

MPEG2 MP@ML/HL リアルタイムエンコーダ VLSI チップセットの アーキテクチャとその実装

Architecture and VLSI Implementation of a Complete Chip Set for an MPEG2 MP@ML/HL Real-Time Encoder

(株)グラフィックス・コミュニケーション・ラボラトリーズ 小林 孝之

(株)グラフィックス・コミュニケーション・ラボラトリーズ 藤原 洋

(株)グラフィックス・コミュニケーション・ラボラトリーズ 正会員 西 和彦

Graphics Communication Laboratories Takayuki Kobayashi

Graphics Communication Laboratories Hiroshi Fujiwara

Graphics Communication Laboratories Kazuhiko Nishi

<あらまし>

国際標準 ISO/IEC 13818-2 | ITU-T H.262 (MPEG2 Video) の Main Profile @ Main/High Level に準拠したリアルタイムエンコーダチップセットの開発を行った。

本チップセットでは、膨大な処理量が必要な動きベクトル探索方式において、新たに予測フレーム間距離に適応させて探索範囲を切り替える粗密探索と階層探索を組み合わせた方式を考案し、2次元可変ストリックアレー構造を持つ2種類のチップに実装した。

本チップセットは、上記2種類のチップに前処理および圧縮符号化処理を行う3種類のチップを加えて、SDTV レベルの符号化を行うことができ、TM5 モデルに対し SNR 劣化 0.5dB 以下を実現した。

さらに本チップセットは、符号化画像のレベルにスケーラブルな拡張機能を搭載し、各レベルに合わせて複数セットを組み合わせることで、例えば、プログレッシブ符号化や HDTV 画像の符号化へ簡単に拡張可能である。本拡張機能は、チップセット中に画面分割機能とビットストリーム結合機能を持たせることで実現した。さらに、画質制御の連続性を保つために符号化制御パラメータの伝送機能を持つコミュニケーションバスを持たせることで、各チップセット間の量子化制御の連続性を保つことが可能となり、画面境界での不連続部を解消した。

<Summary> We have developed a real-time encoder chipset which completely complies to the standard ISO/IEC 13818-2 | ITU-T H.262 (also known as "MPEG2 Video"), for the Main Profile at Main/High Level. In the motion vector search which is required the enormous amount of computational, we developed new search method, based on a two-step hierarchical coarse/fine search adapting to the frame distance between a coding frame and a reference frame. We implemented new search method to two systolic-array based chips: a coarse motion estimation (CME) for the wide range subsampled search, and a fine motion estimation (FME) for the smaller range half-pel resolution refinement.

One set of these five chips is sufficient for real-time encoding of source images equivalent to the ITU-R BT.601 standard, and by connecting several sets together, progressive encoding and HDTV level real-time encoding are possible. In the preprocessing stage, the input image is divided into several segments, and several chipsets work on these segments mostly independently. But a communication bus connects the chips and internal RISC controller, so that coding control information can be shared, allowing the code generation to be managed as if by a single process.

キーワード：MPEG2，動きベクトル探索，HDTV，RISC，シストリック・アレー

1. まえがき

動画像データに関する汎用符号化標準として、ISO/IEC MPEG committee と ITU-T SG15 は、それぞれ 1994 年 11 月に ISO/IEC 13818-2^[1]（いわゆる MPEG2 ビデオ）の国際標準化（IS 化）、および 1995 年 5 月に ITU-T H.261^[1]の勧告化を行った。

この MPEG2 ビデオは、放送、通信および蓄積系の伝達メディア全般における汎用符号化標準として、現在実用化が急ピッチで進められている。これらのメディアにおいて、放送系や通信系の両メディアでは、特にリアルタイムエンコーダが求められている。さらにデジタル放送では、プログレッシブ符号化や HDTV (High Definition Television) 符号化などの高品質な符号化が求められている。

リアルタイムで符号化可能な MPEG2 ビデオエンコーダシステムの実現には、膨大な演算量が必要になり、小型で安価なシステムの実現の障壁となっていた。また HDTV 符号化では、ITU-R BT.601 レベル、いわゆる SDTV (Standard Definition Television) 符号化に比較し、約 6 倍の処理量が必要となり、システムの実現にはさらに大きな障壁となっていた。このようにリアルタイムエンコーダの実現の難しさが、多様なアプリケーションの発展の妨げとなっていた。

我々は、現行の半導体プロセス（ $0.6\mu\text{m}$ のスタンダードセルを想定）を使用して、MP@ML/HL 符号化が可能なリアルタイムエンコーディングチップを実現するために、まずその処理量、メモリバンド幅、およびメモリ容量を算出した。また各種実現手法による画質への影響を、ソフトウェアシミュレータにより、詳細に検討した。

その結果、動きベクトル探索方式において、符号化画像と参照画像のフレーム間距離に応じて探索範囲と探索密度を切り替えることが可能な粗密探索方式と階層探索方式を組み合わせた新しい方式を考案し、CME と FME と呼ぶ 2 種類の 2 次元シストリックアレーのチップに実装した。この 2 つのチップにより、探索範囲が水平方向 $-24*fd - 4 \sim +23*fd + 4$ （ここで、 fd は符号化画像と参照画像との時間的な距離を意味する）、垂直方向 $-18 \sim +17$ を実現した。またこのときの符号化処理量およびメモリバンド幅は、全探索方式に比べて約 $1/8.5$ 、 $1/10$ と極めて低い値を実現した。

これらの新しい探索方式と符号化処理におけるメモリバンド幅の均一化等を行い、5 種類のチ

チップセットと3つのSDRAMによりSDTVレベルの符号化機能を実現した。またこのときの画質劣化は、TM5モデルに対しSNR劣化0.5dB以下を実現した。

さらにSDTVレベルからHDTVレベルまでのリアルタイム符号化への拡張機能についても検討した。その結果、画面分割機能とビットストリーム結合機能をチップセット内に持たせることで、付加回路無しにチップセットの連結のみで1080IのHDTVレベルの符号化を実現した。また拡張機能の実装に際し、各チップセット間で符号化制御パラメータの伝送機能を持たせ、スライス毎に符号化制御のタイミングをずらすことで、各チップセット間の量子化制御の連続性を保つことが可能となり、画面分割符号化で問題となる画面境界での不連続部を解消した。

本論文では、これらのチップセットのアーキテクチャと特徴について述べる。

2. MPEG2 符号化処理の検討

2.1 MPEG2 符号化機能

図1に、MPEG2符号化アルゴリズムの構成を示す。

MPEG2符号化アルゴリズムは、大きく分けて前処理と圧縮符号化処理の2つの処理に分けられる。

まず前処理では、入力されるビデオデータからマクロブロック構造の符号化原画像を生成する。前処理で行われる処理は、入力される4:2:2画像から4:2:0画像やSIF画像へのフォーマット変換処理、インタレース画像を符号化単位であるマクロブロックに変換するスキャン変換処理、双方向予測のためのフレームリオーダリング処理が行われる。また、この前処理で画面のノイズなどを除去するためのプリフィルタ処理が、行われることもある。

次に、前処理で得られた符号化原画像から符号化ビットストリームを生成する圧縮符号化処理が行われる。圧縮符号化処理では、動きベクトル探索処理 (Motion Estimation)、動き補償処理 (Motion Compensation)、予測符号化処理 (Predictive Coding)、DCT/逆DCT処理 (IDCT)、量子化 (Q)/逆量子化処理 (IQ)、可変長コード変換処理 (VLC)、ビットストリーム伝送処理 (Buffer)、符号化モード制御 (Coding Mode Control)、量子化制御 (Quantisation Control) が行われる。

このようにMPEG2符号化アルゴリズムでは、上記のような多くの処理とそれに伴う多くの適応処理が必要とされる。

2.2 符号化処理量

MPEG2メインプロファイルビデオでは、MPEG1とは異なり、その多様性のため、様々な技術が加えられている。まずテレビ放送で使われるインタレース画像に適応した多様な予測モードが加えられた。また放送局の1次分配品質レベルの画質を実現するための、様々な高画質化技術が採用された。さらに多様なメディア形態に適応するための、符号化技術が導入された。このような様々な技術は、主にエンコーダへの処理の負担を必要としている。我々は、まず設計に先立ち

符号化処理における処理量を求めた。

表 1 に MPEG2 エンコーダのそれぞれの符号化処理における処理量を示す。表からも判るように、符号化処理における処理量の大部分を動きベクトル探索処理が占めている。また DCT/IDCT 処理、可変長符号化処理も比較的処理量が大きく、これらの処理に対し現行のプロセスを使用した場合、動作周波数の制約からプロセッサ構成では多数のプロセッサが必要となってしまう。このため、我々は、各処理に対する処理量と複雑さを考慮し、各処理に応じたアーキテクチャを採用することにした。

2.3 メモリバンド幅とメモリ容量

MPEG2 ビデオでは、リオーダメモリや予測メモリ、および伝送バッファ等の多くのメモリが必要とされる。そして、これらのメモリに対し、特に符号化における様々な適応処理のため、多くのメモリアクセス量が必要とされる。そのため、小型で安価なエンコーダシステムを構成するには、メモリの最適な構成と、メモリとのデータ転送帯域の有効利用が重要な課題となる。

表 2 に各処理におけるメモリ量を、表 3 に各処理におけるメモリバンド幅を示す。

入力されるビデオデータから符号化原画像を生成する前処理系において必要とされるメモリは、B ピクチャ符号化のためのリオーダバッファ、インタレースされたフィールド画像をフレーム画像に変換するためのフォーマット変換バッファである。また符号化原画像を使用し動きベクトル探索を行う場合、前方および後方予測画像のための参照画像バッファが必要とされる。

一方符号化マクロブロックから伝送ビットストリームを生成する符号化処理系においては、前方および後方予測のための予測バッファ、生成されたビットストリームを伝送路に合わせて伝送するためのデコーダの VBV (Video Buffering Verifier) バッファと逆の動作を行うバッファが必要とされる。

その結果、ITU-R BT.601 レベルの符号化を行う場合、前処理系においては、約 2.5MBytes（動きベクトル探索を入力ビデオから行う場合、約 3.3MBytes）が、符号化処理系においては、約 1.5MBytes が必要となる。

また前処理および符号化処理におけるメモリバンド幅は、動きベクトル探索を除いて算出すると、フレーム - フィールド変換、4:2:2/4:2:0 変換、プリフィルタ処理、フレームリオーダ処理、符号化ブロック読み出し、再生画像読み込み、再生画像書き込み、符号化データ書き込み、符号化データ読み出しが必要となり、I ピクチャで約 100MBytes/sec、P ピクチャで約 119MBytes/sec、B ピクチャで約 123MBytes/sec となる。

動きベクトル探索処理は、フレーム間距離に応じた探索範囲を採り、整数画素精度で全探索を、全探索で求めた整数画素の周辺 8 画素を 1/2 画素精度で探索を行う場合、約 780 MBytes/sec となる。このように動きベクトル探索を上記のような方法で実現する場合、多くのメモリバンド幅が必要となるため、何らかの削減手段が必要となる。

3. チップセットアーキテクチャ

3.1 機能分割

我々は、現行の半導体プロセスを想定して、前節の解析をふまえ、MPEG2 符号化アルゴリズムを実現するための適切な回路構成について検討した。

まず前処理では、フィルタリングなどの画素単位の定型的な処理が行われる。SDTV レベルでは、13.5MHz 単位で処理が行われ、HDTV レベルでは、74.25MHz 単位で処理が行われる。特に HDTV レベルを想定した場合、高速な処理が必要とされる。このため定型処理に適したパイプライン型の専用ハードウェアで構成することにした。

次に、圧縮符号化処理は、定型的な処理と、プログラマブルな適応処理とに分けることができる。

まず動きベクトル探索処理、DCT／逆 DCT 変換処理、量子化／逆量子化処理、可変長コード変換処理などは、定型的な処理である。従って、専用ハードウェアを使用したパイプライン処理により、ゲート規模を削減することが可能である。

一方、量子化制御処理では、いろいろな画像に対し高画質を維持させるため、各種画像タイプや伝送レートなど符号化条件に適応させて、処理を柔軟に変更することが必要となる。また量子化制御処理は、統計量を算出するために、乗算や除算など比較的複雑な処理が必要とされる。しかしながら、この量子化制御を行うための基本制御時間は、マクロブロック期間で約 $24\mu\text{s}$ と比較的時間余裕があるためプロセッサで実現することが可能である。

我々は、このように MPEG2 符号化アルゴリズムを機能的に分割し、各処理量とハードウェア量、メモリバンド幅、メモリ容量から、機能毎に適応したアーキテクチャを持つ5種類のチップに分割した^[2]。量子化制御には、RISC 型の専用プロセッサをチップセット中に内蔵し、それ以外の部分は、専用ハードウェアにより構成した。また量子化制御を行う制御プログラムについては、ダウンロード可能とし、各種画像に対し高画質を維持できる柔軟な構成とした。

3.2 動きベクトル探索処理

動きベクトル探索処理は、MPEG2 符号化処理の内、大部分の処理量を占め、メモリバンド幅も大きく、また探索方法が圧縮効率に大きく影響する。さらに探索範囲が足りない場合、予測効率が大きく低下する。このため、まず我々は、符号化に必要とされる動きベクトル探索範囲と、効率的な探索アルゴリズムについて、検討を行った^[3]。

3.2.1 動きベクトル探索範囲

通常、TV プログラムにおける画像は、パンニングが多く使用され、チルティングはあまり使用されない。また水平方向の動きベクトル探索の場合、パイプライン処理によりメモリバンド幅を変えずに、探索範囲を広げることが可能である。一方垂直方向の場合、インタレース画像に対

する動きベクトルの誤りは、符号化効率に大きく影響する。さらに探索範囲を広げると、それに比例してメモリバンド幅が大きくなってしまう。

我々は、動きベクトル探索について、これらの特性を考慮し、水平方向を広い範囲で探索し、垂直方向の探索については、インタレース構造に注意して探索を実行することにした。またこの考察を元に、種々の画像を評価し、水平方向探索ではフレーム間距離当たり-24～+23画素を、垂直方向探索では固定の-18～+17画素の探索を実現することにした。

3.2.2 処理量とメモリバンド幅の削減

MPEG2 ビデオでは、プログレッシブ符号化に加えて、インタレース画像の予測効率を高めるためにフレーム／フィールド適応動き補償が行われる。またM=1の符号化では、デュアルプライム予測も行われる。このため、これらの動きベクトルを探索する処理量およびメモリバンド幅は、さらに膨大になる。このため動きベクトル探索処理をLSI化するに当たって、動きベクトル探索処理の削減とメモリアクセス量の削減が重要である。

しかしながら動きベクトルの誤りは、符号化の圧縮効率や主観評価上の画質に大きく影響する。このため動きベクトル探索処理の削減とメモリアクセス量の削減を行いながらも、画質の劣化を最小に留めることが、さらに重要である。

一方動きベクトル探索処理は、繰り返し演算が行われるため、アレーによる並列処理とパイプライン処理が効果的である。さらに近傍のマクロブロックに対する参照画像がオーバーラップするため、内部バッファを有効に利用することにより、メモリバンド幅を削減することが可能である。

我々は、この動きベクトル探索を、階層探索と粗密探索を組み合わせ使用した2階層の探索とした。以下に、我々の実現したアルゴリズムについて詳述する。

(1) 階層探索

動きベクトル探索で、階層探索を使用すると、その縮小率に応じて、その処理量とメモリアクセス量を低減させることができる。一方、同時に2つのフィールドブロックの予測誤差を求めるようにすれば、フレームブロックの予測誤差はそれらの和として求めることができる。

インタレース画像において、フィールド／フレーム予測モード判定の誤判定は、符号化効率に大きく影響する。このため我々は、水平方向のみ1／2にサブサンプリングした画像を使用することにした。しかも、2つのフィールドブロックの予測誤差を同時に求めることにし、これらの予測誤差からフレームブロックの予測誤差を算出することにした。またサブサンプル時のフィルタの効果について、種々の標準画像でシュミレーションを行い評価を行った。その結果水平方向を1／2に圧縮するフィルタの効果は、ITU-R BT.601 放送レベルの画像においては、有意差がほとんどないことが判った。このためハードウェア量とメモリバンド幅を考慮し、サブサンプルにおいてフィルタを使用しないことにした。

これらにより、処理量及びメモリバンド幅をそれぞれ1/4に削減することが可能となった。加えて、2つのフィールドブロックの予測誤差を求める演算器を時分割で使用するにより、更にハードウェア量の削減を図った。

一方フィールド予測では、同一フィールドベクトル（トップフィールドからトップフィールド、ボトムフィールドからボトムフィールド）と異なるフィールドベクトル（トップフィールドからボトムフィールド、ボトムフィールドからトップフィールド）を求めるためのサーチエリア画像は、垂直位置が1ライン異なるだけである。我々は、シストリックアレー上でサイドレジスタを垂直方向のみばかりでなく水平方向にも設けるアーキテクチャを新たに開発した。これにより、一度探索処理を終えたサーチエリア画像を再度初期化位置に戻すことが可能となった。この初期化位置に戻したサーチエリア画像に対し、符号化ブロックをフィールド入れ替えし、再度探索処理を行うことで、同一フィールドベクトルと異なるフィールドベクトルを時分割で探索させることが可能となった。これにより、従来同一フィールドベクトルの探索と異なるフィールドの探索を2回に分けていたものを1回のみで済ますことが可能となった。このことにより、後述するように1マクロブロック列のレジスタ（水平サイドレジスタ）の増加と逆方向バスの選択回路および位相調整用のセレクタの追加のみで、メモリバンド幅をさらに1/2に削減した。

(2) フレーム間隔適応粗密探索

動きベクトルの性質として、符号化画像と参照画像とのフレーム間距離に応じて、探索範囲を切り替えた方が効率よい。この方法としては、テレスコピックサーチ法があるが、各段階でパイプライン処理ができなくなることと、各段階でフルサーチを行うため、処理量及びメモリバンド幅ともかなり大きくなる欠点がある。

我々は、各種画像のシミュレーションから、表4に示すように、フレーム間距離に応じて探索点の密度を変え、探索範囲を切り替える方法を開発した。さらに、フレーム間距離が3のPピクチャの場合、動きベクトルの大きさが小さいときは密探索し、大きいときは粗探索するようにした。この方法により、全て密探索のみを行う方法と比べて、処理量を削減しながら SNR 劣化 0.1dB 以下と有意差なく探索が行えるようになった。

我々は、このアルゴリズムを実現するため、可変シストリックアレーを2分割させて探索密度を変更できるアーキテクチャを、新たに開発した。我々は、この階層探索と粗密探索を組み合わせたハイブリッド探索方式を、CME と呼ばれる新しい可変シストリックアレー構造の LSI に実装した。

(3) 近傍ブロックによる重み付け

上記のような粗い探索を行う場合や、サブサンプル画像により探索を行う場合、検出動き量の空間解像度が劣化するために、雑音などの影響を受けやすく検出誤りが発生しやすくなる。我々は、これらの検出誤りをメモリバンド幅や処理量を増やすことなく低減させるため、近傍ベクトルによる予測誤差の重み付け比較を行うことにした。重み付け関数については、指数関数、1次

関数、ステップ関数等により、その効果を検証した。その結果、重み付けによる SNR 改善効果は、関数による有意差がほとんどないため、下記のような単純なステップ関数による重み付け関数を採用した。

$$\text{dis}' = a * \text{dis}$$

$$a = 1 \quad x + y \leq R$$

$$a = 1 + \text{Const.} \quad x + y > R$$

ここで、dis は予測誤差、dis'は重み付けされた予測誤差、a は重み付け関数、x, y は中心からの距離、R は重み付けの設定境界、Const.は重み付け定数を示す。

これらの手法により、CME チップでは、広い探索範囲を実現しながら、局所最小による誤ったベクトルを選択しにくくし、符号化効率を改善した。

(4) フルサーチ

粗い探索で求めた2つのフィールドベクトルとフレームベクトルは、次段でその周辺を1/2画素精度で探索する。従来、本処理は粗い探索で求めた周辺で、まず整数画素の探索を行い、その後、近傍8画素の探索が行われる。このため最低でも外部メモリに対し2回参照画像を読み出すか、メモリバンド幅に余裕が無い場合、内部にバッファを設ける必要があった。加えて1/2画素探索では、1/2画素フィルタにより8種類の画像を生成した後に、予測誤差を算出していた。このため、1/2画素画像を1つずつ生成し、予測誤差算出を行うか、8種類の画像生成を同時に行い、予測誤差を算出する必要があった。前者の場合、回路規模は小さくなるが、バッファアクセスを8回行う必要がある。後者の場合、バッファアクセスは1回のみであるが、制御が複雑となり回路規模が大きくなる。また1/2画素フィルタには、ラインバッファが必要で比較的回路規模が大きくなる。このように従来方法には多くの欠点を持っていた。

1/2画素は、下記に示すように、整数画素から線形フィルタによって求まる。従って、1/2画素の予測誤差は、隣接する整数画素の予測誤差から求めることが可能である。

$$\text{DI} = x_{00} - y_{00}$$

$$\begin{aligned} \text{DHh} &= x_{00} - (y_{00} + y_{10})/2 \\ &= ((x_{00} - y_{00}) + (x_{00} - y_{10}))/2 \end{aligned}$$

$$\begin{aligned} \text{DHv} &= x_{00} - (y_{00} + y_{01})/2 \\ &= ((x_{00} - y_{00}) + (x_{00} - y_{01}))/2 \end{aligned}$$

$$\begin{aligned} \text{DHd} &= x_{00} - (y_{00} + y_{01} + y_{10} + y_{11})/4 \\ &= (((x_{00} - y_{00}) + (x_{00} - y_{01})) \\ &\quad + ((x_{00} - y_{10}) + (x_{00} - y_{11}))) / 4 \end{aligned}$$

ここで、 x は符号化画像の画素、 y は参照画像の画素を示す。

我々は、この性質を利用して、1回の画像読み出しだけで整数画素と1/2画素の予測誤差を、同時に算出できるようにしたフルサーチ2次元シストリックアレーを新たに開発し、FME と呼ばれるチップに実装した。本 LSI は、従来の2次元シストリックアレーの探索器に、上記の1/2画素演算器と1/2画素用の予測誤差累算器のみの回路追加でフルサーチ探索機能を実現している。

(5) デュアルプライムベクトル探索

デュアルプライム予測や内挿予測の再生画像は、2つのフィールド（或いはフレーム）を線形フィルタリングする事により求められる。従って、デュアルプライム予測や内挿予測においても、1/2画素探索と同じように、フィルタ処理画像を生成することなく探索が可能である。つまり密探索により求めた同一フィールド画像を使用し、符号化画像からこの参照画像の画素値の1/2を引いた画像を新しいテンプレートとする。そして、もう一方の異なるフィールドの参照画像の画素値を1/2した画像をサーチエリアとして、動きベクトル探索を行うことでデュアルプライム探索を実現できる。

$$\begin{aligned} DD &= x - (y_s + y_o)/2 \\ &= (x - y_s/2) - y_o/2 \end{aligned}$$

ここで、 x は符号化画像を、 y_s は同一フィールド画像を、 y_o は異なるフィールド画像をいう。

我々は、FME と呼ばれる LSI に1/2画素と整数画素を同時に求めることが可能なシストリックアレーを新たに開発し、このシストリックアレーを使用し、通常の密探索とデュアルプライムのような内挿画像の探索を切り替えて使用できるようにした。

(6) フィールド／フレーム判定

粗い探索を行った後の密探索では、探索範囲が小さくてもかなりのメモリバンド幅が必要となる。Bピクチャを含む MPEG 符号化では、Bピクチャは、元々予測誤差エントロピが少なく、また次の予測のために使用されないのので、IピクチャやPピクチャより量子化精度を落とすことで圧縮効率を高めている。従って、動き補償の精度を落としても主観評価上影響は少ない。そこで、各種画像のシュミレーションから、Bピクチャに於いて、フィールド／フレーム判定は粗い探索で求めたベクトルの予測誤差を使用することで、画質劣化が0.2dB以下と最小に留めながら、密探索の処理量およびメモリバンド幅ともに1/2に削減した。かつ密探索においてPピクチャと同じ処理量とバンド幅にすることを可能とし、処理の平滑化により符号化制御を容易にした。これらの判定機構は、FME を制御する COD チップ内で実現した。

3.3 メモリアクセスの均一化

動きベクトル探索では、前節で論じたように、参照画像の読み込み、テンプレート画像の読み

込みが必要で、特に参照画像の読み込みは、多くのメモリバンド幅を必要とする。我々はこの動きベクトル探索において、前節で詳述したような新しい動きベクトル探索アルゴリズムとシストリックアレーアーキテクチャを開発した。この2つのチップのメモリバンド幅は、CME で1フレームの片方向探索を行う場合約 16MByte/sec, FME で1回のフレーム探索または2回のフィールド探索を行う場合、それぞれ約 24MByte/sec, デュアルプライム予測の DMV (Differential Motion Vector) 探索の場合約 5MByte/sec となる。

次に、MPEG2 符号化アルゴリズムとこの動きベクトルチップを使用した場合の符号化特性とそのメモリバンド幅について検討した。

MPEG2 ビデオの符号化では、一般にバッファの増加を避けるために入力するビデオデータに合わせて、ピクチャ単位やマクロブロック単位などの一定時間間隔で符号化を行う。この場合、MPEG2 では、予測方式がピクチャタイプに応じて切り替えられるために、ピクチャタイプによるメモリバンド幅の偏りが発生する。従って、エンコーダシステムにおいて、効率的なメモリ構成を実現するためには、画質の劣化を最小に抑えながらも、このメモリバンド幅の偏りを低減することが重要となる。

MPEG では、B ピクチャは時間的に前のフレームと後のフレームの双方から予測を行うのに対し、P ピクチャは時間的に前のフレームからのみ予測を行う。これは、動きベクトル探索処理において、B ピクチャの場合、P ピクチャに対し約2倍のメモリバンド幅が必要となることを示している。また一方 P ピクチャは、符号化画像と参照画像とのフレーム間距離が B ピクチャより大きいので、P ピクチャは B ピクチャより広い範囲を探索した方がよい。従って、粗い動きベクトル探索を行う CME チップでは、P ピクチャの探索を左半分の探索と右半分の探索の2ステップに分けて探索し、B ピクチャでは前方探索と後方探索の2ステップで行うことにした。これにより、CME チップにおける動きベクトル探索のメモリバンド幅は、B ピクチャと P ピクチャほぼ均一になった。

また、FME では CME で求めたフレームベクトルと2本のフィールドベクトルの周辺を1/2画素精度で探索を行うが、B ピクチャの場合、前方探索と後方探索が有るため、P ピクチャに対し2回の探索に相当する2倍のメモリバンド幅が必要になる。そこで、前節で詳述したように、B ピクチャでは CME で求めた予測誤差を使用し、フレーム/フィールド予測判定を密探索前に行うようにした。これにより、前述のように SNR 劣化を 0.2dB 以内と最小に留めながら、FME チップにおける動きベクトル探索のメモリバンド幅も、B ピクチャと P ピクチャほぼ均一にすることが可能となった。

さらに MPEG ビデオでは、ピクチャタイプに応じた適応的な予測符号化が行われるため、B ピクチャ、P ピクチャ、I ピクチャの順に発生符号量が大きくなる。一方これら適応的な予測処理に伴うメモリバンド幅は、I ピクチャ、P ピクチャ、B ピクチャの順に大きくなる。従って、各ピクチャ毎に発生符号量の制御を行うことで、さらにメモリバンド幅の均一化を行うことが可能

となる。

そこで我々は、量子化されスキャン変換された DCT 係数データについて、可変長符号化を行う前にブロック毎にその符号長を調べ、ピクチャタイプに応じて、その発生符号量がある一定値を超えた場合に、以降の高周波係数を"0"に置き換える発生符号量制御を行った。本制御方法は、高ビットレート時に、高周波係数のカットが発生しやすくなるため、画質に影響が出てくることが予想される。そこで高ビットレート時の画像について画質を評価したが、動画を止めて観察することで解像度の劣化を確認できるが、通常再生では検知することができなかった。

このようにメモリバンド幅の均一化処理を、注意深く行うことで、SNR 劣化が 0.2dB 以内と最小に抑えながらも、各ピクチャタイプのメモリバンド幅を均一にすることが可能となった。

これらのアーキテクチャを実装したリアルタイムエンコーダチップセットは、前処理系においてのメモリのバンド幅が最大約 244MByte/sec、圧縮符号化処理系においてのメモリバンド幅が最大約 130MByte/sec を実現した。その結果、前処理系において 2 個の 16Mbits SDRAM と、圧縮符号化処理系において 1 個の 16Mbits SDRAM を、72MHz のデータ転送クロックで動作させることで、画質の劣化を最小に抑えながらも、B ピクチャを含む MP@ML レベルのリアルタイムエンコードを、前処理系を含めて実現した。

3.4 HDTV 符号化アーキテクチャ

HDTV レベルの符号化処理は、SDTV レベルに比べて、さらに 6 倍の演算量が必要であるため、多数のチップが必要となる。この HDTV レベルの符号化機能を、機能分割による符号化処理で実現しようとする、他種類のチップが必要となってしまう。また、各処理毎に複数チップを使用する構成では、いろいろなレベルの符号化機能を実現する場合、接続が複雑になってしまう。このため我々は、SDTV レベルを基本構成単位とし、各レベルに対応して基本構成単位毎に画面分割した部分画面を符号化するスケーラブルアーキテクチャを採用した。

しかし画面分割により符号化処理を行う場合、分割画面にまたがる動き補償が発生する。この隣接する部分画像の動き補償を行った場合、メモリバンド幅が局所的に変動することになる。これを避けるため、隣接する部分画像の動き補償を行わずに符号化すると、分割画面の境界部の境界線が見えてしまうことがある。例えば水平方向にパンニングしている画像を符号化する場合、両画面の境界部分でイントラマクロブロックとインターマクロブロックが切り替わることになる。この場合、符号化モードで量子化特性が異なるため、境界部分が見知されることがある。さらに量子化制御を部分画像毎に独立させて実行すると、量子化ステップが画面境界で不連続になり、さらにはっきり境界線が見えてしまうことになる。

このため我々は、図 2、3 に示すように、各画面毎の符号化処理を、1 スライス分時間をずらして符号化を行い、1 スライス分の符号化情報を右隣の符号化処理部に送ることで、符号化処理の連続性と符号化モードの切り替わりに合わせた補償を行うようにした。これにより部分画面に

またがった動き補償を行わなくても、境界部分を見知させずに符号化できるようになった。

4. VLSI 実装

4.1 PRE

PRE チップは、図 4 に示すような専用ハードウェアにより構成した。PRE チップでは、スキャン変換処理、プリフィルタ処理、4:2:2 から 4:2:0 や SIF へのフォーマット変換処理、フレームのリオーダリング処理と CME チップの制御およびメモリ制御が行われる。また PRE チップ内に、符号化原画像に対しアクティビティなどの各種統計量を算出する機能と、雑音除去フィルタを実装し、適応的に制御可能とした。PRE で扱うメモリは、1M*16bits の SDRAM を 2 個使用し、スキャンおよび 4:2:0 変換バッファ、リオーダバッファおよび CME 用参照画像バッファとして使用する。PRE では、CME の制御も行い、その粗探索ベクトルを COD チップに転送する。

4.2 COD

COD チップは、動き補償処理、予測符号化処理、ビットストリーム伝送処理、FME チップの制御およびメモリ制御を行い、図 5 で示すような専用ハードウェアで実現した。また予測モード等の判定方法については、専用ハードウェアを柔軟に変更できるようにした。

COD チップでは、SDRAM 1 個を使用し、予測符号化のための 2 面の予測バッファと可変長符号化されたビットストリームのための伝送バッファとして使用する。なお COD チップでは、次段の MPEG2 システムビットストリーム生成のために PES ヘッダ生成処理機能も実装した。

4.3 QVL

QVL チップは、図 6 に示すように、プログラムにより量子化制御を実行する RISC と、DCT／逆 DCT 処理、量子化／逆量子化処理、スキャン変換処理、可変長コード変換処理を行う専用ハードウェアで構成した。RISC では、ホストプロセッサとの交信とチップセットの全体処理および量子化制御を行い、適応的に処理を行う。

4.4 CME[3]

CME は、粗探索の動きベクトル探索を行い、図 7 に示すように 2 組のプロセッサエレメント(PE)アレーを持つ 2 次元可変シストリックアレー構造により構成される。各 PE アレーは、図 8 に示すようなシストリックアレー構造をして、図 9 に示すプロセッサエレメント 192 個と、図 10 に示す中間レジスタ 176 個により構成される。

プロセッサエレメントは、セレクトタにより F_i を選択することで、中間レジスタをアレー中から切り離すことができる。また FA_i を選択することで、中間レジスタをアレー中に配置することができる。アレー中から中間レジスタを切り離すことで、全ての画素位置を探索し、中間レジスタをアレー中に配置することで、2 画素毎に粗く探索することが可能となる。また 2 組の PE ア

レーは、上記のアレー構造の変更を独立に変更することが可能である。探索距離が1のBピクチャの場合、両アレー共に F_i を選択することで、水平方向-24画素から+22画素までの探索ができる。また、探索距離が2のBピクチャの場合、両アレー共に FA_i を選択することで、水平方向-48画素から+44画素までの探索ができる。さらに、探索距離が3のPピクチャの場合、左側探索で PER を FA_i に、 PEL を F_i にし、右側探索で PEL を FA_i に、 PER を F_i にすることで、水平方向-72画素から+70画素までの探索ができる。このとき-24画素から+23画素までは、全ての位置が探索され、その外側では2画素単位で探索される。このように本シストリックアレー構造により、参照画像とテンプレートとのフレーム間距離に従い、3段階に探索範囲を切り替えることが可能となった。

また、本チップには、従来の垂直サイドレジスタ (SR_v) に加えて、水平サイドレジスタ (SR_h) を設けている。そして、前方転送路 (F_i, FA_i, Fo) と垂直転送路 (V_i, Vo) を使用し、参照画像を2次元転送し、予測誤差を算出する。参照画像は、予測誤差算出中に水平サイドレジスタ (SR_h) 中に蓄えられていき、予測誤差算出後、1マクロブロック列の参照画像が蓄えられる。その後、水平サイドレジスタ中の参照画像が、逆方向転送路 (R_i, Ro) により再度初期化位置に戻される。PEアレー中のシフタは、2組のPEアレーの構造が異なるとき、逆方向転送時にデータの位相を補正するために使用される。本アレー構造により、まず同一のフィールド同士で探索を行い、再度この同じサーチウィンドウを使用し、テンプレート画像の入れ替えを行った後、異なるフィールド同士で探索を行う。このようにしてメモリバンド幅を増やすことなく1本のフレームベクトルおよび4本のフィールドベクトルを求めることができる。

4.5 FME[3]

FMEチップは、CMEで探索された粗い動きベクトルの周辺、水平-4～+4画素、垂直-2～+2画素を1/2画素精度でフルサーチ探索する。FMEもCMEと同じく、図11に示すような2次元シストリックアレーにより構成される。

各プロセッサエレメントは、図12の構成をしていて、 H, V, D バスを介して隣接するプロセッサから図11に示すような差分値が入力され、前節3.2.2(4)の式に示されるように1/2画素画像の差分値の算出が行われる。このように1/2画素画像を生成することなく、整数画素と1/2画素の予測誤差が同時に算出される。

さらに図13に示すような参照画像と符号化画像の前処理部で、前節3.2.2(5)の式に示されるような前処理が行われ、1/2画素精度の探索とデュアルプライム探索を同じアレー上で切り替えて行えるようにした。

このようにFMEチップは、1回の参照画像読み出しにより、整数画素の予測誤差と3つの1/2画素位置の予測誤差を同時に算出する。またデュアルプライム用のベクトルを算出する場合も、1/2画素の場合と同様に、1回の参照画像読み出しにより、内挿画像を生成することなく探索を行うことができる。

4.6 実装結果

本チップセットを使用した ITU-R BT.601 レベルのエンコーダシステムの構成を図 13 に示す。

本エンコーダチップセットは、CME チップを $0.6\mu\text{m}$ CMOS プロセスで、PRE, COD, QVL および FME チップは、 $0.4\mu\text{m}$ CMOS プロセスで実現した。5 チップの総ゲート数は、表 5 に示すように約 1200kGates になった。また符号化処理は、36MHz で実行され、SDRAM メモリ間のデータ転送は、72MHz で行われる。またこれらのチップセットを使用したエンコーダシステムは、メインレベルの符号化をグルーロジック無しに、本チップセットと、3 個の SDRAM と、チップにダウンロードするための量子化制御用プログラムを内蔵した ROM のみで構成される。この構成で、前処理を含む符号化処理の全ての機能を実現している。

また、本チップセットを使用した HDTV レベルのエンコーダシステムの構成を図 14 に示す。図で示すように、部分画像の切り出し機能を PRE チップ内に実装し、部分画像毎に符号化されたビットストリームの連結機能を COD チップ内に実装した。また、量子化制御における隣接画像の符号化情報については、QVL チップ内に RISC チップの専用コミュニケーションバスを介した送受信機能を実装した。このように、本エンコーダチップセットは、特別な回路無しに並列接続を行うことで、HDTV レベルの符号化機能を実現している。

4.7 評価および考察

本チップセットを使用したエンコーダシステムの仕様を表 6 に示す。本チップセットでは、量子化制御に RISC チップを実装し、それ以外を専用回路により構成している。このため HDTV 拡張機能を有しながら、SDTV 符号化を、総ゲート数が約 1200k ゲートの 5 種類のチップにより実現している。

本チップセットでは、SDRAM 3 個のメモリのみで、符号化周波数 36MHz、メモリアクセス周波数 72MHz により、SDTV 符号化を実現している。

一方、従来のプロセッサ型のエンコーダチップは、約 2500k トランジスタ（約 630k ゲート）の規模で実現したことが報告されている[4]。このチップを使用し SDTV レベルの符号化を行う場合、動きベクトル探索を除いた符号化処理のみで 2 チップを必要としている。また、このときの動作周波数は 80MHz で、符号化のために多数のメモリが必要となっている。

このように専用回路と RISC を用いた本方式の優位性がわかる。

また、動きベクトル探索において、本チップセットの FME と CME では、2 個のチップにより探索範囲が水平方向 $-24*fd - 4 \sim +23*fd + 4$ 、垂直方向 $-18 \sim +17$ を実現している。またこのときのメモリバンド幅は、CME で約 32MByte/sec、FME で 1 回のフレーム探索と 2 回のフィールド探索を行う場合、約 47MByte/sec、デュアルプライム予測の DMV (Differential Motion Vector) 探索の場合約 5MByte/sec で可能で、全探索方式に比べて、約 $1/10$ (デュアルプライム予測を含む場合約 $1/8.3$) と極めて小さい値を実現している。

また、このときの探索処理量も 33GOPS と約 $1/8.5$ を実現した。このため動きベクトルチップの総ゲート数は、768k ゲートとなり、上記探索の動作周波数は、CME が約 24MHz で FME が約 24MHz（デュアルプライム予測を含む場合は約 30MHz）で探索可能となり、極めて低い動作スピードで動きベクトル探索を実現している。

一方、従来の全探索方式を採用した動きベクトル探索チップ^[5]では、8 個のチップで水平方向 -63.5 から +63.5、垂直方向 -15.5 から +15.5 の探索が可能で、チップのトランジスタ数は、850k トランジスタ（約 210k ゲート）で、SDTV レベルの探索時の動作周波数は 40.5MHz で、このときの参照画像入力のバンド幅は、162MByte/sec となっている。

このように本方式では、ゲート数、動作周波数、およびメモリバンド幅共に、全探索方式に比べて極めて小さな値で実現していることがわかる。

さらに、本構成によるエンコーダシステムによる SNR 劣化は、NTSC 標準画像を使用して評価した結果、TM5 シミュレーションモデルに対し 0.5dB 以内の劣化となり、目標通りの性能を達成できた。

5. むすび

MPEG2 リアルタイムエンコーダのためのチップセットを開発した。本チップセットは、PRE, COD, QVL, CME, FME の 5 種類のチップから構成される。

これらのチップセットを使用したエンコーダシステムは、メインレベルの符号化をグルーロジック無しに、本チップセットと、3 個の SDRAM と、チップにダウンロードするための量子化制御用プログラムを内蔵した ROM のみで前処理を含む符号化処理の全ての機能を実現している。

本構成によるエンコーダシステムによる SNR 劣化は、TM5 シミュレーションモデルに対し 0.5dB 以内の劣化となり、目標通りの性能を達成できた。

さらに本チップセットは、画像のレベルに応じた拡張機能を有し、8 組のチップセットを並列接続することで、HDTV レベルの符号化を実現した。

参考資料

- 1) ISO/IEC 13818-2 | ITU-T H.262: Coding of Motion Pictures and Associated Audio, (1995).
- 2) Kobayashi, Saito, Kimura, Nakatomi, Nagai, Arai, Wuertele, Fujiwara, Nishi, Okada: "VLSI Implementation of a Complete Chip Set for an MPEG2 Real-Time Encoder", ISCAS (May-1996)
- 3) Kobayashi, Wuertele, Okada, Otsubo, Asada: "Algorithm and Architecture of LSIs for MPEG2 Motion Vector Estimation", PCS'96, Australia (March-1996)
- 4) M.Tokumura, et. al., "A Video Digital Signal Processor with a Vector-Pipeline Architecture", ISSCC Digest of Technical Paper, (Feb., 1994)
- 5) K.Ishihara, et. al., "A Half-pel Precision MPEG2 Motion Estimation Processor with Concurrent Three-

vector Search Scheme”, ISSCC Digest of Technical Paper, (Feb., 1995)

**Table 1 Required performance of processing
(ITU-R BT.601)**

Function	Performance [MOPS]
Pre-processing	200
Predictive coding	100
ME*1	280,000
DCT	1,280
IDCT	1,280
Q	40
IQ	40
VLC	400
Q control	80

*1 (M=3)

Horizontal range:-24*(frame distance) ~ +23*(frame distance)

Vertical range:-18 ~ +17

Table 2 Requirements of Memory Capacity

Function	Capacity [MByte]		Remarks
	NTSC	PAL	
Preprocessing	2.08	2.50	4 frames + 4 line(C)*2
Predictive frame	1.04	1.24	2 frames
VBV Buffer	0.23	0.23	
ME*1	0.69	0.83	2 frames (Y)

*1 motion vector searches using original memory

*2 using 4 tap filter for the conversion from 4:2:2 to 4:2:0

Table 3 Requirements of Memory Bandwidth

Function	Bandwidth [MByte/sec]		
	I	P	B
Preprocessing	80	80	80
MC	16	35	39
VBV Buffer	4	4	4
ME*1	0	780	780

*1 (M=3)

Horizontal range: $-24 * (\text{frame distance}) \sim +23 * (\text{frame distance})$

Vertical range: $-18 \sim +17$

**Table 4 Search Range of Frame Distance Adaptive
Coarse/Fine Search Method**

Distance between Template and Search Frames	Horizontal Search Range	Precision of Motion Vector*
1 Frame	-24 ~ +22	2 pels
2 Frames	-48 ~ +44	4 pels
3 Frames	-72 ~ +70	Near(-24 ~ +22): 2 pels Far: 4 pels

* accounting for 1/2 Horizontal Subsampling Hierarchical Search

Table 4 Implementation Results

LSI	Process	Number of Gates	Package
PRE	0.4 micron CMOS	135 Kgates + RAM	240pins Plastic QFP
COD	0.4 micron CMOS	115 Kgates + RAM	240pins Plastic QFP
QVL	0.4 micron CMOS	158 Kgates + RAM	240pins Plastic QFP
FME	0.4 micron CMOS	150 Kgates + RAM	160pins Plastic QFP
CME	0.6 micron CMOS	618 Kgates	168pins Plastic QFP

Table 5 Functional Characteristics of Encoder Chipset

Algorithm	MPEG1
	MPEG2 MP@ML(1 Chipset) MP@HL (8 Chipsets)
Resolution	8 Chipsets: 1920*1080 pixel
	1 Chipset: 720*480/720*576/352*240 pixel etc.
Frame Rate	30, 25, 24 frame/sec
Bitrate	HDTV: Max 60 Mbit/sec
	ITU-R BT.601: Max 20 Mbit/sec (including P,B) Max 60 Mbit/sec (only I)
Pic. Structure	Frame Structure
Picture Type	I/P/B Picture, P pic. with INTRA Slice
Motion Estimation	Two Step Hierarchical coarse/fine search
	Search Range: -24*fd-4/-22*fd+4 (Hor.) -18/+17(Vert.)
	Accuracy: Integer Pel/Half Pel
DCT	Frame/Field DCT
	Zig-Zag/Alternate Scan
Coding Type	Intra, Forward/BackwardBidirectional Pred.
Pred. Type	Frame/Field Prediction

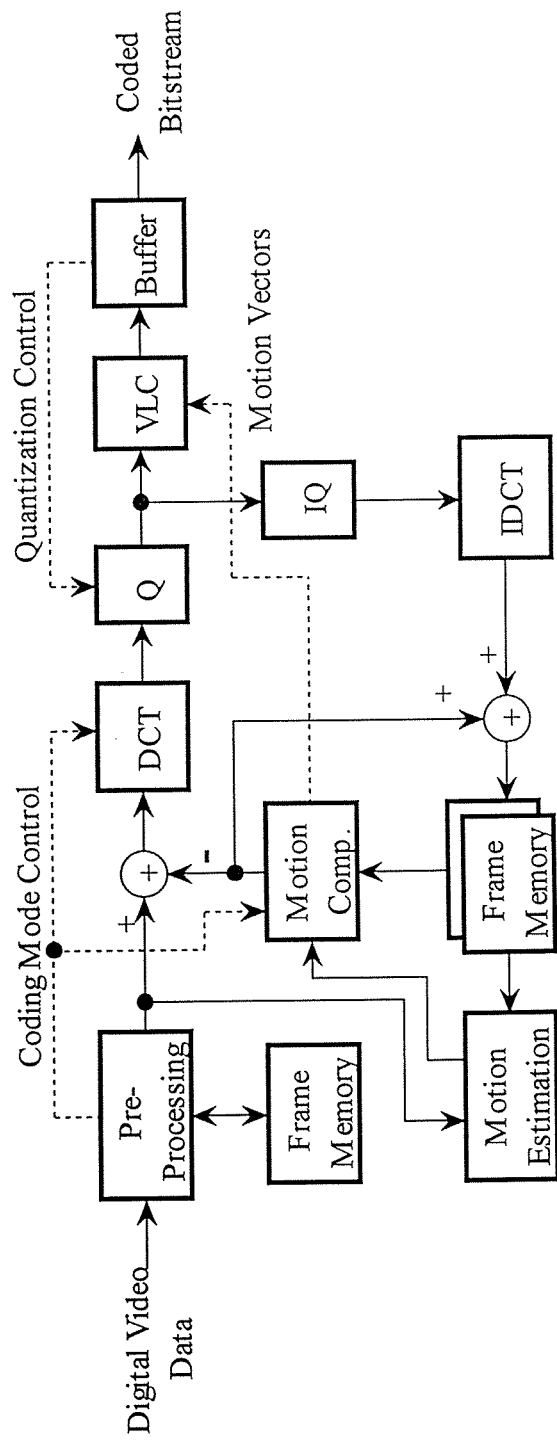


Fig.1 MPEG Encoder

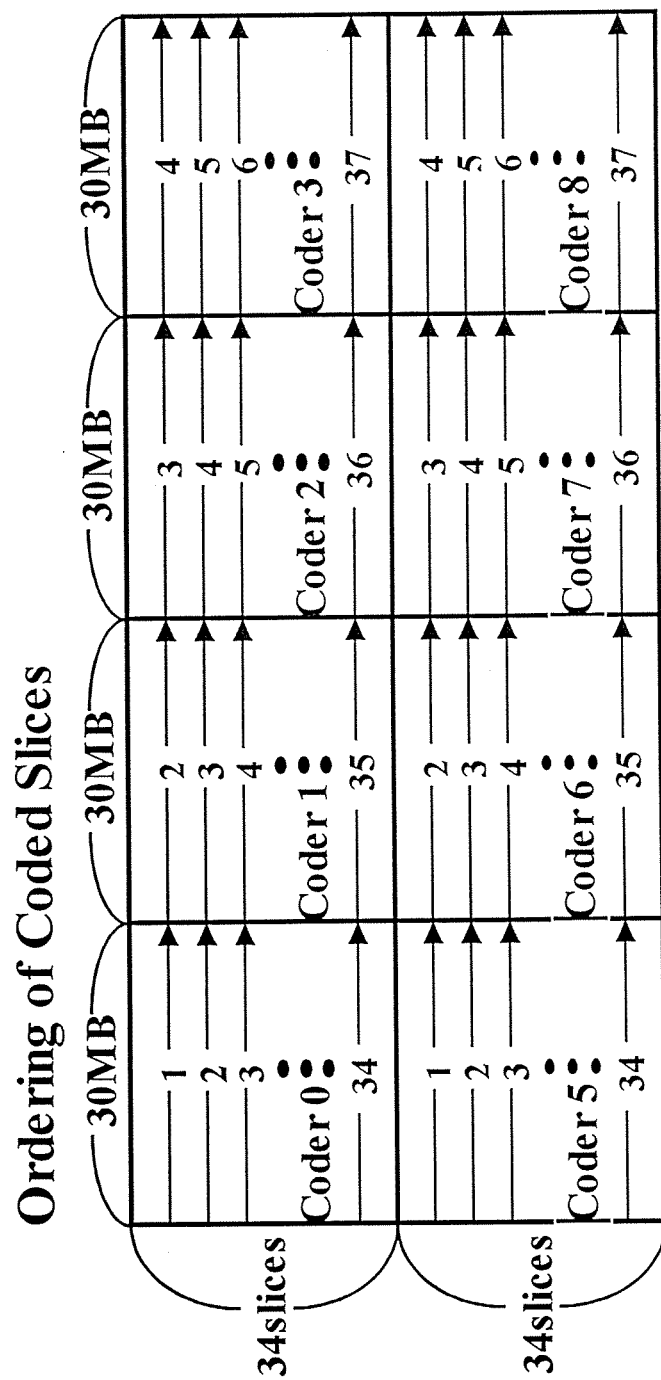
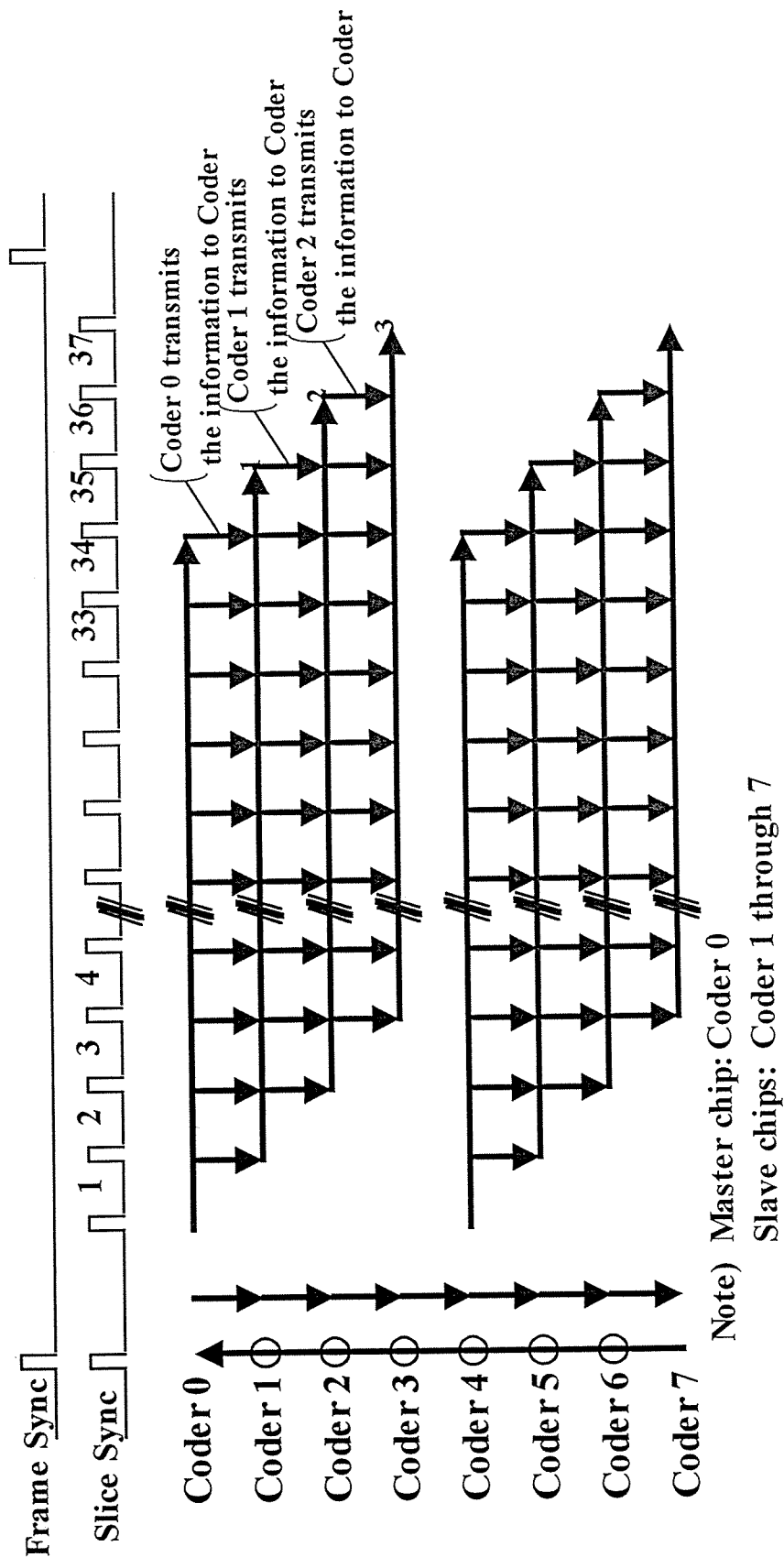


Fig.2 Segmentation of HDTV Encoding



**Fig.3 Transmission of Coding Information
in HDTV Encoding**

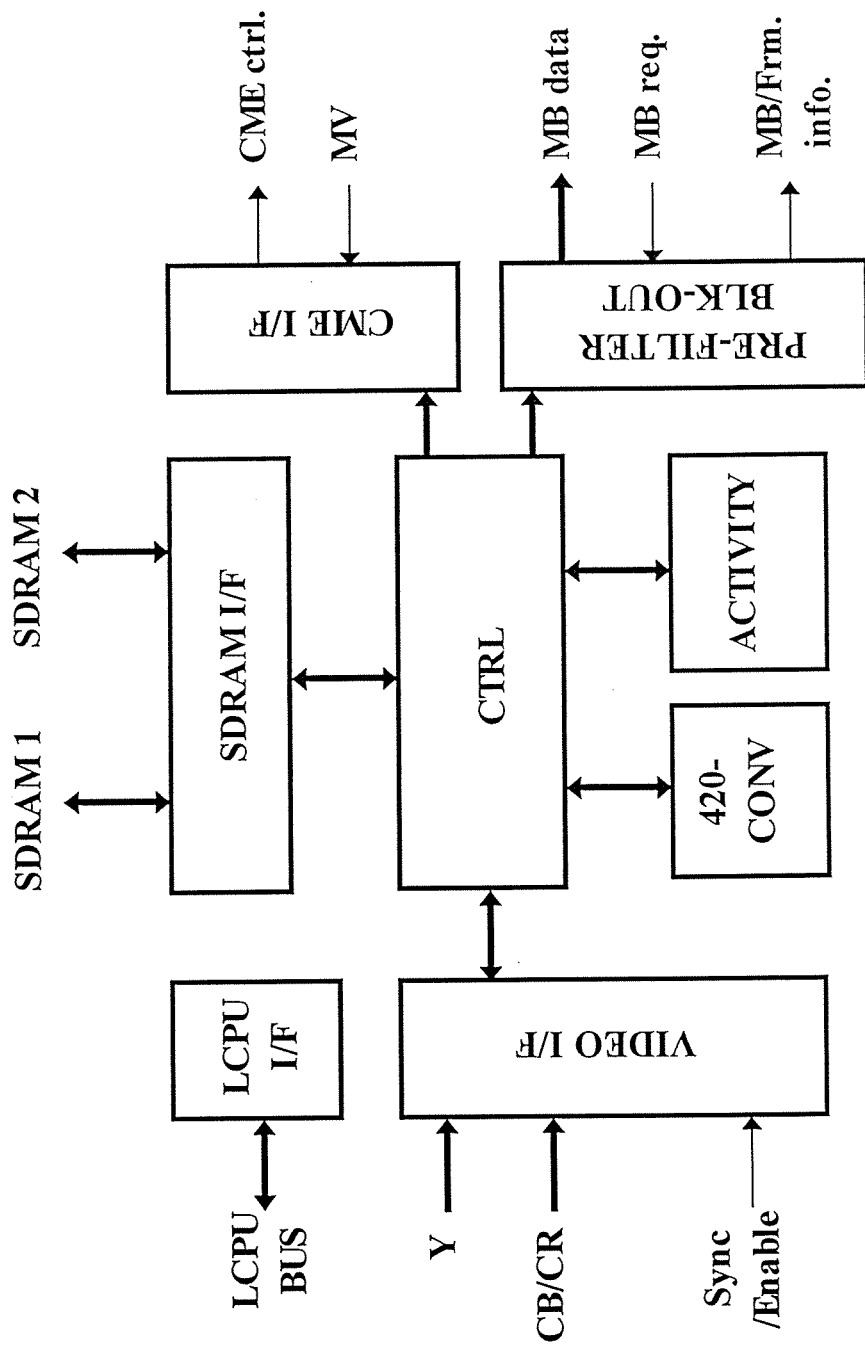


Fig.4 Block Diagram of PRE-LSI

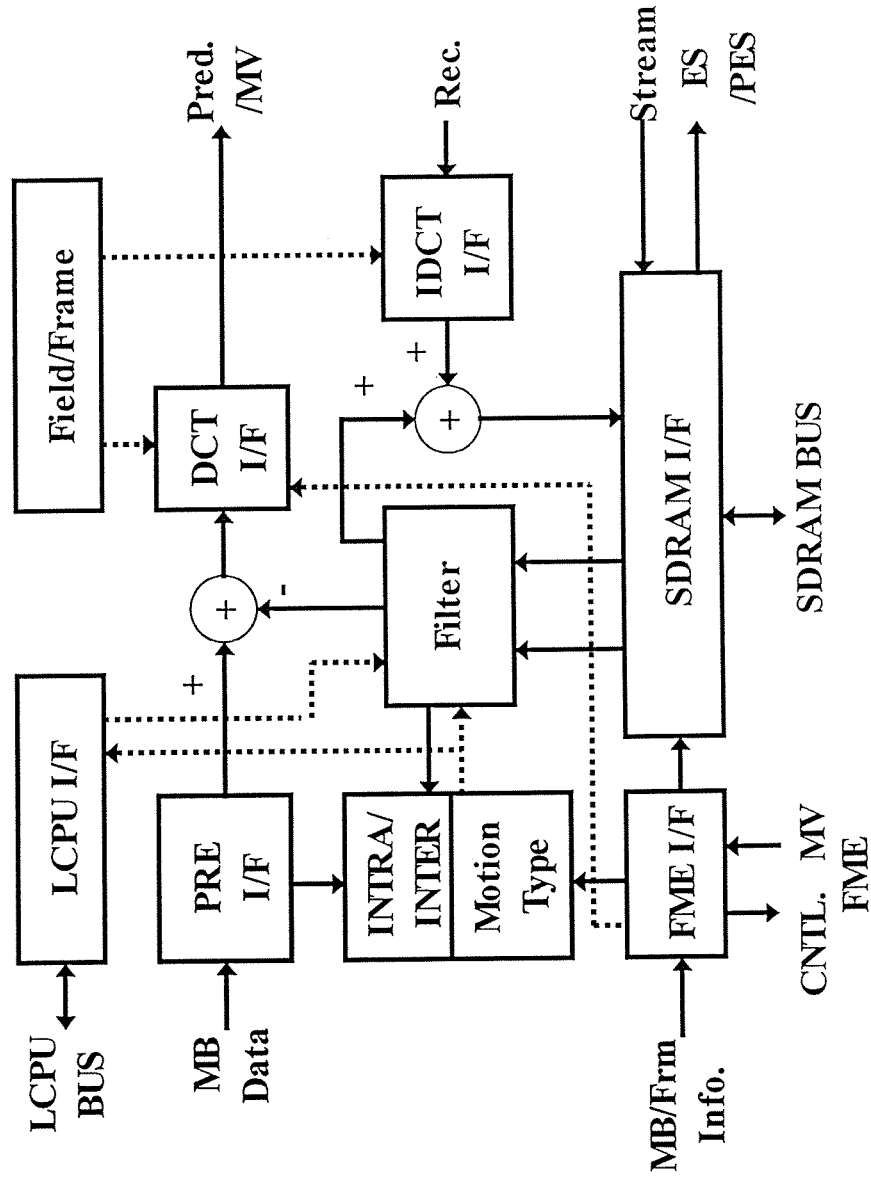


Fig.5 Block Diagram of COD-LSI

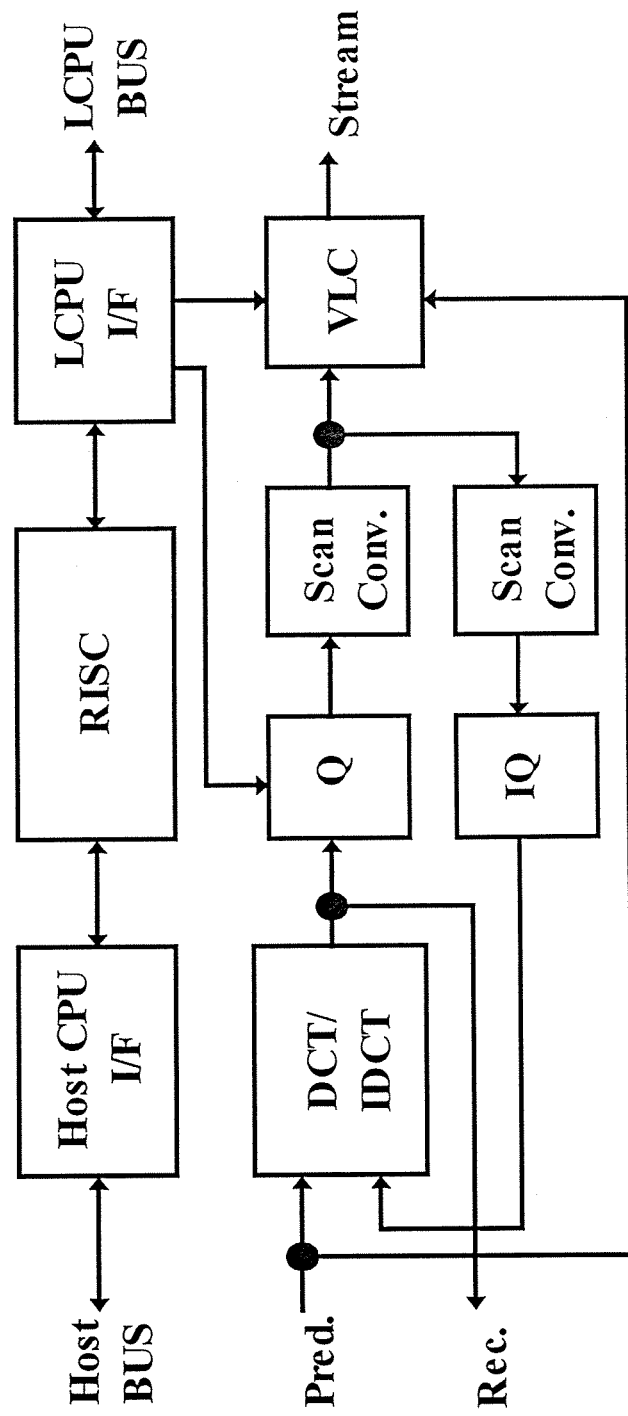


Fig.6 Block Diagram of QVL-LSI

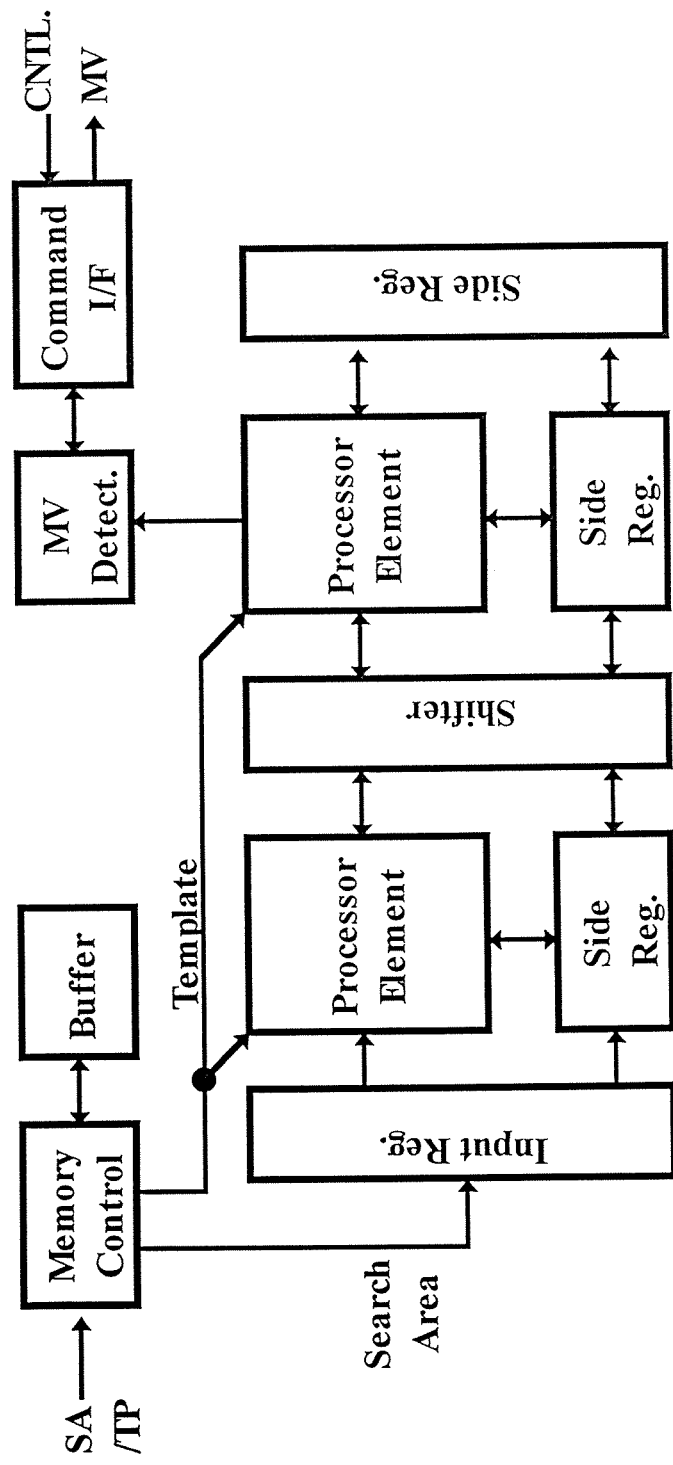


Fig.7 Block Diagram of CME-LSI

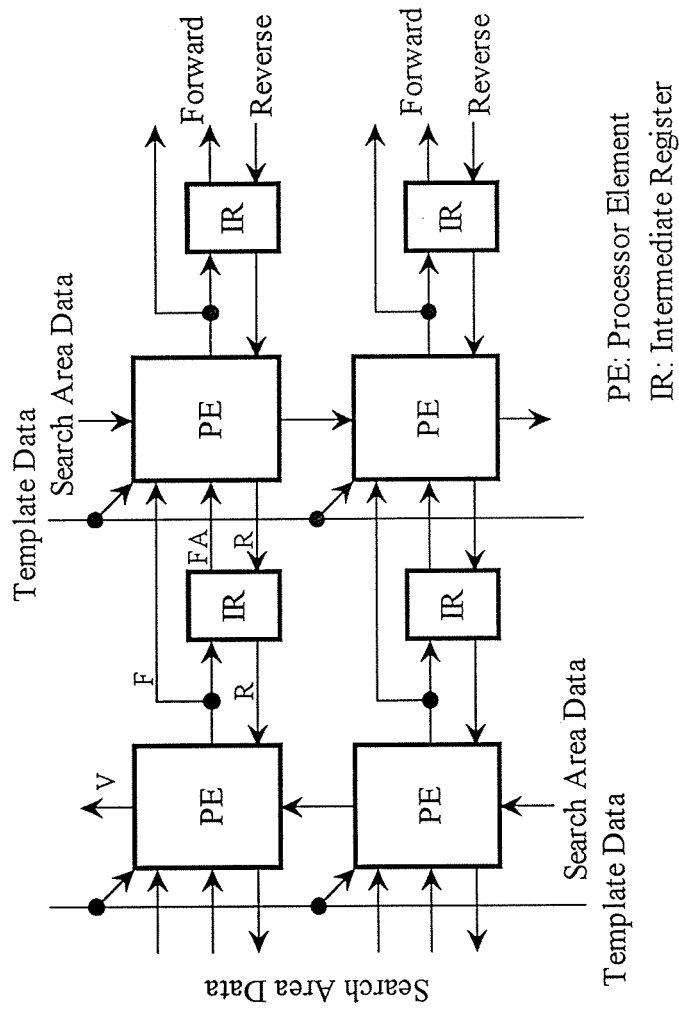


Fig.8 Array Architecture of the CME

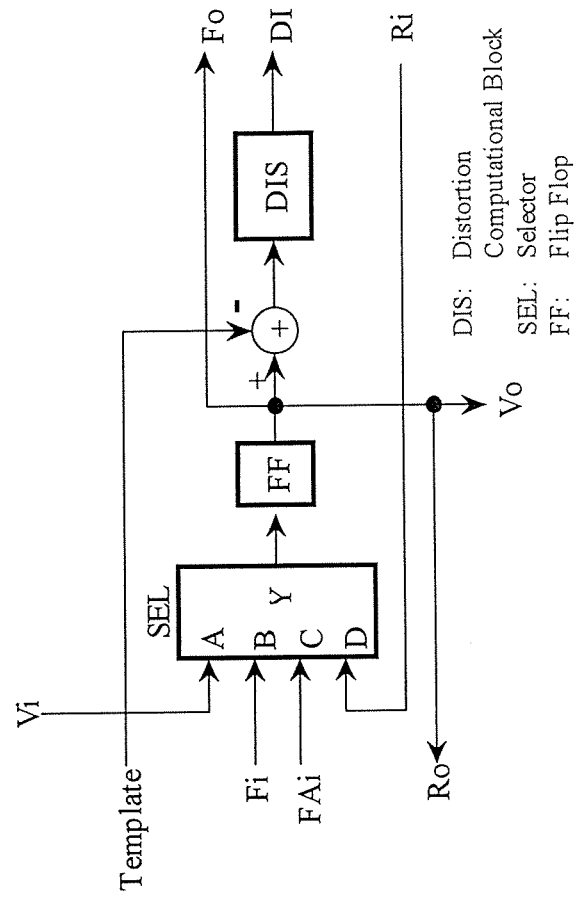


Fig.9 Processor Element of the CME

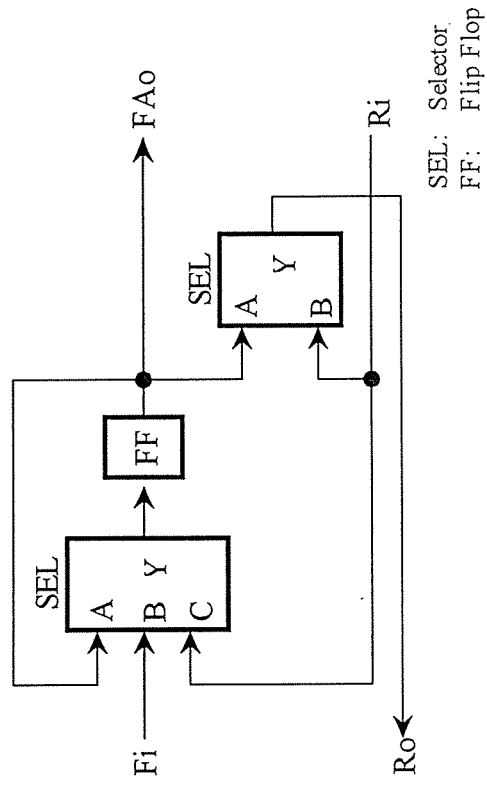


Fig.10 Intermediate Register of the CME

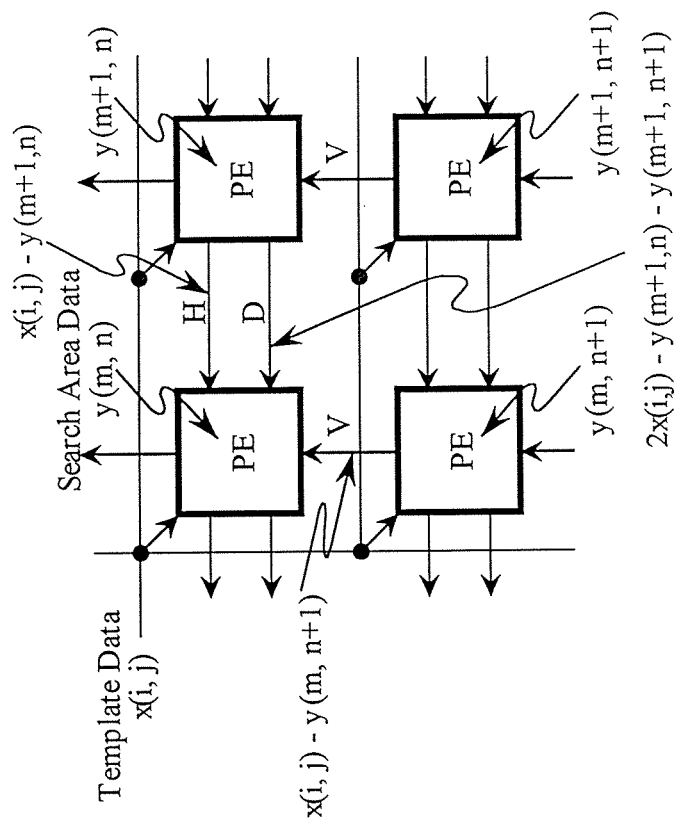


Fig.11 Array Architecture of the FME

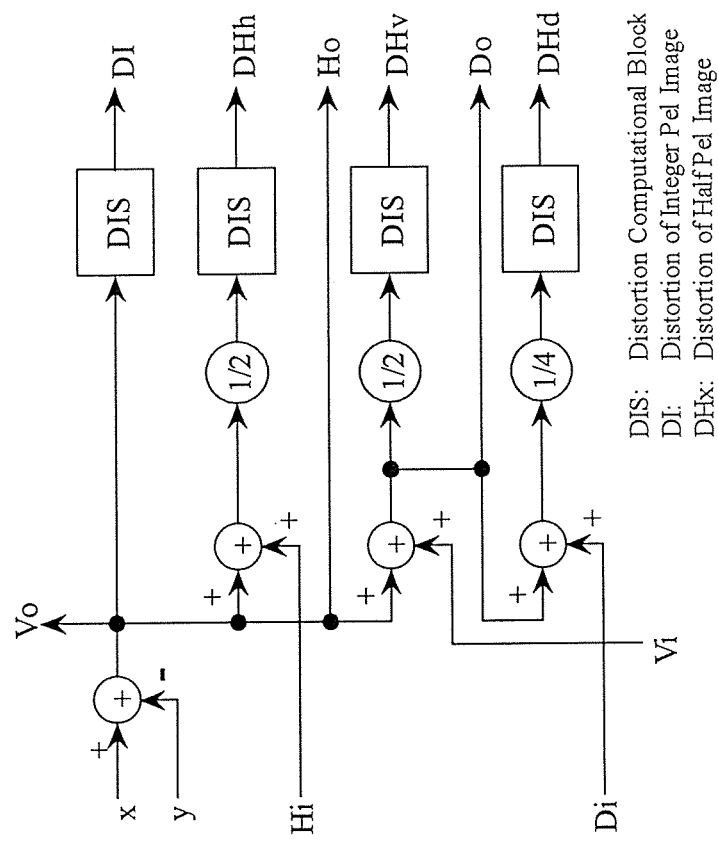
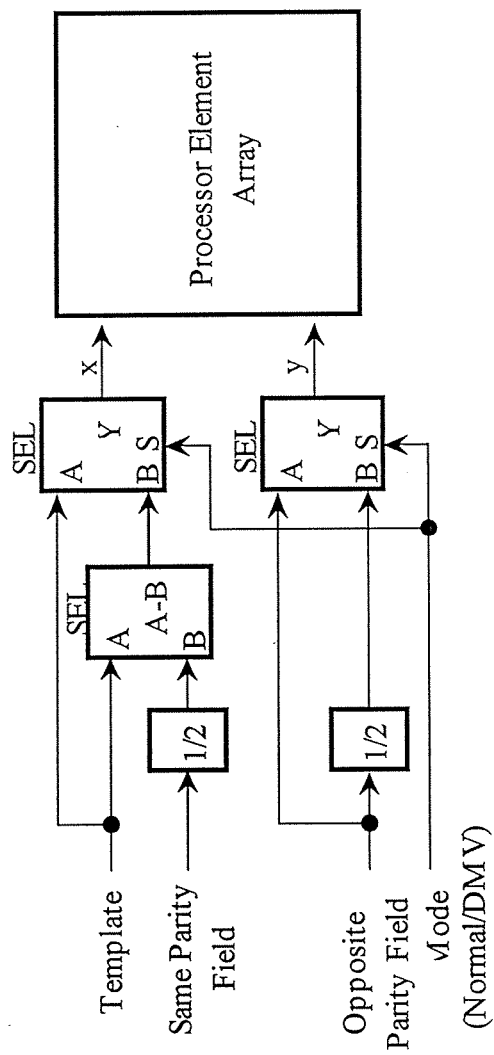


Fig.12 One Processor Element of the FME



**Fig.13 Input Reformatting for
Dual Prime Vector Evaluation**

(Normal/DMV)

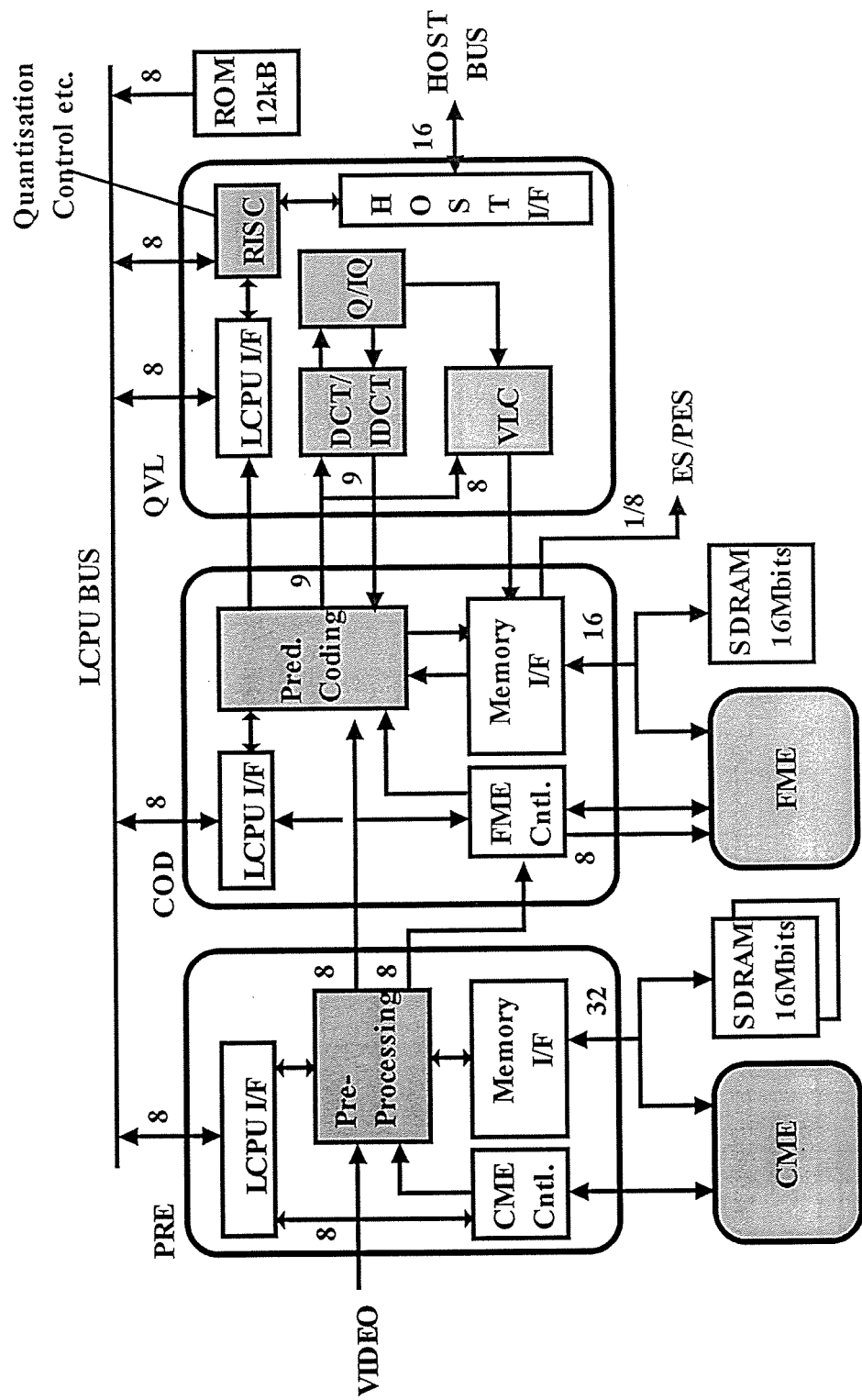


Fig.14 System Block Diagram of MPEG2 Video Encoder

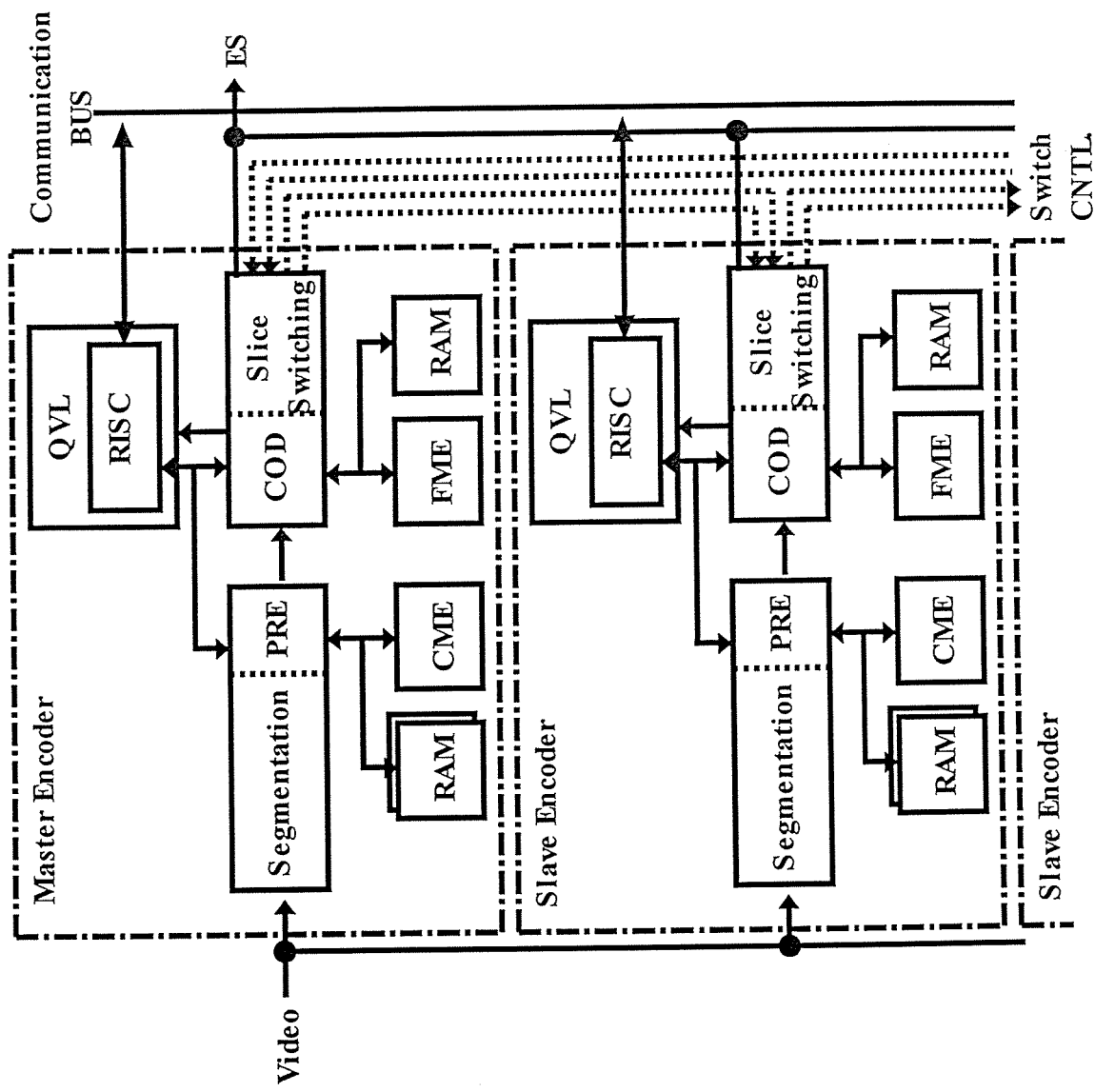


Fig.15 System Block Diagram of HDTV Encoder