehensive statistical analysis a video feedback from the storage is necessary, thus allowing interframe (picture to picture) analysis, and processor 30 is operable in "feed forward" manner, only. Accordingly, video input processor 30 cannot detect developments in the video contents, such as scene change, flash, sudden motion, fade in/fade out, etc.

Reference is now made to Fig. 3 which is a block diagram illustration of a prior art video encoding circuit 50, similar to encoding circuit 10, however, connected to a plurality of external memory units. As an example, Fig. 3 depicts circuit 50 connected to a pre-encoding memory unit 60, a reference memory unit 62 and a post-encoding memory unit 64, respectively. Reference is made in parallel to Fig. 4, a chart depicting the flow of data within circuit 50.

Encoding circuit 50 includes a video input processor 52, a motion estimation processor 54, a digital signal processor 56 and a bitstream processor 58. Processors 54 to 58, respectively, are generally connected in series.

In the present example, video encoding circuit 50 operates under MPEG video/audio compression standards. Hence, for purposes of clarity, reference to a current frame refers to a frame to be encoded. Reference to a reference frame refers to a frame that has already been encoded and reconstructed, preferably by digital signal processor 56, and transferred to and stored in reference memory unit 62. Reference frames are compared to current frames during the motion estimation task, which is generally performed by motion estimation processor 54.

Video input processor 52 captures a video signal, which contains a current frame, or a plurality of current frames, and processes and transfers them to external pre-encoding memory unit 60. External pre-encoding memory unit 60 implements an input frame buffer (not shown) which accumulates and re-orders the frames according to the standard required for the MPEG compression scheme.

External pre-encoding memory unit 60 transfers the current frames to motion estimation processor 54. External reference memory unit 62 transfers the reference frames also to motion estimation processor 54. Motion estimation processor 54, reads and compares both sets of frames, analyzes the motion of the video signal, and transfers the motion analysis to digital signal processor 56.

Digital signal processor 56 receives the current frames from the external pre-encoding memory 60, and according to the motion analysis received from motion estimation processor 54, processes and compresses the

video signal. Digital signal processor 56 then transfers the compressed data to the bitstream processor 58. Digital signal processor 56 further reconstructs the reference frame and stores it in reference memory 62. Bitstream processor 58 encodes the compressed data and transfers an encoded video bitstream to external post-encoding memory unit 64.

It will be appreciated by those skilled in the art that such an encoding circuit has several disadvantages. For example, one disadvantage of encoding circuit 50 is that a plurality of separate memory units are needed to support its operations, thereby greatly increasing the cost and complexity of any encoding system based on device 50.

Another disadvantage is that encoding circuit 50 has a plurality of separate memory interfaces. This increases the data traffic volume and the number of external connections of encoding circuit 50, thereby greatly increasing the cost and the complexity of encoding circuit 50. Another disadvantage is that encoder circuit 50 does not implement video and audio multiplexing, which is typically required in compression schemes.

Further limitations and disadvantages of conventional and traditional approaches will become apparent to one of skill in the art, through comparison of such systems with embodiments of the present invention as set forth in the remainder of the present application with reference to the drawings.

### BRIEF SUMMARY OF THE INVENTION

Certain embodiments of the present invention provide an apparatus for performing video and audio encoding. In particular, certain embodiments provide for performing multi-channel video and audio encoding on a single chip.

Apparatus of the present invention provides for performing real time multi-channel video/audio encoding on a single chip. Within the single chip, a multi-channel video encoder generates at least two channels of encoded video data from uncompressed video data and a multi-channel audio encoder generates at least two channels of encoded audio data from uncompressed audio data. A mux processor within the single chip generates at least two multiplexed output streams of data from the at least two channels of encoded video data and the at least two channels of encoded audio data.

These and other advantages and novel features of the present invention, as well as details of an illustrated embodiment thereof, will be more fully understood from the following description and drawings.

### BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block diagram of a prior art video encoding circuit.

Fig. 2 is a block diagram of a prior art video input processor.

Fig. 3 is a block diagram of a prior art video encoding circuit linked to a plurality of external memory units.

Fig. 4 is a flow chart of the data flow within the prior art circuit illustrated in Fig. 3.

Fig. 5 is a block diagram of a video and audio encoding video/audio/data multiplexing device constructed and operative on a single chip in accordance with an embodiment of the present invention.

## DETAILED DESCRIPTION OF THE INVENTION

An embodiment of the present invention provides a dual channel video/audio encoder on a single chip to generate compressed video and audio multiplexed into transport streams. One embodiment of the encoder of the present invention supports MPEG standards and AC-3 standards, for example. With a single firmware change, however, the encoder can support any number of other standards as well. Applications for the encoder of the present invention may include personal video recorders, DVD recorders, set top box recorders, PC TV tuners, digital camcorders, video streaming, video conferencing, and game consoles.

Reference is now made to Fig. 5, a block diagram of video encoding video/audio/data multiplexing device 100, constructed and operative in accordance with an embodiment of the present invention.

An embodiment of the present invention overcomes the disadvantage of the prior art by providing a novel approach to video/audio compression and encoding, and, as per this approach, a novel encoding device structure which comprises a plurality of processors with a defined, optimized work division scheme.

Typically, a sequence of compression commands are instructions or a sequence of instructions, such as, removal of temporal redundancy, removal of spatial redundancy, and entropy redundancy of data, and the like. Device 100 operates according to an optimized compression labor division, thus segmenting the compression tasks between the different processors and reducing, in comparison to prior art, the compression time.

According to an embodiment of the present invention, device 100 is a parallel digital processor implemented on a single chip and designed for the purposes of real-time video/audio compression and multiplexing, such as for MPEG encoding and the like. For purposes of clarity herein, multiplexing refers to the creating of synchronized streams of a plurality of unsynchronized

audio and video streams. Device 100 may be incorporated in digital camcorders, recordable digital video disk (DVD), game machines, desktop multimedia, video broadcast equipment, video authoring systems, video streaming and video conferencing equipment, security and surveillance systems, and the like.

According to an embodiment of the present invention, device 100 efficiently performs video compression tasks such as removing temporal redundancy (i.e., motion between frames), spatial redundancy (i.e., motion within frame), and entropy redundancy of data. Device 100 has a plurality of processors, each processor designed to perform a segment of the compression task, hence, achieving optimal performance of each such task.

The number of processors, the architecture of each processor, and the task list per processor, achieves the optimal tradeoff between device implementation cost and efficiency.

In an embodiment of the present invention, device 100 is a dual channel device incorporating both video encoding and audio encoding on a single chip. Device 100 is able to simultaneously encode two channels of video and two channels of audio in one chip.

Device 100 includes a dual channel video input buffer (VIB) 102, a dual channel global controller 104, dual channel motion estimation processors P4 105 and MEF 106, a dual channel digital signal processor (DSP) 108, a dual memory controller 110, a dual channel bitstream processor (BSM) 112, a dual channel audio encoder (AUD) 113, and a multiplexing processor (MUX) 114.

Together, the VIB 102, MEF 106, P4 105, DSP 108, and BSM 112 constitute a video encoder in an embodiment of the present invention.

Device 100 may be connectable to an external video interface, an external audio interface, an external memory unit, and an external host interface. Typically, for example, the video interface supplies a digital video

signal in CCIR 656 format and the audio interface supplies a digital audio signal in I2S/AC97 formats.

The host interface typically connects to an external host (not shown) and acts as a user interface between device 100 and the user. The host interface accepts microcodes, commands, data parameters and the like received from a user or a supervising system. The host interface also transfers information from device 100 to the user. The host interface provides access to the compressed data and may be used to transfer uncompressed digitized video and/or audio and/or user data into device 100.

In an alternative embodiment of the present invention, the host interface is also integrated on the single chip.

In an embodiment of the present invention, device 100 is operable either in a programming mode or an operational mode, and is capable of operating in both modes simultaneously.

In the programming mode, an external host transfers, via the host interface, microcodes, commands and data parameters to dual channel global controller 104. Dual channel global controller 104 transfers the microcodes, commands and data parameters to dual channel video input buffer 102, dual channel motion estimation processors 105 and 106, dual channel digital signal processor 108, dual memory controller 100, dual channel bitstream processor 112, and multiplexing processor 114.

In the operational mode, video input buffer 102 is responsible for acquiring an uncompressed CCIR-656 video signal from an external video source (not shown) and storing it via the memory controller 110. In an alternative embodiment, VIB 102 captures an uncompressed video signal, via the host interface. VIB 102 is responsible for acquiring an uncompressed CCIR-656 video and storing it via the memory controller 110 in an external memory unit in a raster-scan manner.

In an embodiment of the present invention, the memory controller 110 is a SDRAM controller and the external memory unit is an SDRAM memory unit. The SDRAM controller is responsible for communication between the single chip and the external SDRAM memory unit, which is used as a frame buffer and an output buffer for compressed data. The SDRAM controller operations are controlled and scheduled by special instructions issued by the global controller 104.

The memory controller 110 includes two SDRAM controllers that may each interface to an external SDRAM memory unit. Therefore, two streams of data may be written to two external SDRAM memory units at full data resolution. Alternatively, if just one external SDRAM memory unit is provided, one stream of data may be written at full data resolution or two streams of data may be written at half data resolution by arbitrating between the two SDRAM controllers.

Video input buffer 102 performs statistical analysis of the video signal, thereby detecting 3-2 pulled up sequences and developments in the video contents, such as scene change, sudden motion, fade in/fade out and the like. Video input buffer 102 also performs resolution down-scaling, thereby allowing or enabling compression not only of the original resolution frames, but also reduced resolution frames (such as SIF, half D1 etc.). Additionally, video input buffer 102 also pre-processes the video signal, such as spatial filtering, noise reduction, image enhancement and the like. Furthermore, video input buffer 102 decreases the frame rate by decimating (dropping) frames thus allowing flexible rate control.

Video input buffer 102 accumulates the scaled and processed video data and transfers the data in bursts to an external memory unit, via dual memory controller 110. Dual memory controller 110 stores the video data in the external memory unit.

In an embodiment of the present invention, device 100 operates under MPEG video/audio compression standards. Hence, a data block represents a

macroblock, which is a sixteen by sixteen matrix of luminance pixels and two, four or eight by eight matrices of chrominance pixels as defined by MPEG standards. For purposes of clarity herein, reference to a reference frame refers to a frame that has already been encoded, reconstructed and stored in an external memory unit, and which is compared to the current frame during the motion estimation performed by motion estimation processors 105 and 106.

Motion estimation processor 105 (P4) is a level 1 motion estimation engine that is responsible for downscaling current and original reference pictures and for motion vector search. Motion estimation processor 105 finds motion vectors with a 2-pel accuracy by applying a fully exhaustive search in the range of +/- 96 pels horizontally and +/- 64 pels vertically.

Motion estimation processor 106 (MEF) is a level 2 motion estimation engine that is responsible for finding final (half pel) motion vectors. Additionally, the MEF performs horizontal and vertical interpolation of a chrominance signal. The MEF employs a fully exhaustive search in the range of +/-2 pels horizontally and vertically. After the full-pel motion vector is found, the MEF performs half-pel motion search in eight possible positions surrounding the optimal full-pel vector.

The dual memory controller 110 retrieves a current frame macroblock, and certain parts of the reference frames (referred hereto as search area) from the external memory unit and loads them into motion estimation processors 105 and 106. The motion estimation processors compare the current frame macroblock with the respective reference search area in accordance with a sequence of compression commands, thereby producing an estimation of the motion of the current frame macroblock. The estimation is used to remove temporal redundancy from the video signal.

Motion estimation processors 105 and 106 transfer the resulting motion estimation to global controller 104. Motion estimation processors 105

and 106 also transfer the current frame macroblock and the corresponding reference frames macroblocks to digital signal processor 108.

Digital signal processor 108 performs a series of macroblock processing operations intended to remove the spatial redundancy of the video signal, such as discrete cosine transform, macroblock type selection, quantization, rate control and the like. Digital signal processor 108 transfers the compressed data to the bitstream processor 112. Digital signal processor 108 further processes the compressed frame, thus reconstructing the reference frames, and transfers the reconstructed reference frames to the external memory unit via memory controller 110, thereby overwriting some of the existing reference frames.

Bitstream processor 112 encodes the compressed video data into a standard MPEG format, in accordance with a sequence known in the art of encoding commands. Bitstream processor 112 transfers compressed video data streams to multiplexing processor 114.

Audio encoder 113 is a processor responsible for audio encoding. In an embodiment of the present invention, audio encoder 113 supports MPEG-1 Layer II and Dolby AC-3 two channel encoding and may be reprogrammed to support various additional audio compression schemes. The audio encoder 113 is also responsible for acquiring the uncompressed audio signal (I2S and AC97 standards are supported, for example) and buffering the compressed audio. In one embodiment, the audio encoder 113 supports two audio channels.

Multiplexing processor 114 multiplexes the encoded video and the encoded audio and/or user data streams (as received from bitstream processor 112 and audio encoder 113) and generates, according to a sequence of optimized multiplexing commands, MPEG standard format streams such as packetized elementary stream, program stream, transport stream and the like. Multiplexing processor 114 transfers the multiplexed video/audio/data streams to a compressed data stream output and to dual

memory controller 110. Multiplexing processor 114 may output two multiplexed streams, each containing one video stream and one audio stream or one multiplexed stream containing all four input streams (two video streams and two audio streams).

Global controller 104 controls and schedules the video input buffer 102, the motion estimation processors 105 and 106, the digital signal processor 108, the memory controller 110, the bitstream processor 112, and the multiplexing processor 114. Global controller 104 is a central control unit that synchronizes and controls all of the internal chip units and communicates with all of the internal chip units using data-instruction-device buses. The single chip contains two global controller units where each unit corresponds to a different channel.

In an embodiment of the present invention, internal memory of device 100 may be accessed without disturbing the encoding processes. The internal memory may be accessed externally via the host interface. The internal memory includes data such as interrupts, bit rate, total bits encoded, etc. A clock rate to the internal memory is doubled such that internal workings of the chip are performed on even rising clock edges of the doubled clock signal and external accesses are performed on odd rising edges of the doubled clock signal. As a result, time does not have to be scheduled to stop encoding processes to gain access to internal memory.

The various elements of device 100 may be combined or separated according to various embodiments of the present invention.

Also, the various elements may be implemented as various combinations of programmable and non-programmable hardware elements.

In summary, certain embodiments of the present invention afford an approach to perform multi-channel video and audio encoding on a single chip to generate multiple streams of encoded video and audio data for use in various applications such as personal video recorders, DVD recorders, and set top box recorders. In other words, the system of the present invention

enables a single chip that encodes video and audio (and any other system data desired), and generates therefrom one or more streams of encoded data.

While the invention has been described with reference to certain embodiments, it will be understood by those skilled in the art that various changes may be made and equivalents may be substituted without departing from the scope of the invention. In addition, many modifications may be made to adapt a particular situation or material to the teachings of the invention without departing from its scope. Therefore, it is intended that the invention not be limited to the particular embodiment disclosed, but that the invention will include all embodiments falling within the scope of the appended claims.

## CLAIMS

What is claimed is:

1. A single chip digital signal processing apparatus for real time multi-channel video/audio encoding, said apparatus comprising:
  - a multi-channel video encoder for generating at least two channels of encoded video data from uncompressed video data;
  - a multi-channel audio encoder for generating at least two channels of encoded audio data from uncompressed audio data; and
  - a mux processor to generate at least two multiplexed output streams of data from said at least two channels of encoded video data and said at least two channels of encoded audio data.
2. The apparatus of claim 1 wherein an encoding type of said multi-channel video encoder and said multi-channel audio encoder is programmable.
3. The apparatus of claim 1 further comprising internal memory that may be externally accessed without disrupting any encoding processes of said single chip.
4. The apparatus of claim 1 further comprising at least two memory controllers to write at least two channels of encoded data to at least two external memory devices at full data resolution.
5. The apparatus of claim 1 further comprising at least two memory controllers to write at least two channels of encoded data to at least one external memory device at less than full data resolution.

6. The apparatus of claim 1 further comprising at least one memory controller to write at least one channel of encoded data to at least one external memory device at full data resolution.
7. The apparatus of claim 1 wherein said uncompressed video data and said uncompressed audio data are encoded with any one of the encoding types including the various MPEG standards and Dolby AC-3.
8. The apparatus of claim 1 wherein said uncompressed video data comprises CCIR-656 video data.
9. The apparatus of claim 1 wherein said uncompressed audio data comprises one of I2S audio data and AC97 audio data.

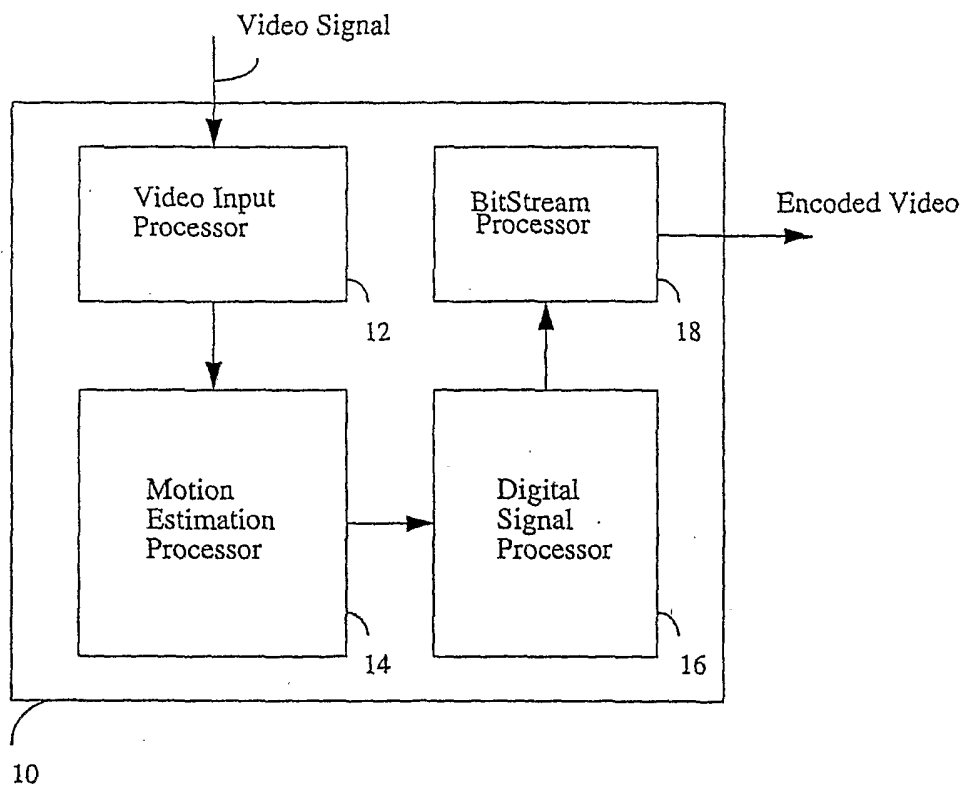


FIG. 1

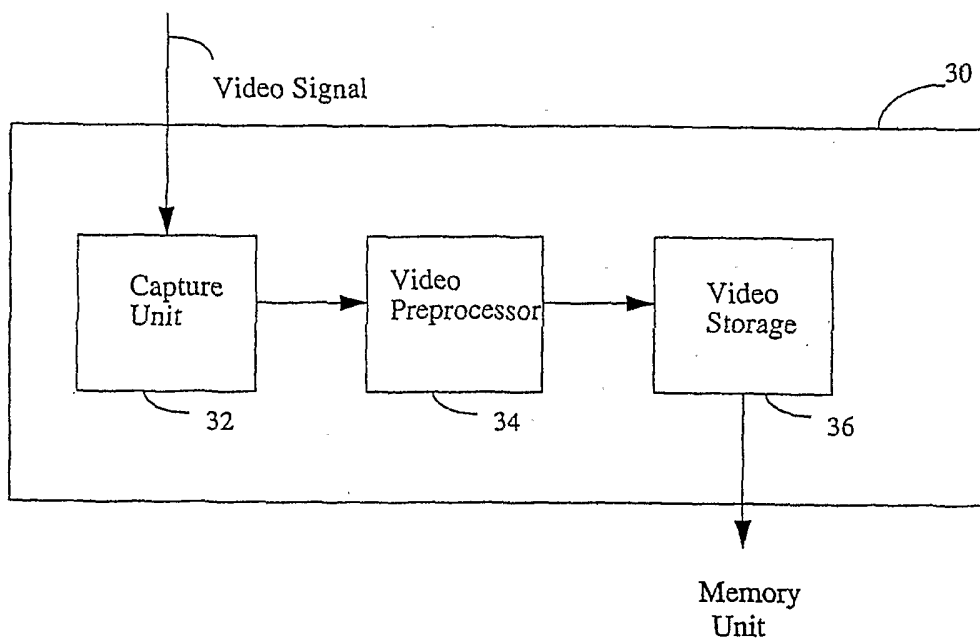
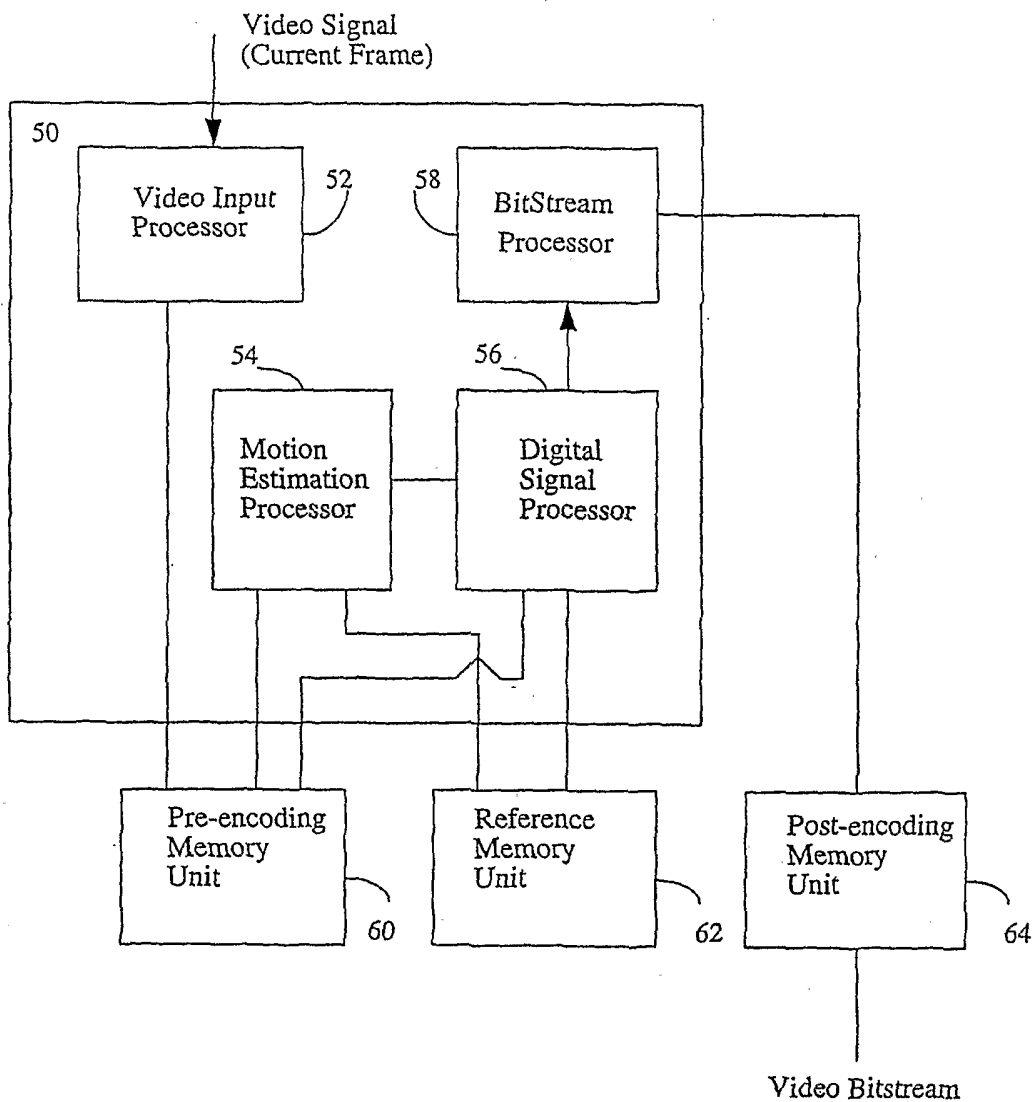


FIG. 2

FIG. 3



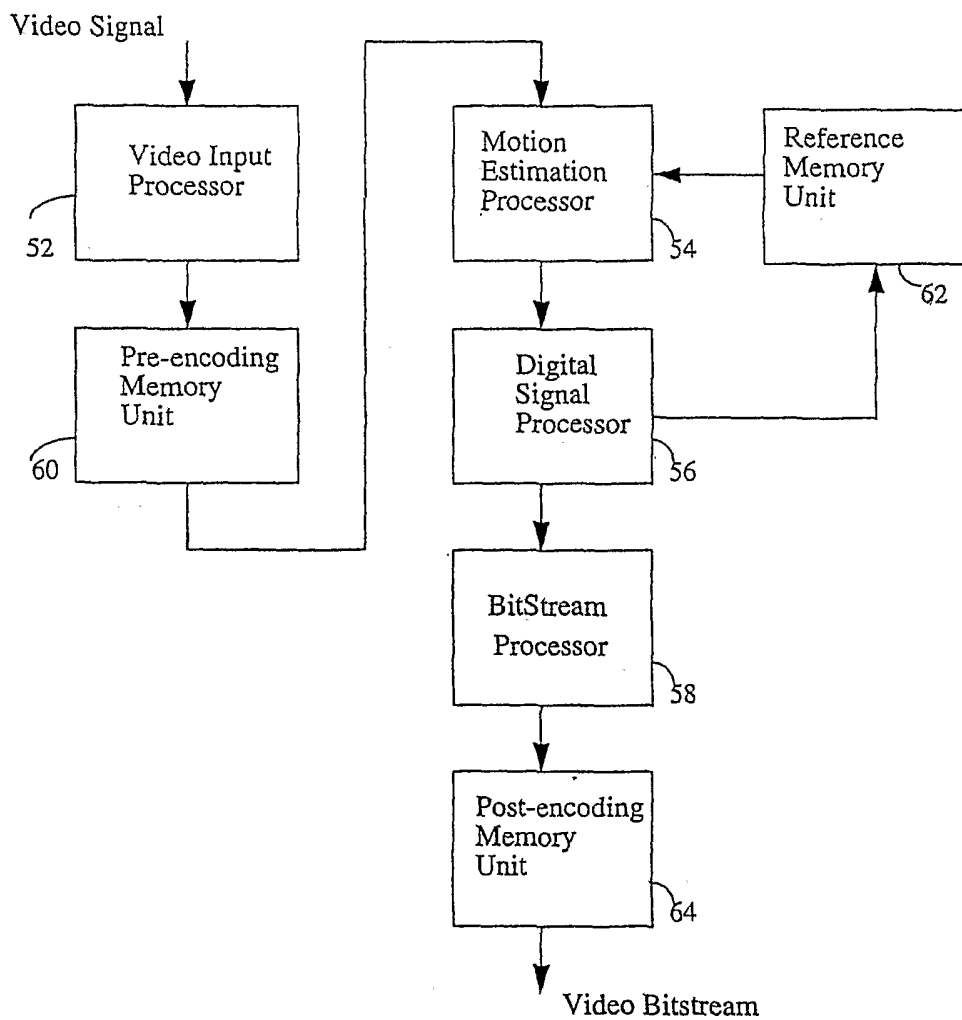


FIG. 4

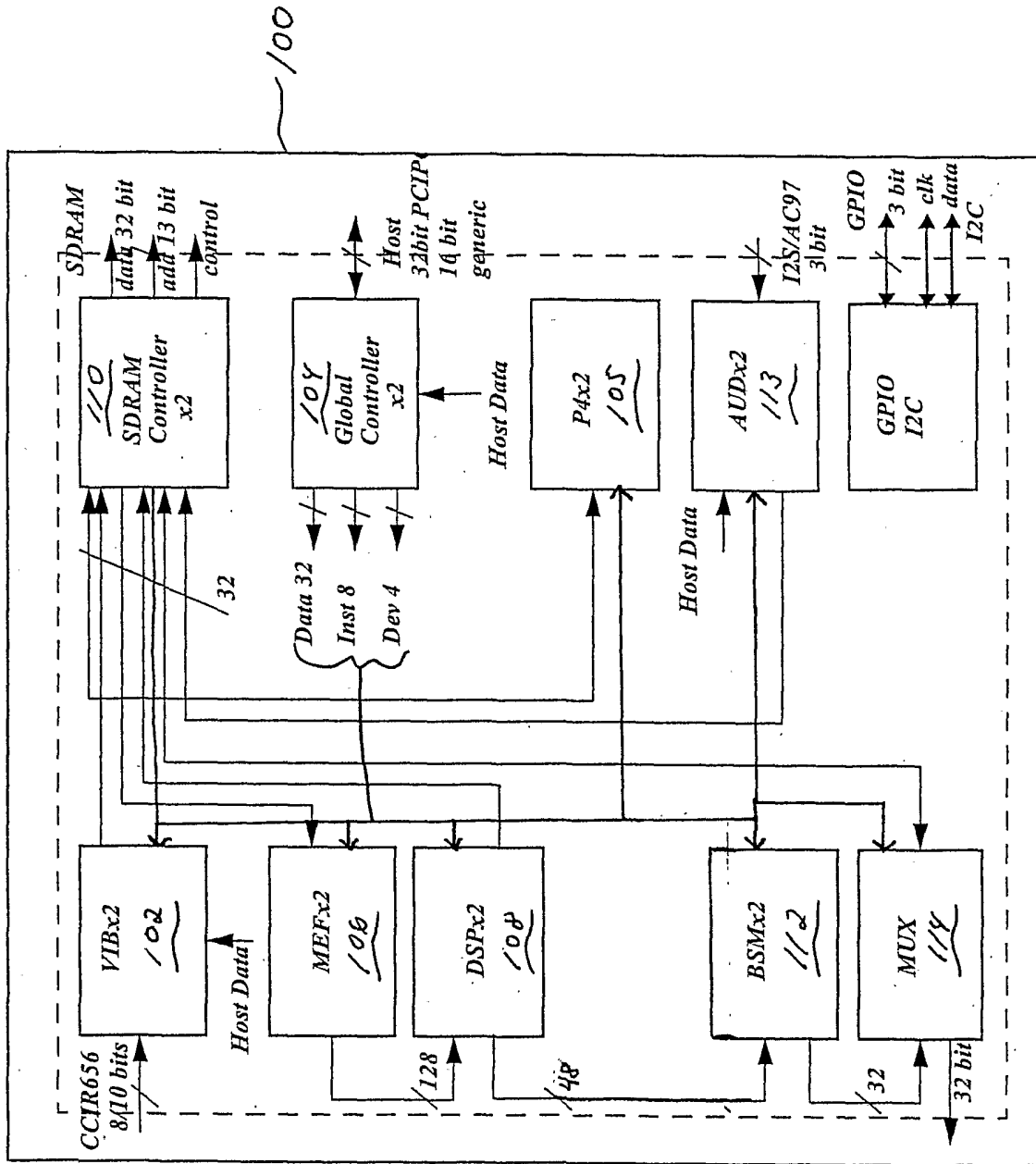


FIG. 5