

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号
特表2005-500754
(P2005-500754A)

(43) 公表日 平成17年1月6日(2005. 1. 6)

(51) Int.Cl.⁷
H04N 7/32
H03M 7/36

F I
H04N 7/137
H03M 7/36

Z
5C059
5J064

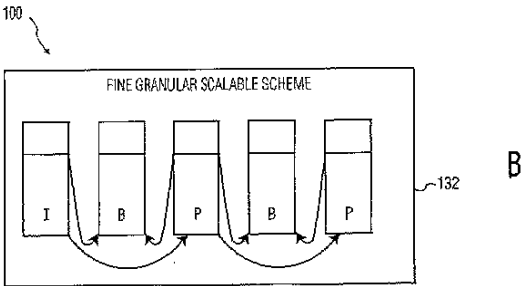
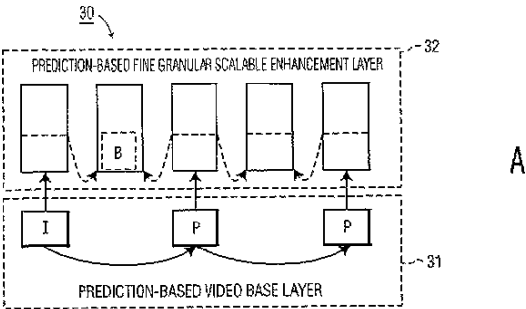
テーマコード (参考)

審査請求 未請求 予備審査請求 未請求 (全 45 頁)	
(21) 出願番号 特願2003-521624 (P2003-521624)	(71) 出願人 590000248
(86) (22) 出願日 平成14年7月11日 (2002. 7. 11)	コーニンクレッカ フィリップス エレク
(85) 翻訳文提出日 平成16年2月10日 (2004. 2. 10)	トロニクス エヌ ヴィ
(86) 国際出願番号 PCT/IB2002/002924	Koninklijke Philips
(87) 国際公開番号 W02003/017672	Electronics N. V.
(87) 国際公開日 平成15年2月27日 (2003. 2. 27)	オランダ国 5621 ペーアー アイン
(31) 優先権主張番号 09/930, 672	ドーフエン フルーネヴァウツウェッハ
(32) 優先日 平成13年8月15日 (2001. 8. 15)	1
(33) 優先権主張国 米国 (US)	Groenewoudseweg 1, 5
(81) 指定国 EP (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), CN, JP, KR	621 BA Eindhoven, The Netherlands
	(74) 代理人 100070150 弁理士 伊東 忠彦
	(74) 代理人 100091214 弁理士 大貫 進介
	最終頁に続く

(54) 【発明の名称】 動き補償を備えた完全に組み込まれた FGS 映像符号化

(57) 【要約】

スケーラブル・コーデックで完全に符号化された双方向予測フレーム（Bフレーム）又は予測フレーム及び双方向予測フレーム（P及びBフレーム）を生成する単一の動き補償ループを有するスケーラブル映像符号化の仕組みである。



【特許請求の範囲】

【請求項 1】

符号化されていない映像をエンコードして拡張基本レイヤの参照フレームを生成するステップであって、それぞれの拡張基本レイヤの参照フレームが基本レイヤの参照フレームと、関連する上位レイヤの参照フレームの少なくとも一部とを有するステップと、前記符号化されていない映像と前記拡張基本レイヤの参照フレームからフレームの残差を予測するステップとを有する映像符号化方法。

【請求項 2】

請求項 1 に記載の映像符号化方法であって、

DCTに基づくコーデック又はウェーブレット (wavelet) に基づくコーデックを有するグループから選択されたスケーラブル・コーデックで、前記フレームの残差を符号化し、上位レイヤのフレームを生成するステップを更に有する方法。 10

【請求項 3】

請求項 1 に記載の映像符号化方法であって、

細粒スケーラブル (fine granular scalable) ・コーデックで、前記フレームの残差を符号化し、細粒スケーラブル (fine granular scalable) 上位レイヤのフレームを生成するステップを更に有する方法。

【請求項 4】

請求項 1 に記載の映像符号化方法であって、

前記フレームの残差が B フレームの残差を有する方法。 20

【請求項 5】

請求項 4 に記載の映像符号化方法であって、

前記フレームの残差が P フレームの残差を更に有する方法。

【請求項 6】

請求項 1 に記載の映像符号化方法であって、

前記フレームの残差が P フレームの残差を有する方法。

【請求項 7】

基本レイヤのストリームと上位レイヤのストリームとを有する圧縮された映像をデコードする方法であって、

前記基本レイヤと上位レイヤのストリームをデコードして拡張基本レイヤの参照フレームを生成するステップであって、それぞれの前記拡張基本レイヤの参照フレームが基本レイヤの参照フレームと、関連する上位レイヤの参照フレームの少なくとも一部とを有するステップと、前記拡張基本レイヤの参照フレームからフレームの残差を予測するステップとを有する方法。 30

【請求項 8】

請求項 7 に記載の映像符号化方法であって、

DCTに基づくデコード又はウェーブレット (wavelet) に基づくデコードを有するグループから選択されたスケーラブル・デコードで、前記フレームの残差をデコードするステップを更に有する方法。 40

【請求項 9】

請求項 8 に記載の映像符号化方法であって、

前記フレームの残差から上位レイヤのフレームを生成するステップと、

前記基本レイヤのフレームと前記上位レイヤのフレームから拡張された映像を生成する生成するステップと

を更に有する方法。

【請求項 10】

請求項 7 に記載の映像符号化方法であって、

前記フレームの残差が B フレームの残差を有する方法。

【請求項 11】

請求項 10 に記載の映像符号化方法であって、
前記フレームの残差が P フレームの残差を更に有する方法。

【請求項 12】

請求項 7 に記載の方法であって、
前記フレームの残差が P フレームの残差を有する方法。

【請求項 13】

映像を符号化するメモリ媒体であって、
符号化されていない映像をエンコードして拡張基本レイヤの参照フレームを生成するコードであって、それぞれの前記拡張基本レイヤの参照フレームが基本レイヤの参照フレームと、関連する上位レイヤの参照フレームの少なくとも一部とを有するコードと、
前記符号化されていない映像と前記拡張基本レイヤの参照フレームからフレームの残差を予測するコードと
を有するメモリ媒体。

10

【請求項 14】

請求項 13 に記載の映像を符号化するメモリ媒体であって、
前記フレームの残差をスケーラブル・エンコードするコードを更に有するメモリ媒体。

【請求項 15】

請求項 13 に記載の映像を符号化するメモリ媒体であって、
前記フレームの残差を細粒スケーラブル (fine granular scalable) エンコードするコードを更に有するメモリ媒体。

20

【請求項 16】

請求項 13 に記載の映像を符号化するメモリ媒体であって、
前記フレームの残差が B フレームの残差を有するメモリ媒体。

【請求項 17】

請求項 13 に記載の映像を符号化するメモリ媒体であって、
前記フレームの残差が P フレームの残差を更に有するメモリ媒体。

【請求項 18】

請求項 13 に記載の映像を符号化するメモリ媒体であって、
前記フレームの残差が P フレームの残差を有するメモリ媒体。

【請求項 19】

基本レイヤのストリームと上位レイヤのストリームとを有する圧縮された映像をデコードするメモリ媒体であって、
前記基本レイヤと上位レイヤのストリームをデコードして拡張基本レイヤの参照フレームを生成するコードであって、それぞれの前記拡張基本レイヤの参照フレームが基本レイヤの参照フレームと、関連する上位レイヤの参照フレームの少なくとも一部とを有するコードと、
前記拡張基本レイヤの参照フレームからフレームの残差を予測するステップと
を有するメモリ媒体。

30

【請求項 20】

請求項 19 に記載の圧縮された映像をデコードするメモリ媒体であって、
前記フレームの残差をスケーラブル・デコードするコードを更に有し、
前記スケーラブル・デコードするコードは、DCT に基づくコード又はウェーブレット (wavelet) に基づくコードを有するグループから選択されるメモリ媒体。

40

【請求項 21】

請求項 20 に記載の圧縮された映像をデコードするメモリ媒体であって、
前記フレームの残差から上位レイヤのフレームを生成するコードと、
前記基本レイヤのフレームと前記上位レイヤのフレームから拡張された映像を生成する生成するコードと
を更に有するメモリ媒体。

【請求項 22】

50

請求項 19 に記載の圧縮された映像をデコードするメモリ媒体であって、
前記フレームの残差が B フレームの残差を有するメモリ媒体。

【請求項 23】

請求項 22 に記載の圧縮された映像をデコードするメモリ媒体であって、
前記フレームの残差が P フレームの残差を更に有するメモリ媒体。

【請求項 24】

請求項 19 に記載の圧縮された映像をデコードするメモリ媒体であって、
前記フレームの残差が P フレームの残差を有するメモリ媒体。

【請求項 25】

映像を符号化する装置であって、
符号化されていない映像をエンコードして拡張基本レイヤの参照フレームを生成する手段
であって、それぞれの前記拡張基本レイヤの参照フレームが基本レイヤの参照フレームと
、関連する上位レイヤの参照フレームの少なくとも一部とを有する手段と、
前記符号化されていない映像と前記拡張基本レイヤの参照フレームからフレームの残差を
予測する手段と
を有する装置。

10

【請求項 26】

請求項 25 に記載の映像を符号化する装置であって、
前記フレームの残差をスケーラブル・エンコードする手段を更に有する装置。

【請求項 27】

請求項 25 に記載の映像を符号化する装置であって、
前記フレームの残差を細粒スケーラブル (fine granular scalable) エンコードするコ
ードを更に有する装置。

20

【請求項 28】

請求項 25 に記載の映像を符号化する装置であって、
前記フレームの残差が B フレームの残差を有する装置。

【請求項 29】

請求項 28 に記載の映像を符号化する装置であって、
前記フレームの残差が P フレームの残差を更に有する装置。

【請求項 30】

請求項 25 に記載の映像を符号化する装置であって、
前記フレームの残差が P フレームの残差を有する装置。

30

【請求項 31】

基本レイヤのストリームと上位レイヤのストリームとを有する圧縮された映像をデコード
する装置であって、
前記基本レイヤと上位レイヤのストリームをデコードして拡張基本レイヤの参照フレーム
を生成する手段であって、それぞれの前記拡張基本レイヤの参照フレームが基本レイヤの
参照フレームと、関連する上位レイヤの参照フレームの少なくとも一部とを有する手段と
、
前記拡張基本レイヤの参照フレームからフレームの残差を予測する手段と
を有する装置。

40

【請求項 32】

請求項 31 に記載の圧縮された映像をデコードする装置であって、
前記フレームの残差をデコードするスケーラブル・デコード手段を更に有し、
前記スケーラブル・デコード手段は、DCTに基づくデコード手段又はウェーブレット (wav
elet) に基づくデコード手段を有するグループから選択される装置。

【請求項 33】

請求項 32 に記載の圧縮された映像をデコードする装置であって、
前記フレームの残差から上位レイヤのフレームを生成する手段と、
前記基本レイヤのフレームと前記上位レイヤのフレームから拡張された映像を生成する生

50

成する手段と
を更に有する装置。

【請求項 3 4】

請求項 3 1 に記載の圧縮された映像をデコードする装置であって、
前記フレームの残差が B フレームの残差を有する装置。

【請求項 3 5】

請求項 3 4 に記載の圧縮された映像をデコードする装置であって、
前記フレームの残差が P フレームの残差を更に有する装置。

【請求項 3 6】

請求項 3 1 に記載の圧縮された映像をデコードする装置であって、
前記フレームの残差が P フレームの残差を有する装置。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は映像符号化に関し、特に、細粒スケーラブル (FGS) 符号化 (fine granular scalable coding) で完全に符号化された双方向予測フレーム (B フレーム) 又は予測フレーム及び双方向予測フレーム (P 及び B フレーム) を生成する単一の動き補償ループを用いるスケーラブル映像符号化の仕組みに関するものである。

【背景技術】

【0002】

20

スケーラブル上位レイヤの映像符号化が、インターネットのような変化する帯域を有するコンピュータネットワーク上で送信される映像の圧縮に用いられてきた。(ISO MPEG-4 標準で採用されている) FGS 符号化技術で用いる現在の上位レイヤの映像符号化の仕組みが図 1 に示されている。図示の通り、映像符号化の仕組み 10 は、ビットレート R_{BL} で符号化された予測に基づく基本レイヤ 11 と、 R_{EL} で符号化された FGS 上位レイヤを含む。

予測に基づく基本レイヤ 11 は、フレーム内符号化 I フレームと、動き推定補償を用いて以前の I 又は P フレームから時間的に予測されたフレーム間符号化 P フレームと、動き推定補償を用いて B フレームに隣接した以前と次のフレームの双方から時間的に予測されたフレーム間符号化双方向 B フレームとを含む。基本レイヤ 11 における予測符号化及び / 又は内挿的符号化、すなわち動き推定と対応する補償は、基本レイヤのフレームのみが予測に使われるため、ある程度のみ、その時間的な冗長を減少する。

30

【0003】

上位レイヤ 12 は、それぞれ再構成された基本レイヤのフレームをそれぞれの元のフレームから差し引くことによって導き出された FGS 上位レイヤの I 及び P 及び B フレームを含む (前記の差し引くことは、動き補償領域でも行われる)。従って、上位レイヤにおける FGS 上位レイヤの I 及び P 及び B フレームは、動き補償されない (FGS 残差が同時にフレームから引き出される)。このことの主要な理由は、送信時に利用可能な帯域に応じて、個別に各 FGS 上位レイヤのフレームの切捨てを可能にする柔軟性を提供するためである。特に、上位レイヤ 12 の細粒スケーラブル符号化 (fine granular scalable coding) は、FGS 映像ストリームが $R_{min} = R_{BL}$ から $R_{max} = R_{BL} + R_{EL}$ までの範囲の利用可能な帯域を備えた何らかのネットワーク上で送信されることを可能にする。例えば、送信機と受信機の間で利用可能な帯域が $B = R$ である場合、送信機は、速度 R_{BL} で基本レイヤを送信し、速度 $R_{EL} = R - R_{BL}$ で上位レイヤのフレームの一部のみを送信する。図 1 からわかる通り、上位レイヤにおける FGS 上位レイヤの一部は、送信において細粒スケーラブル符号化 (fine granular scalable coding) 方法で選択され得る。従って、単一の上位レイヤで広範囲の送信帯域に対応する柔軟性のために、全体の送信ビットレートは $R = R_{BL} + R_{EL}$ である。

40

【0004】

図 2 は、図 1 の映像符号化の仕組みの基本レイヤ 11 と上位レイヤ 12 を符号化するため

50

の従来のFGSエンコーダのブロック図である。図に示すように、フレーム i の上位レイヤの残差 ($FGSR(i)$) は $MCR(i) - MCRQ(i)$ と等しく、 $MCR(i)$ はフレーム i の動き補償された残差であり、 $MCRQ(i)$ は量子化及び逆量子化処理の後のフレーム i の動き補償された残差である。

【0005】

図1の現在のFGS上位レイヤの映像符号化の仕組み10は非常に柔軟性があるが、同じ送信ビットレートで機能するスケーラブルではないコードと比較して、映像の画像品質に関する性能が比較的低いという不利を有する。画像品質の減少は、上位レイヤ12の細粒スケーラブル符号化 (fine granular scalable coding) のためではなく、主に上位レイヤ12内のFGS残差フレーム間の時間的な冗長の減少された利用のためである。特に、上位レイヤ12のFGS上位レイヤのフレームは、それぞれの基本レイヤのI及びP及びBフレームの動き補償された残差からのみ導き出され、上位レイヤ12の他のFGS上位レイヤのフレーム、又は基本レイヤ11の他のフレームを予測するためにFGS上位レイヤのフレームは用いられない。

10

【0006】

従って、改善された映像の画像品質を有するスケーラブル映像符号化の仕組みが必要とされる。

【発明の開示】

【課題を解決するための手段】

【0007】

本発明は、細粒スケーラブル (FGS) 符号化 (fine granular scalable coding) で完全に符号化された双方向予測フレーム (Bフレーム) 又は予測フレーム及び双方向予測フレーム (P及びBフレーム) を生成する単一の動き補償ループを用いるスケーラブル映像符号化の仕組みを対象とする。本発明の一形態は、符号化されていない映像をエンコードして拡張基本レイヤの参照フレームを生成するステップであって、それぞれの前記拡張基本レイヤの参照フレームが基本レイヤの参照フレームと、関連する上位レイヤの参照フレームの少なくとも一部とを有するステップと、前記符号化されていない映像と前記拡張基本レイヤの参照フレームからフレームの残差を予測するステップとを有する映像符号化方法を有する。

20

【0008】

本方法の他の形態は、基本レイヤのストリームと上位レイヤのストリームとを有する圧縮された映像をデコードする方法を有し、前記基本レイヤと上位レイヤのストリームをデコードして拡張基本レイヤの参照フレームを生成するステップであって、それぞれの前記拡張基本レイヤの参照フレームが基本レイヤの参照フレームと、関連する上位レイヤの参照フレームの少なくとも一部とを有するステップと、前記拡張基本レイヤの参照フレームからフレームの残差を予測するステップとを有する。

30

【0009】

本発明の更なる他の形態は、映像を符号化するメモリ媒体を有し、符号化されていない映像をエンコードして拡張基本レイヤの参照フレームを生成するコードであって、それぞれの前記拡張基本レイヤの参照フレームが基本レイヤの参照フレームと、関連する上位レイヤの参照フレームの少なくとも一部とを有するコードと、前記符号化されていない映像と前記拡張基本レイヤの参照フレームからフレームの残差を予測するコードとを有する。

40

【0010】

本発明の更なる他の形態は、基本レイヤのストリームと上位レイヤのストリームとを有する圧縮された映像をデコードするメモリ媒体を有し、前記基本レイヤと上位レイヤのストリームをデコードして拡張基本レイヤの参照フレームを生成するコードであって、それぞれの前記拡張基本レイヤの参照フレームが基本レイヤの参照フレームと、関連する上位レイヤの参照フレームの少なくとも一部とを有するコードと、前記拡張基本レイヤの参照フレームからフレームの残差を予測するコードとを有する。

【0011】

50

本発明の更なる他の形態は、映像を符号化する装置を有し、符号化されていない映像をエンコードして拡張基本レイヤの参照フレームを生成する手段であって、それぞれの前記拡張基本レイヤの参照フレームが基本レイヤの参照フレームと、関連する上位レイヤの参照フレームの少なくとも一部とを有する手段と、前記符号化されていない映像と前記拡張基本レイヤの参照フレームからフレームの残差を予測する手段とを有する。

【0012】

本発明の更なる他の形態は、基本レイヤのストリームと上位レイヤのストリームとを有する圧縮された映像をデコードする装置を有し、前記基本レイヤと上位レイヤのストリームをデコードして拡張基本レイヤの参照フレームを生成する手段であって、それぞれの前記拡張基本レイヤの参照フレームが基本レイヤの参照フレームと、関連する上位レイヤの参照フレームの少なくとも一部とを有する手段と、拡張基本レイヤの参照フレームからフレームの残差を予測する手段とを有する。

10

【発明を実施するための最良の形態】

【0013】

本発明の利点と本質と多様な追加の特徴が、添付の図面と共に詳細に説明される例示的な実施例の検討で、更に完全に現れる。図面において同様の参照番号は、図面を通じて同様の要素を特定する。

【0014】

図3Aは、本発明の第1の例示的な実施例によるスケーラブル映像符号化の仕組み30を示したものである。スケーラブル映像符号化の仕組み30は、予測に基づく基本レイヤ31と単一のループの予測に基づく上位レイヤ32とを有する。

20

【0015】

従来は基本レイヤの（スケーラブルではない）符号化の間に標準的な基本レイヤI及びP参照フレームから生成されたフレーム内符号化Iフレームとフレーム間符号化Pフレームを含むように、予測に基づく基本レイヤ31が符号化される。フレーム間符号化双方向Bフレームは基本レイヤで符号化されない。

【0016】

本発明の原理によると、基本レイヤの符号化の間に“向上した”又は“拡張された”基本レイヤのI及びP又はP及びP参照フレーム（以下、拡張基本レイヤのI及びP参照フレームと言う）から動き予測されたフレーム間符号化双方向Bフレームを含むように、予測に基づく上位レイヤ32が符号化される。それぞれの拡張基本レイヤの参照フレームは、標準的な基本レイヤの参照フレームと、関連する上位レイヤの参照フレームの少なくとも一部からなる（関連する上位レイヤの参照フレームの1つ以上のビットプレーン又はわずかのビットプレーンが用いられ得る）。

30

【0017】

従来はそれぞれの元の基本レイヤのフレームの残差からそれぞれの再構成された（デコードされた）基本レイヤのフレームの残差を差し引くことによって生成された上位レイヤのI及びPフレームを含むように、上位レイヤ32もまた符号化される。上位レイヤのI及びB及びPフレームは、何らかの適切なスケーラブル・コーデックで符号化され得る。例えば、スケーラブル・コーデックは、DCTに基づくコーデック（FGS）や、ウェーブレット（wavelet）に基づくコーデックや、何らかの他の組み込まれたコーデックである場合がある。図3Aに示された実施例において、スケーラブル・コーデックはFGSからなる。

40

【0018】

その技術に通常熟練した人は、本発明の映像符号化の仕組み30は、映像の画像品質を改善することがわかるだろう。このことは、映像符号化の仕組み30が上位レイヤのBフレームにおける時間的な冗長を減少するために拡張基本レイヤの参照フレームを用いているためである。

【0019】

図4は、図3Aのスケーラブル映像符号化の仕組みを作るために用いられ得る、本発明の例示的な実施例によるエンコーダ40のブロック図である。図に示すように、エンコーダ

50

40は、基本レイヤのエンコーダ41と上位レイヤのエンコーダ42とを含む。基本レイヤのエンコーダ41は、元の映像シーケンスと、フレームメモリ60に保存された基本レイヤ及び拡張基本レイヤの参照フレームとから動き情報(動きベクトル及び予測モード)を生成する動き推定手段43を含む。動き情報と従来の参照フレームとフレームメモリ60に保存された拡張基本レイヤのI及びP参照フレームとを用いて、前記動き情報は、従来の動き補償された基本レイヤの参照フレームと、本発明の拡張基本レイヤのI及びP参照フレームの動き補償されたバージョン(全て $Ref(i)$ で示される)とを生成する動き補償手段44に適用される。第1の減算手段45は、元の映像シーケンスから従来の動き補償された参照フレームを差し引き、基本レイヤのI及びPフレームの動き補償された残差を生成する。第1のフレームフロー制御装置62は、離散コサイン変換(DCT)エンコーダ46と、量子化手段47と、エントロピーエンコーダ48によって処理する基本レイヤのI及びPフレームの動き補償された残差 $MCR(i)$ の経路を定め、圧縮された基本レイヤのストリームの一部を形成する基本レイヤのI及びPフレームを生成する。動き推定手段43によって生成された動き情報もまた、マルチプレクサ49に適用され、基本レイヤのI及びPフレームと動き情報とを組み合わせ、圧縮された基本レイヤのストリームを完成する。量子化手段47の出力で生成された、量子化された基本レイヤのI及びPフレームの動き補償された残差 $MCR(i)$ は、逆量子化手段50で逆量子化され、逆DCTデコーダ51でデコードされる。この処理が、逆DCT51の出力で、基本レイヤのI及びPフレームの動き補償された残差の量子化/逆量子化されたバージョン $MCRQ(i)$ を生成する。逆DCT51の出力における量子化された/逆量子化された基本レイヤのI及びPフレームの動き補償された残差は、第1の加算手段61に適用され、前記第1の加算手段は対応する動き補償された基本レイヤの参照フレーム $Ref(i)$ とそれを合計し、それ故に、前述の通りフレームメモリ60に保存された従来の基本レイヤの参照フレームを生成する。

【0020】

量子化された/逆量子化された基本レイヤのI及びPフレームの動き補償された残差もまた、上位レイヤのエンコーダ42の第2の減算手段53に適用される。第2の減算手段53は、量子化された/逆量子化された基本レイヤのI及びPフレームの動き補償された残差を、対応する基本レイヤのI及びPフレームの動き補償された残差を差し引き、差分のI及びPフレームの残差を生成する。第2の減算手段53の出力は、FGSエンコーダ54又は同様のスケラブル・エンコーダによってスケラブル符号化が行われる。FGSエンコーダ54は、圧縮された上位レイヤのストリームの一部を形成するスケラブル(FGS)エンコードされたI及びPフレームを生成するために、従来のビットプレーンのDCTスキミングと従来のエントロピーエンコードに続いて、従来のDCTエンコードを用いる。マスキング装置55が、スケラブル・エンコードされたI及びPフレームの1つ以上の符号化ビットプレーンを受け取り、第3のフロー制御装置65を通じて選択的に経路を定められ、前記データを第2の加算手段56の第1の入力57に適用する。基本レイヤのエンコーダ41で生成された、I及びPフレームの動き補償された残差の量子化された/逆量子化されたバージョン $MCRQ(i)$ は、第2の加算手段56の第2の入力58に更に適用される。上位レイヤのエンコードされたI及びPフレームの1つ以上の符号化ビットプレーンと、それぞれのI及びPフレームの残差 $MCRQ(i)$ とを合計することにより、第2の加算手段56は上位レイヤのI及びP参照フレームを生成する。第2の加算手段56によって計算された上位レイヤのI及びP参照フレームは、基本レイヤのエンコーダ41の第3の加算手段52に適用される。第3の加算手段52は、上位レイヤのI及びP参照フレームと、対応する動き補償された基本レイヤのI及びP参照フレーム $Ref(i)$ と、対応する量子化された/逆量子化された動き補償された基本レイヤのI及びPフレームの残差とを合計し、拡張基本レイヤのI及びP参照フレームを生成し、それらはフレームメモリ60に保存される。

【0021】

動き補償手段44は、動き情報とフレームメモリ60に保存された拡張基本レイヤのI及びP参照フレームを用いて、拡張基本レイヤのI及びP参照フレームの動き補償されたバ

ージョンを生成する。第1の減算手段45は、元の映像シーケンスから動き補償された上位レイヤの参照フレームを差し引き、動き補償されたBフレームの残差を生成する。第1のフレーム制御装置62は、スケーラブル・エンコードを行うために、動き補償されたBフレームの残差を上位レイヤのエンコーダ42のスケーラブル(FGS)エンコーダ54に経路を定める。スケーラブル(FGS)エンコードされたBフレームは、圧縮された上位レイヤのストリームの残りの部分を形成する。動き推定手段43により生成されたBフレームに関する動き情報はまた、第3のフレーム制御装置63を介して、上位レイヤのエンコーダ42の第2のマルチプレクサ64に適用される。第2のマルチプレクサ64は、Bフレームの動き情報と上位レイヤのフレームを組み合わせて、圧縮された上位レイヤのストリームを完成する。

10

【0022】

図6は、図4のエンコーダ40で生成された圧縮された基本レイヤと上位レイヤのストリームをデコードするために用いられ得る、本発明の例示的な実施例によるデコードのブロック図を示したものである。図に示す通り、デコーダ70は基本レイヤのデコーダ71と上位レイヤのデコーダ72を含む。基本レイヤのデコーダ71は、エンコードされた基本レイヤのストリームを受信し、前記ストリームを動き情報を含む第1のデータストリーム75aと、テクスチャ情報を含む第2のデータストリーム75bに逆多重化するデマルチプレクサ73を含む。上位レイヤのデコーダ72は、エンコードされた上位レイヤのストリームを受信し、前記ストリームを、テクスチャ情報を含む第3のデータストリーム74aと、動き情報を含む第4のデータストリーム74bに逆多重化するデマルチプレクサ92を含む。動き補償手段76は第4のデータストリーム74bの動き情報と、関連する基本レイヤのフレームメモリ77に保存された拡張基本レイヤの参照フレームを用いて、動き補償された拡張基本レイヤの参照(I及びP)フレームを再構成する。動き補償手段76は第1のデータストリーム75aのI及びP動き情報と、基本レイヤのフレームメモリ77に保存された従来の基本レイヤの参照フレームを用いて、従来の動き補償された基本レイヤの(I及びP)参照フレームを再構成する。動き補償された拡張基本レイヤの参照フレームと従来の動き補償された基本レイヤの参照フレームは、以下に説明される通り、第2のフレームのフロー制御装置93によって処理される。

20

【0023】

第2のデータストリーム75bのテクスチャ情報は、デコードするために基本レイヤの可変長デコーダ81に適用され、逆量子化するために逆量子化手段82に適用される。逆量子化係数は、逆離散コサイン変換デコーダ83に適用され、そこで逆量子化されたコードが第1の加算手段78の第1の入力80に適用される基本レイヤのフレームの残差に変換される。第1の加算手段78は、基本レイヤのPフレームの残差と、第2のフレームのフロー制御装置93によって第1の加算手段の第2の入力79に選択的に経路を定められたそれぞれの動き補償された基本レイヤの参照フレームとを合計し、動き予測されたPフレームを出力する。(基本レイヤのIフレームの残差は、第1の加算手段78によって基本レイヤのIフレームとして出力される。)第1の加算手段78によって出力されたI及びP基本レイヤフレームは、基本レイヤのフレームメモリ77に保存され、従来の基本レイヤの参照フレームを形成する。更に、第1の加算手段78によって出力されたI及びPフ

30

40

【0024】

上位レイヤのデコーダ72は、FGSビットプレーンのデコーダ84、又は圧縮された上位レイヤのストリームをデコードし、差分のI及びPフレームの残差とBフレームの残差を再構成する同様のスケーラブル・デコーダを含み、前記差分のI及びPフレームの残差とBフレームの残差は第2の加算手段90に適用される。I及びP差分のフレームの残差は、第1のフレームのフロー制御装置85によって差分のI及びPフレームの残差の1つ以上の再構成された上位レイヤのビットフレーム(又はその一部分)を受け取るマスキング装置86に選択的に経路を定められ、それを第3の加算手段87の第1の入力88に適用する。第3の加算手段87は、I及びPフレームの残差と、基本レイヤのデコーダ71に

50

よって第2の入力89に適用される対応する基本レイヤのI及びPフレームとを合計し、拡張基本レイヤのI及びP参照フレームを再構成し、それらはフレームメモリ77に保存される。

【0025】

動き補償された拡張基本レイヤのI及びP参照フレームは、第2のフレームのフロー制御装置83によって第2の加算手段90に選択的に経路を定められ、前記第2の加算手段は、動き補償された拡張基本レイヤのI及びP参照フレームと、対応するBフレームの残差とBフレームの動き情報（圧縮された上位レイヤのストリームにおいて送信される）とを合計し、上位レイヤのBフレームを再構成する。

【0026】

基本レイヤのデコーダ71の第1の加算手段78によって出力された基本レイヤのI及びPフレームは、第3のフレームのフロー制御装置91によって第2の加算手段90に選択的に経路を定められ、前記第2の加算手段は、上位レイヤのI及びPフレームとそれぞれの基本レイヤのI及びPフレームとを合計し、拡張I及びPフレームを生成する。拡張I及びPフレームと上位レイヤBは、拡張された映像として第2の加算手段90によって出力される。

【0027】

図3Bは本発明の第2の例示的な実施例によるスケーラブル映像符号化の仕組みを示したものである。第2の実施例のスケーラブル映像符号化の仕組み100は、フレーム内符号化Iフレームと、フレーム間符号化動き予測Pフレームと、フレーム間符号化動き双方向予測Bフレームを有する単一のループの予測に基づくスケーラブル・レイヤ132のみを有する。この実施例において、全てのフレーム（I及びP及びBフレーム）は、スケーラブル・コーデックで完全に符号化される。スケーラブル・コーデックは、DCTに基づくもの（FGS）や、ウェーブレット（wavelet）に基づくものや、何らかの他の組み込まれたコーデックである場合がある。P及びBフレームは、エンコードの間に、拡張基本レイヤI及びP又はP及びP参照フレームから完全に動き予測される。

【0028】

その技術に通常熟練した人は、基本レイヤの削除は、上位レイヤのP及びBフレームの双方の時間的な冗長を減少させるため、前記の符号化の仕組みを効率的にし、更に映像の画像品質を改善することがわかるだろう。

【0029】

図5は、図3Bのスケーラブル映像符号化の仕組みを作るために用いられ得る、本発明の例示的な実施例によるエンコーダ140のブロック図を示したものである。図に示す通り、図5のエンコーダ140は、動き補償及び推定ユニット141とスケーラブル・テクスチャ・エンコーダ142とを含む。動き補償及び推定ユニット141は、拡張基本レイヤのI及びP参照フレームを含むフレームメモリ60を有する。動き推定手段43は、元の映像シーケンスと、フレームメモリ60に保存された拡張基本レイヤのI及びP参照フレームから動き情報（動きベクトルと予測モード）を生成する。前記動き情報は、動き補償手段44とマルチプレクサ49に適用される。動き補償手段44は、動き情報とフレームメモリ60に保存された拡張基本レイヤのI及びP参照フレームを用いて、拡張基本レイヤのI及びP参照フレームRef(i)の動き補償されたバージョンを生成する。減算手段45は、拡張基本レイヤの参照フレームRef(i)の動き補償されたバージョンから元の映像シーケンスを差し引き、動き補償されたフレームの残差MCR(i)を生成する。

【0030】

スケーラブル・テクスチャ・エンコーダ142は、従来のFGSエンコーダ54又は同様のスケーラブル・エンコーダを含む。FGSエンコーダ54の場合、基本レイヤのエンコーダ41の減算手段45によって出力された動き補償されたフレームの残差が、DCTエンコードが行われ、ビットプレーンのDCTスキャンが行われ、エントロピーエンコードが行われ、圧縮された上位レイヤの（FGS符号化）フレームを生成する。マルチプレクサ49は、圧縮された上位レイヤのフレームと動き推定手段43によって生成された動き情報とを組

10

20

30

40

50

み合わせることにより、圧縮された出力ストリームを生成する。マスキング装置 55 は、上位レイヤの符号化 I 及び P フレームの 1 つ以上の符号化ビットプレーンを受け取り、それを加算手段 52 に適用する。加算手段 52 は、前記データと、対応する動き補償された上位レイヤの I 及び P 参照フレーム Ref(i) とを合計し、フレームメモリ 60 に保存される新しい拡張基本レイヤの I 及び P 参照フレームを生成する。

【0031】

本発明のスケラブル映像符号化の仕組みは、映像シーケンスの多様な部分又は多様な映像シーケンスについて、図 1 の現在の映像符号化の仕組みと交換する又は切り替えることができる。更に、図 3 A と、3 B と、図 1 の現在の映像符号化の仕組み、及び / 又は前述の関連する同時係属の米国特許出願において説明された映像符号化の仕組み、及び / 又は他の映像符号化の仕組みとの間で切り替えが実行され得る。前記の映像符号化の切り替えは、チャネル特性に基づいて行うことができ、エンコード時又は送信時に実行されることができる。更に、本発明の映像符号化の仕組みは、複雑性のわずかな増加のみ（図 3 A ）又は減少（図 3 B ）で、符号化効率における大幅な利益を達する。

10

【0032】

図 7 は、図 5 のエンコーダ 140 で生成された出力ストリームをデコードするために用いられ得る、本発明の例示的な実施例によるデコーダ 170 のブロック図を示したものである。図に示す通り、デコーダ 170 は、エンコードされたスケラブル・ストリームを受信し、前記ストリームを第 1 と第 2 のデータストリーム 174 と 175 に逆多重化するデマルチプレクサ 173 を含む。動き情報（動きベクトルと動き予測モード）を含む第 1 のデータストリーム 174 は、動き補償手段 176 に適用される。動き補償手段 176 は、前記動き情報と、基本レイヤのフレームメモリ 177 に保存された拡張基本レイヤの I 及び P 参照フレームとを用いて、動き補償された拡張基本レイヤの I 及び P 参照フレームを再構成する。

20

【0033】

デマルチプレクサ 173 によって逆多重化された第 2 のデータストリーム 175 は、テクスチャ・デコーダ 172 に適用され、前記テクスチャ・デコーダは、FGS のビットプレーンのデコーダ 184、又は第 2 のデータストリーム 175 をデコードする同様のスケラブル・デコーダを含み、第 1 の加算手段 190 に適用される I 及び P 及び B フレームの残差を再構成する。I 及び P フレームの残差はまた、I 及び P フレームの残差の 1 つ以上の符号化ビットプレーン（又はその一部分）を受け取り、それを第 2 の加算手段 187 の第 1 の入力 188 に適用するフレームのフロー制御装置 185 を介して、マスキング装置 186 に適用する。第 2 の加算手段 187 は、I 及び P フレームの残差データと、動き補償手段 176 によって第 2 の入力 189 に適用された、対応する再攻勢された動き補償された拡張基本レイヤの I 及び P フレームとを合計し、新しい拡張基本レイヤの I 及び P 参照フレームを再構成し、それらはフレームメモリ 177 に保存される。

30

【0034】

動き補償された拡張基本レイヤの I 及び P 参照フレームはまた、第 1 の加算手段 190 に経路を定められ、前記第 1 の加算手段は、それと、（FGS デコーダ 184 からの）対応する再構成されたフレームの残差とを合計し、拡張された I 及び P 及び B フレームを生成し、それらは拡張された映像として第 1 の加算手段 190 によって出力される。

40

【0035】

図 8 は、本発明の原理を実現するために用いられ得るシステム 200 の例示的な実施例を示したものである。システム 200 は、テレビや、セットトップボックスや、デスクトップ又はラップトップ又はパームトップのコンピュータや、個人情報端末（PDA）や、ビデオカセットレコーダ（VCR）のような映像 / 画像保存装置や、デジタルビデオレコーダ（DVR）や、TiVo 装置等や、それに加えてこれらや他の装置の一部又は組み合わせを表し得る。本システム 200 は、1 つ以上の映像 / 画像のソース 201 と、1 つ以上の入出力装置 202 と、プロセッサ 203 と、メモリ 204 とを含む。映像 / 画像のソース（群）201 は、例えばテレビ受信機又は VCR 又は他の映像 / 画像保存装置を表し得る。ソース（群）

50

）２０１は、例えば、インターネットや、広域ネットワークや、メトロポリタンエリアネットワークや、ローカルエリアネットワークや、地上波放送システムや、ケーブルネットワークや、衛星ネットワークや、無線ネットワークや、電話ネットワークや、それに加えてこれらや他の形式のネットワークの一部又は組み合わせのようなグローバルなコンピュータ通信ネットワーク上で、サーバ又はサーバ群から映像を受信するための１つ以上のネットワーク接続を択一的に表し得る。

【００３６】

入出力装置２０２と、プロセッサ２０３とメモリ２０４は、通信媒体２０５上で通信し得る。通信媒体２０５は、例えば、バスや、通信ネットワークや、回路又は回路カード又は他の装置の１つ以上の内部接続や、それに加えてこれらや他の通信媒体の一部及び組み合わせを表し得る。ソース（群）２１０からの入力映像データは、ディスプレイ装置２０６に供給される出力映像／画像を生成するために、メモリ２０４に保存された１つ以上のソフトウェアプログラムに従って処理され、プロセッサ２０３によって実行される。

10

【００３７】

好ましい実施例において、本発明の原理を使用する符号化とデコードは、システムによって実行されるコンピュータ読み取り可能なコードによって実現され得る。前記コードは、メモリ２０４に保存され得る、又はCD-ROMやフロッピー（Ｒ）ディスクのようなメモリ媒体から読み取られる／ダウンロードされ得る。他の実施例において、本発明を実現するために、ソフトウェアの命令の代わりに又はそれと組み合わせてハードウェアの回路構成が用いられ得る。例えば、図４－７に示される要素はまた、分離したハードウェア要素として実現され得る。

20

【００３８】

本発明は特定の実施例について前述したが、本発明はここで開示される実施例に制限又は限定されることを意図されるのではないことがわかる。例えば、DCTの他に、ウェーブレット（wavelet）又は他のマッチングパズル（matching-pursuits）を含むが、それに限定されない他の変換が用いられ得る。前記の及び他の全ての改良と変更が特許請求の範囲内であると考えられる。

【図面の簡単な説明】

【００３９】

【図１】現在の上位レイヤの映像符号化の仕組みを示したものである。

30

【図２】図１の映像符号化の仕組みの基本レイヤと上位レイヤを符号化するための従来のエンコーダのブロック図を示したものである。

【図３Ａ】本発明の第１の例示的な実施例によるスケーラブル映像符号化の仕組みを示したものである。

【図３Ｂ】本発明の第２の例示的な実施例によるスケーラブル映像符号化の仕組みを示したものである。

【図４】図３Ａのスケーラブル映像符号化の仕組みを作るために用いられ得る、本発明の例示的な実施例によるエンコーダのブロック図を示したものである。

【図５】図３Ｂのスケーラブル映像符号化の仕組みを作るために用いられ得る、本発明の例示的な実施例によるエンコーダのブロック図を示したものである。

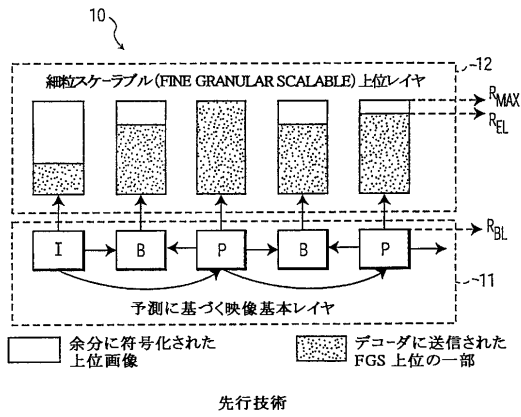
40

【図６】図４のエンコーダで生成された圧縮された基本レイヤと上位レイヤのストリームをデコードするために用いられ得る、本発明の例示的な実施例によるデコーダのブロック図を示したものである。

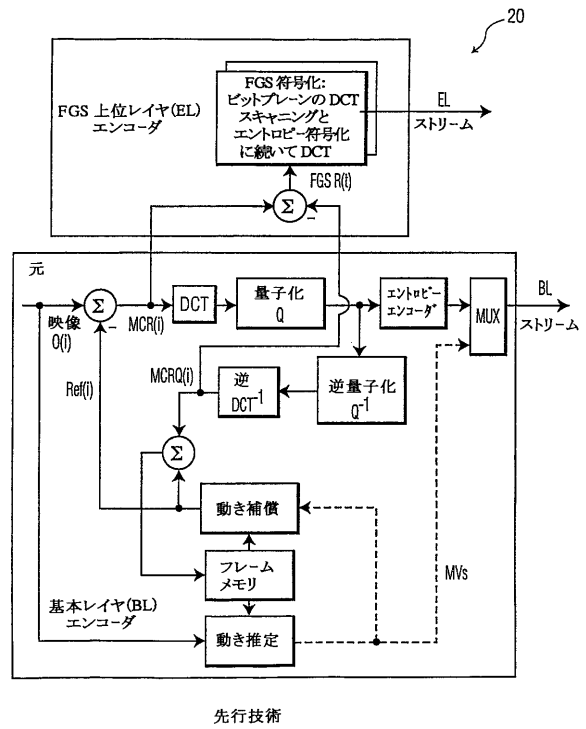
【図７】図５のエンコーダで生成された圧縮された基本レイヤと上位レイヤのストリームをデコードするために用いられ得る、本発明の例示的な実施例によるデコーダのブロック図を示したものである。

【図８】本発明の原理を実現するために用いられ得るシステムの例示的な実施例を示したものである。

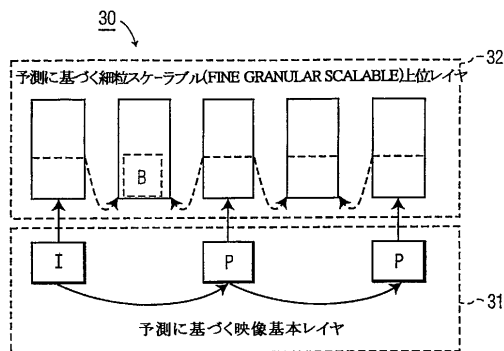
【図 1】



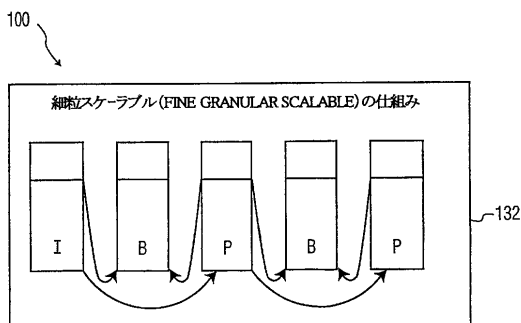
【図 2】



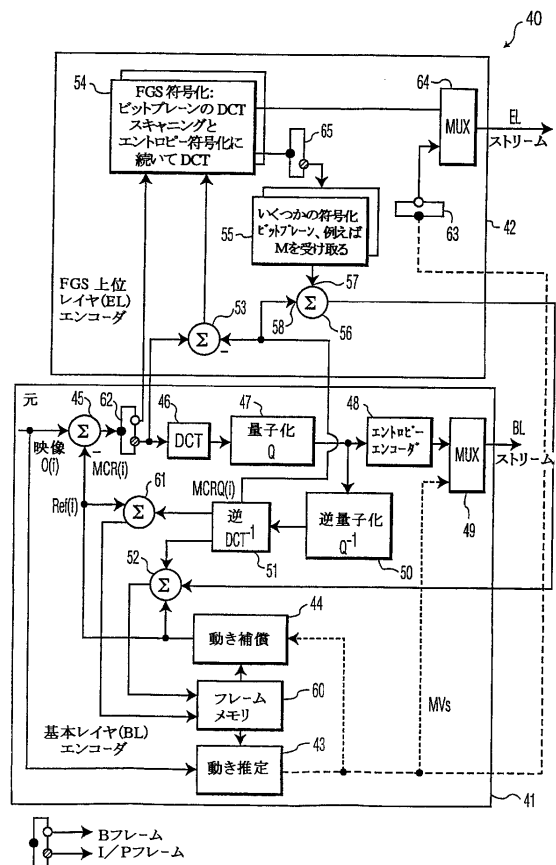
【図 3 A】



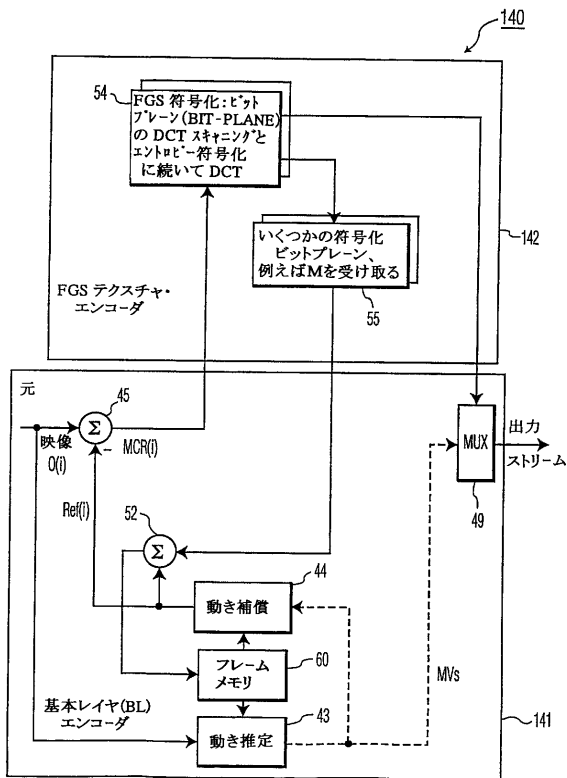
【図 3 B】



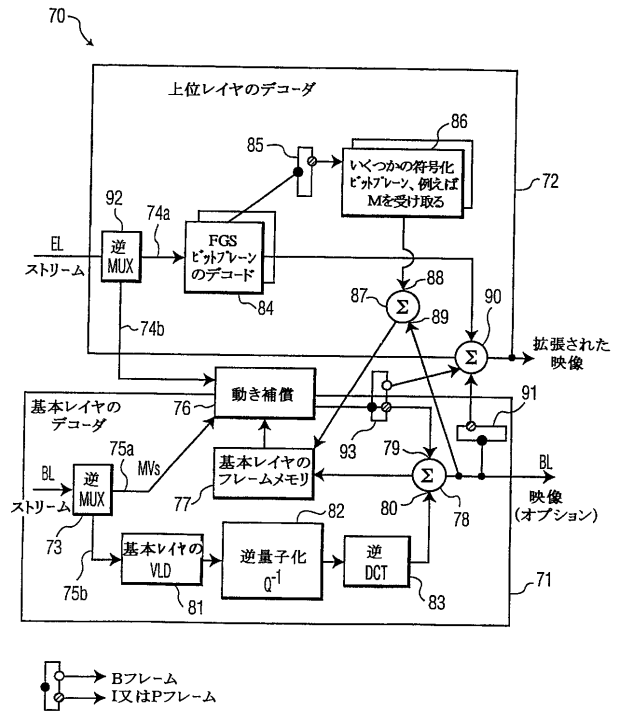
【図 4】



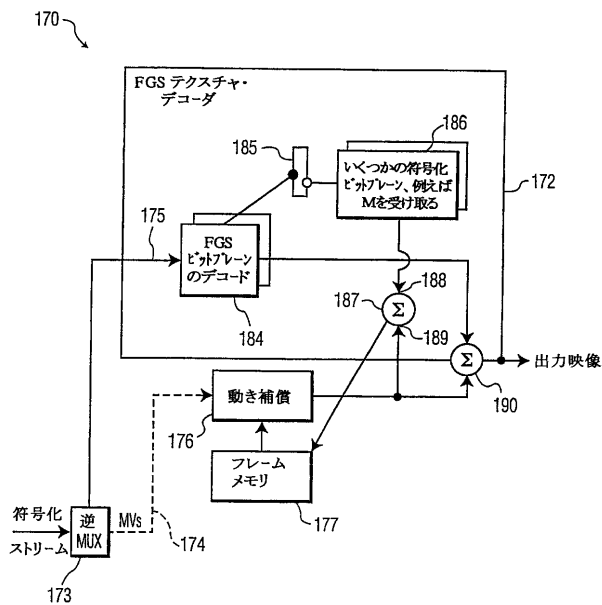
【図 5】



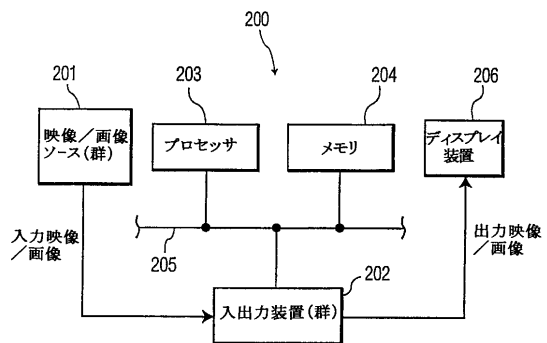
【図 6】



【図 7】



【図 8】



【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

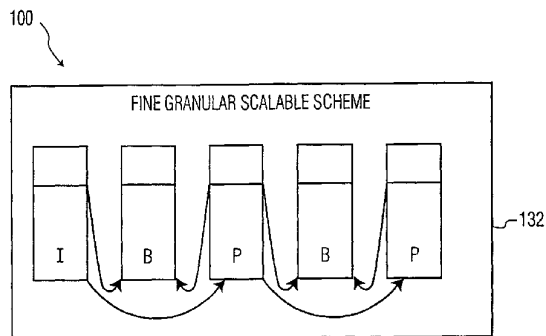
(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
27 February 2003 (27.02.2003)

PCT

(10) International Publication Number
WO 03/017672 A2

- (51) International Patent Classification: **H04N 7/26** (74) Agent: LANDOUSY, Christian; Internationaal Octrooibureau B.V., Prof. Holslaan 6, NL-5656 AA Eindhoven (NL).
- (21) International Application Number: PCT/IB02/02924
- (22) International Filing Date: 11 July 2002 (11.07.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 15 August 2001 (15.08.2001) US 09/930,672
- (71) Applicant: KONINKLIJKE PHILIPS ELECTRONICS N.V. (NL/NL); Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).
- (72) Inventor: VAN DER SCHAAR, Mihaela; Prof. Holslaan 6, NL-5656 AA Eindhoven (NL).
- (81) Designated States (national): CN, JP, KR.
- (84) Designated States (regional): European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, 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 "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: TOTALLY EMBEDDED FGS VIDEO CODING WITH MOTION COMPENSATION



(57) Abstract: A scalable video coding scheme having a single motion compensation loop that generates bi-directional predicted frames (B frames) or predicted frames and bi-directional predicted frames (P and B frames) coded entirely with a scalable codec.

WO 03/017672 A2

WO 03/017672

1

PCT/IB02/02924

Totally embedded FGS video coding with motion compensation

FIELD OF THE INVENTION

The present invention relates to video coding, and more particularly to a scalable video coding scheme that employs a single motion compensation loop for generating bi-directional predicted frames (B frames) or predicted frames and bi-directional predicted frames and (P and B frames) coded entirely with fine granular scalable (FGS) coding.

BACKGROUND OF THE INVENTION

Scalable enhancement layer video coding has been used for compressing video transmitted over computer networks having a varying bandwidth, such as the Internet. A current enhancement layer video coding scheme employing FGS coding techniques (adopted by the ISO MPEG-4 standard) is shown in FIG. 1. As can be seen, the video coding scheme includes a prediction-based base layer 11 coded at a bit rate R_{BL} , and an FGS enhancement layer 12 coded at R_{EL} .

The prediction-based base layer 11 includes intraframe coded I frames, interframe coded P frames which are temporally predicted from previous I or P frames using motion estimation-compensation, and interframe coded bi-directional B frames which are temporally predicted from both previous and succeeding frames adjacent the B frame using motion estimation-compensation. The use of predictive and/or interpolative coding i.e., motion estimation and corresponding compensation, in the base layer 11 reduces temporal redundancy therein, but only to a limited extent, since only base layer frames are used for prediction.

The enhancement layer 12 includes FGS enhancement layer I, P, and B frames derived by subtracting their respective reconstructed base layer frames from the respective original frames (this subtraction can also take place in the motion-compensated domain). Consequently, the FGS enhancement layer I, P and B frames in the enhancement layer are not motion-compensated. (The FGS residual is taken from frames at the same time-instance.) The primary reason for this is to provide flexibility which allows truncation of each FGS enhancement layer frame individually depending on the available bandwidth at transmission time. More specifically, the fine granular scalable coding of the enhancement layer 12

WO 03/017672

2

PCT/IB02/02924

permits an FGS video stream to be transmitted over any network session with an available bandwidth ranging from $R_{\min} = R_{BL}$ to $R_{\max} = R_{BL} + R_{EL}$. For example, if the available bandwidth between the transmitter and the receiver is $B=R$, then the transmitter sends the base layer frames at the rate R_{BL} and only a portion of the enhancement layer frames at the rate $R_{EL} = R - R_{BL}$. As can be seen from FIG. 1, portions of the FGS enhancement layer frames in the enhancement layer can be selected in a fine granular scalable manner for transmission. Therefore, the total transmitted bit-rate is $R = R_{BL} + R_{EL}$, because of its flexibility in supporting a wide range of transmission bandwidth with a single enhancement layer.

FIG. 2 shows a block-diagram of a conventional FGS encoder for coding the base layer 11 and enhancement layer 12 of the video coding scheme of FIG. 1. As can be seen, the enhancement layer residual of frame i (FGSR(i)) equals $MCR(i) - MCRQ(i)$, where $MCR(i)$ is the motion-compensated residual of frame i , and $MCRQ(i)$ is the motion-compensated residual of frame i after the quantization and the dequantization processes.

Although the current FGS enhancement layer video coding scheme 10 of FIG. 1 is very flexible, it has the disadvantage that its performance in terms of video image quality is relatively low compared with that of a non-scalable coder functioning at the same transmission bit-rate. The decrease in image quality is not due to the fine granular scalable coding of the enhancement layer 12 but mainly due to the reduced exploitation of the temporal redundancy among the FGS residual frames within the enhancement layer 12. In particular, the FGS enhancement layer frames of the enhancement layer 12 are derived only from the motion-compensated residual of their respective base layer I, P, and B frames, no FGS enhancement layer frames are used to predict other FGS enhancement layer frames in the enhancement layer 12 or other frames in the base layer 11.

Accordingly, a scalable video coding scheme having improved video image quality is needed.

SUMMARY OF THE INVENTION

The present invention is directed to a scalable video coding scheme that employs a single motion compensation loop for generating bi-directional predicted frames (B frames) or predicted frames and bi-directional predicted frames and (P and B frames) coded entirely with fine granular scalable (FGS) coding. One aspect of the invention involves a method of coding video comprising the steps of: encoding an uncoded video to generate extended base layer reference frames, each of the extended base layer reference frames

WO 03/017672

PCT/IB02/02924

3

including a base layer reference frame and at least a portion of an associated enhancement layer reference frame; and predicting frame residuals from the uncoded video and the extended base layer reference frames.

5 Another aspect of the invention involves a method of decoding a compressed video having a base layer stream and an enhancement layer stream, comprising the steps of: decoding the base layer and enhancement layer streams to generate extended base layer reference frames, each of the extended base layer reference frames including a base layer reference frame and at least a portion of an associated enhancement layer reference frame; and predicting frame residuals from the extended base layer reference frames.

10 Still another aspect of the invention involves a memory medium for coding video, comprising: code for encoding an uncoded video to generate extended base layer reference frames, each of the extended base layer reference frames including a base layer reference frame and at least a portion of an associated enhancement layer reference frame; and code for predicting frame residuals from the uncoded video and the extended base layer reference frames.

15 A further aspect of the invention involves a memory medium for decoding a compressed video having a base layer stream and an enhancement layer stream, comprising: code for decoding the base layer and enhancement layer streams to generate extended base layer reference frames, each of the extended base layer reference frames including a base layer reference frame and at least a portion of an associated enhancement layer reference frame; and code for predicting frame residuals from the extended base layer reference frames.

20 Still a further aspect of the invention involves an apparatus for coding video, which comprises: means for encoding an uncoded video to generate extended base layer reference frames, each of the extended base layer reference frames including a base layer reference frame and at least a portion of an associated enhancement layer reference frame; and means for predicting frame residuals from the uncoded video and the extended base layer reference frames.

25 Still another aspect of the invention involves an apparatus for decoding a compressed video having a base layer stream and an enhancement layer stream, which comprises: means for decoding the base layer and enhancement layer streams to generate extended base layer reference frames, each of the extended base layer reference frames including a base layer reference frame and at least a portion of an associated enhancement layer reference frame; and means for predicting frame residuals from the extended base layer reference frames.

30

WO 03/017672

4

PCT/IB02/02924

BRIEF DESCRIPTION OF THE DRAWINGS

The advantages, nature, and various additional features of the invention will appear more fully upon consideration of the illustrative embodiments now to be described in detail in connection with accompanying drawings where like reference numerals identify like elements throughout the drawings:

FIG. 1 shows a current enhancement layer video coding scheme;

FIG. 2 shows a block-diagram of a conventional encoder for coding the base layer and enhancement layer of the video coding scheme of FIG. 1;

FIG. 3A shows a scalable video coding scheme according to a first exemplary embodiment of the present invention;

FIG. 3B shows a scalable video coding scheme according to a second exemplary embodiment of the present invention;

FIG. 4 shows a block-diagram of an encoder, according to an exemplary embodiment of the present invention, that may be used for generating the scalable video coding scheme of FIG. 3A;

FIG. 5 shows a block-diagram of an encoder, according to an exemplary embodiment of the present invention, that may be used for generating the scalable video coding scheme of FIG. 3B;

FIG. 6 shows a block-diagram of a decoder, according to an exemplary embodiment of the present invention, that may be used for decoding the compressed base layer and enhancement layer streams generated by the encoder of FIG. 4;

FIG. 7 shows a block-diagram of a decoder, according to an exemplary embodiment of the present invention, that may be used for decoding the compressed base layer and enhancement layer streams generated by the encoder of FIG. 5; and

FIG. 8 shows an exemplary embodiment of a system which may be used for implementing the principles of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

FIG. 3A shows a scalable video coding scheme 30 according to a first exemplary embodiment of the present invention. The scalable video coding scheme 30 includes a prediction-based base layer 31 and a single-loop prediction-based enhancement layer 32.

WO 03/017672

PCT/IB02/02924

5

The prediction-based base layer 31 is coded to include intraframe coded I frames and interframe coded P frames, which are generated conventionally during base layer (non-scalable) coding from standard base layer I and P reference frames. No interframe coded bi-directional B frames are coded in the base layer.

5 In accordance with the principles of the present invention, the prediction-based enhancement layer 32 is coded to include interframe coded bi-directional B frames, which are motion-predicted from "extended" or "enhanced" base layer I and P or P and P reference frames (hereinafter extended base layer I and P reference frames) during base layer coding. Each extended base layer reference frame comprises a standard base layer reference frame,
10 and at least a portion of an associated enhancement layer reference frame (one or more bitplanes or fractional bit-planes of the associated enhancement layer reference frame can be used).

The enhancement layer 32 is also coded to include enhancement layer I and P frames that are generated conventionally by subtracting their respective reconstructed
15 (decoded) base layer frame residuals from their respective original base layer frame residuals. The enhancement layer I, B, and P frames may be coded with any suitable a scalable codec. For example, the scalable codec may be a DCT-based codec (FGS), a wavelet-based codec, or any other embedded codec. In the embodiment shown in FIG. 3A, the scalable codec comprises FGS.

20 As one of ordinary skill in the art will appreciate, the video coding scheme 30 of the present invention improves the image quality of the video. This is because the video coding scheme 30 uses extended base layer reference frames to reduce temporal redundancy in the enhancement layer B frames.

FIG. 4 shows a block-diagram of an encoder 40, according to an exemplary
25 embodiment of the present invention, that may be used for generating the scalable video coding scheme of FIG. 3A. As can be seen, the encoder 40 includes a base layer encoder 41 and an enhancement layer encoder 42. The base layer encoder 41 includes a motion estimator 43 that generates motion information (motion vectors and prediction modes) from the original video sequence and base layer and extended base layer reference frames stored in
30 frame memory 60. This motion information is then applied to a motion compensator 44 that generates conventional motion-compensated base layer reference frames and motion-compensated versions of the extended base layer I and P reference frames of the present invention (all denoted Ref(i)) using the motion information and conventional reference frames and the extended base layer I and P reference frames stored in the frame memory 60.

WO 03/017672

6

PCT/IB02/02924

A first subtractor 45 subtracts the conventional motion-compensated reference frames from the original video sequence to generate motion-compensated residuals of the base layer I and P frames. A first frame flow control device 62 routes just the motion-compensated residuals of the base layer I and P frames MCR(i) for processing by a discrete cosine transform (DCT) encoder 46, a quantizer 47, and an entropy encoder 48 to generate the base layer I and P frames, which form a portion of a compressed base layer stream. The motion information generated by the motion estimator 43 is also applied to a multiplexer 49, which combines the motion information with the base layer I and P frames to complete the compressed base layer stream. The quantized motion-compensated residuals of the base layer I and P frames MCR(i) generated at the output of the quantizer 47 are dequantized by an inverse quantizer 50, and then decoded by an inverse DCT decoder 51. This process generates quantized/dequantized versions of the motion-compensated residuals of the base layer I and P frames MCRQ(i) at the output of the inverse DCT 51. The quantized/dequantized motion-compensated residuals of the base layer I and P frames at the output of the inverse DCT 51 are applied to a first adder 61, which sums them with corresponding motion-compensated base layer reference frames Ref(i), hence generating the conventional base layer reference frames that are stored in the frame memory 60 as described above.

The quantized/dequantized motion-compensated residuals of the base layer I and P frames are also applied to a second subtractor 53 in the enhancement layer encoder 42. The second subtractor 53 subtracts the quantized/dequantized motion-compensated residuals of the base layer I and P frames from corresponding motion-compensated residuals of the base layer I and P frames to generate differential I and P frame residuals. The output of the second subtractor 53 is scalable coded by an FGS encoder 54 or like scalable encoder. The FGS encoder 54 uses conventional DCT encoding followed by conventional bit-plane DCT scanning and conventional entropy encoding to generate scalable (FGS) encoded I and P frames, which form a portion of a compressed enhancement layer stream. A masking device 55 takes one or more of the coded bit planes of the scalable encoded I and P frames, selectively routed through a third frame flow control device 65, and applies this data to a first input 57 of a second adder 56. The quantized/dequantized versions of the motion-compensated residuals of the I and P frames MCRQ(i) generated by the base layer encoder 41 are further applied to a second input 58 of the second adder 56. The second adder 56 generates enhancement layer I and P reference frames by summing the one or more coded bit planes of the enhancement layer encoded I and P frames with respective I and P frame residuals MCRQ(i). The enhancement layer I and P reference frames computed by the second

WO 03/017672

PCT/IB02/02924

7

adder 56 are applied to a third adder 52 in the base layer encoder 41. The third adder 52 sums the enhancement layer I and P reference frames with corresponding motion-compensated base layer I and P reference frames Ref(i) and corresponding quantized/dequantized motion-compensated base layer I and P frame residuals to generate the extended base layer I and P reference frames, which are stored in the frame memory 60.

The motion compensator 44 generates motion-compensated versions of the extended base layer I and P reference frames using the motion information and the extended base layer I and P reference frames stored in the frame memory 60. The first subtractor 45 subtracts the motion-compensated extended base layer reference frames from the original video sequence to generate motion-compensated B frame residuals. The first frame control device 62 routes the motion-compensated B frame residuals to the scalable (FGS) encoder 54 of the enhancement layer encoder 42, for scalable encoding. The scalable (FGS) encoded B frames form the remaining portion of the compressed enhancement layer stream. The motion information pertaining to the B frames generated by the motion estimator 43 is also applied to a second multiplexer 64 in the enhancement layer encoder 42, via a third frame control device 63. The second multiplexer 64 combines the B frame motion information with the enhancement layer frames to complete the compressed enhancement layer stream.

FIG. 6 shows a block-diagram of a decoder 70, according to an exemplary embodiment of the present invention, that may be used for decoding the compressed base layer and enhancement layer streams generated by the encoder 40 of FIG. 4. As can be seen, the decoder 70 includes a base layer decoder 71 and an enhancement layer decoder 72. The base layer decoder 71 includes a demultiplexer 73 which receives the encoded base layer stream and demultiplexes the stream into a first data stream 75a that contains motion information, and a second data stream 75b that contains texture information. The enhancement layer decoder 72 includes a demultiplexer 92 which receives the encoded enhancement layer stream and demultiplexes this stream into a third data stream 74a that contains texture information, and a fourth data stream 74b that contains motion information. A motion compensator 76 uses the motion information in the fourth data stream 74b and extended base layer reference frames stored in an associated base layer frame memory 77 to reconstruct the motion-compensated extended base layer reference (I and P) frames. The motion compensator 76 uses the I and P motion information in the first data stream 75a and conventional base layer reference frames stored in the base layer frame memory 77 to reconstruct the conventional motion-compensated base layer (I and P) reference frames. The motion-compensated extended base layer reference frames and the conventional motion-

WO 03/017672

8

PCT/IB02/02924

compensated base layer reference frames are then processed by a second frame flow control device 93 as will be explained further on.

The texture information in the second data stream 75b is applied to a base layer variable length code decoder 81 for decoding, and to an inverse quantizer 82 for
5 dequantizing. The dequantized coefficients are applied to an inverse discrete cosine transform decoder 83 where the dequantized code is transformed into the base layer frame residuals which are applied to a first input 80 of a first adder 78. The first adder 78 sums the base layer P frame residuals with their respective motion compensated base layer reference frames selectively routed by the second frame flow control device 93 to a second input 79 of the first
10 adder, and outputs the motion-predicted P frames. (The base layer I frame residuals are outputted by the first adder 78 as base layer I frames.) The I and P base layer frames outputted by the first adder 78 are stored in the base layer frame memory 77 and form the conventional base-layer reference frames. Additionally, the I and P frames outputted by the first adder 78 may be optionally outputted as a base layer video.

15 The enhancement layer decoder 72 includes an FGS bit-plane decoder 84 or like scalable decoder that decodes the compressed enhancement layer stream to reconstruct the differential I and P frame residuals and B frame residuals, which are applied to a second adder 90. The I and P differential frame residuals are also selectively routed by a first frame flow control device 85 to a masking device 86 that takes one or more of the reconstructed
20 enhancement-layer bit-planes (or fractions thereof) of the differential I and P frame residuals and applies them to a first input 88 of a third adder 87. The third adder 87 sums the I and P frame residuals with corresponding base layer I and P frames applied at a second input 89 thereof by the base layer decoder 71 to reconstruct the extended base layer I and P reference frames, which are stored in the frame memory 77.

25 The motion-compensated extended base layer I and P reference frames are selectively routed by the second frame flow control device 93 to the second adder 90, which sums the motion-compensated extended base layer I and P reference frames with corresponding B frame residuals and B frame motion information (transmitted in the compressed enhancement layer stream) to reconstruct the enhancement layer B frames.

30 The base layer I and P frames outputted by the first adder 78 of the base layer decoder 71 are selectively routed by a third frame flow control device 91 to the second adder 90, which sums the enhancement layer I and P frames with respective base layer I and P frames to generate enhanced I and P frames. The enhanced I and P frame and the enhancement layer B are outputted by the second adder 90 as an enhanced video.

WO 03/017672

PCT/IB02/02924

9

FIG. 3B shows a scalable video coding scheme 100 according to a second exemplary embodiment of the present invention. The scalable video coding scheme 100 of the second embodiment only includes a single-loop prediction-based scalable layer 132 having intraframe coded I frames; interframe-coded, motion-predicted P frames; and

5 interframe-coded, motion-bidirectional-predicted B frames. In this embodiment, all the frames (I, P, and B frames) are coded entirely with a scalable codec. The scalable codec can be DCT-based (FGS), wavelet-based, or any other embedded codec. The P and B frames are motion-predicted entirely from extended base layer I and P or P and P reference frames during encoding.

10 As one of ordinary skill in the art will appreciate, the elimination of a base layer makes this coding scheme very efficient and further improves the video image quality because it reduces temporal redundancy in both the enhancement layer P and B frames.

FIG. 5 shows a block-diagram of an encoder 140, according to an exemplary embodiment of the present invention, that may be used for generating the scalable video coding scheme of FIG. 3B. As can be seen, the encoder 140 of FIG. 5 includes a motion-

15 compensation and estimation unit 141 and a scalable texture encoder 142. The motion-compensation and estimation unit 141 includes a frame memory 60 which contains the extended base layer I and P reference frames. A motion estimator 43 generates motion information (motion vectors and prediction modes) from the original video sequence and the

20 extended base layer I and P reference frames stored in frame memory 60. This motion information is then applied to a motion compensator 44 and a multiplexer 49. The motion compensator 44 generates motion-compensated versions of the extended base layer I and P reference frames Ref(i) using the motion information and the extended base layer I and P reference frames stored in the frame memory 60. A subtractor 45 subtracts the motion-

25 compensated versions of the extended base layer reference frames Ref(i) from the original video sequence to generate motion-compensated frame residuals MCR(i).

The scalable texture encoder 142 includes a conventional FGS encoder 54 or like scalable encoder. In the case of the FGS encoder 54, the motion-compensated frame residuals outputted by the subtractor 45 of the base layer encoder 41 are DCT encoded, bit-

30 plane DCT scanned, and entropy encoded to generate compressed enhancement layer (FGS coded) frames. The multiplexer 49 generates a compressed output stream by combining the compressed enhancement layer frames with the motion information generated by the motion estimator 43. A masking device 55 takes one or more of the coded bit planes of the enhancement layer coded I and P frames and applies them to an adder 52. The adder 52 sums

WO 03/017672

10

PCT/IB02/02924

this data with the corresponding motion-compensated extended base layer I and P reference frames Ref(i) to generate new extended base layer I and P reference frames that are stored in the frame memory 60.

The scalable video coding schemes of the present invention can be alternated or switched with the current video coding scheme of FIG. 1 for the various portions of a video sequence or for various video sequences. Additionally, switching can be performed among the scalable video coding schemes of FIGS. 3A, 3B and the current video coding scheme of FIG. 1, and/or the video coding schemes described in the earlier-mentioned related copending U.S. Patent Applications and/or other video coding schemes. Such switching of video coding schemes can be done based on channel characteristics and can be performed at encoding or at transmission time. Further the video coding schemes of the present invention achieve a large gain in coding efficiency with only a slight increase (FIG. 3A), or decrease (FIG. 3B) in complexity.

FIG. 7 shows a block-diagram of a decoder 170, according to an exemplary embodiment of the present invention, that may be used for decoding the output stream generated by the encoder 140 of FIG. 5. As can be seen, the decoder 170 includes a demultiplexer 173 which receives the encoded scalable stream and demultiplexes the stream into first and second data streams 174 and 175. The first data stream 174, which includes motion information (motion vectors and motion prediction modes), is applied to a motion compensator 176. The motion compensator 176 uses this motion information and extended base layer I and P reference frames stored in base layer frame memory 177 to reconstruct the motion-compensated extended base layer I and P reference frames.

The second data stream 175 demultiplexed by the demultiplexer 173 is applied to a texture decoder 172, which includes an FGS bit-plane decoder 184 or like scalable decoder that decodes the second data stream 175 to reconstruct the I, P, and B frame residuals, which are applied to a first adder 190. The I and P frame residuals are also applied to a masking device 186 via a frame flow control device 185 that takes one or more of the coded bit-planes (or fractions thereof) of the I and P frame residuals and applies them to a first input 188 of a second adder 187. The second adder 187 sums the I and P frame residual data with corresponding reconstructed motion-compensated extended base layer I and P frames applied at a second input 189 thereof by the motion compensator 176 to reconstruct new extended base layer I and P reference frames, which are stored in the frame memory 177.

WO 03/017672

11

PCT/IB02/02924

The motion-compensated extended base layer I and P reference frames are also routed to the first adder 190, which sums them with corresponding reconstructed frame residuals (from the FGS decoder 184) to generate enhanced I, P and B frames, which are outputted by the first adder 190 as an enhanced video.

5 FIG. 8 shows an exemplary embodiment of a system 200 which may be used for implementing the principles of the present invention. The system 200 may represent a television, a set-top box, a desktop, laptop or palmtop computer, a personal digital assistant (PDA), a video/image storage device such as a video cassette recorder (VCR), a digital video recorder (DVR), a TiVO device, etc., as well as portions or combinations of these and other
10 devices. The system 200 includes one or more video/image sources 201, one or more input/output devices 202, a processor 203 and a memory 204. The video/image source(s) 201 may represent, e.g., a television receiver, a VCR or other video/image storage device. The source(s) 201 may alternatively represent one or more network connections for receiving
15 video from a server or servers over, e.g., a global computer communications network such as the Internet, a wide area network, a metropolitan area network, a local area network, a terrestrial broadcast system, a cable network, a satellite network, a wireless network, or a telephone network, as well as portions or combinations of these and other types of networks.

 The input/output devices 202, processor 203 and memory 204 may communicate over a communication medium 205. The communication medium 205 may
20 represent, e.g., a bus, a communication network, one or more internal connections of a circuit, circuit card or other device, as well as portions and combinations of these and other communication media. Input video data from the source(s) 201 is processed in accordance with one or more software programs stored in memory 204 and executed by processor 203 in order to generate output video/images supplied to a display device 206.

25 In a preferred embodiment, the coding and decoding employing the principles of the present invention may be implemented by computer readable code executed by the system. The code may be stored in the memory 204 or read/downloaded from a memory medium such as a CD-ROM or floppy disk. In other embodiments, hardware circuitry may be used in place of, or in combination with, software instructions to implement the invention.
30 For example, the elements shown in FIGS. 4-7 may also be implemented as discrete hardware elements.

 While the present invention has been described above in terms of specific embodiments, it is to be understood that the invention is not intended to be confined or limited to the embodiments disclosed herein. For example, other transforms besides DCT can

WO 03/017672

12

PCT/IB02/02924

be employed, including but not limited to wavelets or matching-pursuits. These and all other such modifications and changes are considered to be within the scope of the appended claims.

WO 03/017672

13

PCT/IB02/02924

CLAIMS:

1. A method of coding video, comprising the steps of:
encoding (41, 141, 42, 142) an uncoded video to generate extended base layer
reference frames, each of the extended base layer reference frames including a base layer
reference frame and at least a portion of an associated enhancement layer reference frame;
5 and
generating frame residuals (45) from the uncoded video and the extended base
layer reference frames.
2. A method of coding video according to claim 1, further comprising the step of
10 coding (54) the frame residuals with a scalable codec selected from the group consisting of
DCT based codecs or wavelet based codecs to generate enhancement layer frames.
3. A method of coding video according to claim 1, further comprising the step of
coding (54) the frame residuals with a fine granular scalable codec to generate fine granular
15 scalable enhancement layer frames.
4. A method of coding video according to claim 1, wherein the frame residuals
include B frame residuals.
- 20 5. A method of coding video according to claim 4, wherein the frame residual
further include P frame residuals.
6. A method of coding video according to claim 1, wherein the frame residual
include P frame residuals.
- 25 7. A method of decoding a compressed video having a base layer stream and an
enhancement layer stream, the method comprising the steps of:
decoding (71, 72, 172) the base layer and enhancement layer streams to
generate extended base layer reference frames, each of the extended base layer reference

WO 03/017672

14

PCT/IB02/02924

frames including a base layer reference frame and at least a portion of an associated enhancement layer reference frame; and
predicting (78) frame residuals from the extended base layer reference frames.

5 8. A method of decoding video according to claim 7, further comprising the step
of decoding the frame residuals with scalable decoding (84) selected from the group
consisting of DCT based decoding or wavelet based decoding.

9. A method of decoding video according to claim 8, further comprising the steps
10 of:
generating enhancement layer frames from the frame residuals; and
generating (90) an enhanced video from the base layer frames and the
enhancement layer frames.

15 10. A method of decoding video according to claim 7, wherein the frame residuals
include B frame residuals.

11. A method of decoding video according to claim 10, wherein the frame
residuals further include P frame residuals.

20 12. A method of decoding video according to claim 7, wherein the frame residuals
include P-frame residuals.

13. A memory medium for coding video, the memory medium comprising:
25 code (41, 141, 42, 142) for encoding an uncoded video to generate extended
base layer reference frames, each of the extended base layer reference frames including a
base layer reference frame and at least a portion of an associated enhancement layer reference
frame; and
code (45) for predicting frame residuals from the uncoded video and the
30 extended base layer reference frames.

14. A memory medium for coding video according to claim 13, further comprising
code (54) for scalable encoding the frame residuals.

WO 03/017672

PCT/IB02/02924

15

15. A memory medium for coding video according to claim 13, further comprising code (54) for fine granular scalable encoding the frame residuals.

16. A memory medium for coding video according to claim 13, wherein the frame residuals include B frame residuals.

17. A memory medium for coding video according to claim 16, wherein the frame residuals further include P frame residuals.

18. A memory medium for coding video according to claim 13, wherein the frame residuals include P frame residuals.

19. A memory medium for decoding a compressed video having a base layer stream and an enhancement layer stream, the memory medium comprising:

code (71, 72, 172) for decoding the base layer and enhancement layer streams to generate extended base layer reference frames, each of the extended base layer reference frames including a base layer reference frame and at least a portion of an associated enhancement layer reference frame; and

code (78) for predicting frame residuals from the extended base layer reference frames.

20. A memory medium for decoding a compressed video according to claim 19, further comprising code (84) for scalable decoding the frame residuals, the code for scalable decoding selected from the group consisting of DCT based code or wavelet based code.

21. A memory medium for decoding a compressed video according to claim 20, further comprising:

code for generating enhancement layer frames from the frame residuals; and

code for generating (90) an enhanced video from the base layer frames and the enhancement layer frames.

22. A memory medium for decoding a compressed video according to claim 19, wherein the frame residuals include B frame residuals.

WO 03/017672

16

PCT/IB02/02924

23. A memory medium for decoding a compressed video according to claim 22,
wherein the frame residuals further include P frame residuals.
24. A memory medium for decoding a compressed video according to claim 19,
5 wherein the frame residuals include P frame residuals.
25. An apparatus (40, 140) for coding video, the apparatus comprising:
means (41, 141, 42, 142) for encoding an uncoded video to generate extended
base layer reference frames, each of the extended base layer reference frames including a
10 base layer reference frame and at least a portion of an associated enhancement layer reference
frame; and
means (45) for predicting frame residuals from the uncoded video and the
extended base layer reference frames.
- 15 26. An apparatus for coding video according to claim 25, further comprising
means (54) for scalable encoding the frame residuals.
27. An apparatus for coding video according to claim 25, further comprising code
(54) for fine granular scalable encoding the frame residuals.
- 20 28. An apparatus for coding video according to claim 25, wherein the frame
residuals include B frame residuals.
29. An apparatus for coding video according to claim 28, wherein the frame
25 residuals further include P frame residuals.
30. An apparatus for coding video according to claim 25, wherein the frame
residuals include P frame residuals.
- 30 31. An apparatus (70, 170) for decoding a compressed video having a base layer
stream and an enhancement layer stream, the apparatus comprising:
means (71, 72, 172) for decoding the base layer and enhancement layer
streams to generate extended base layer reference frames, each of the extended base layer

WO 03/017672

17

PCT/IB02/02924

reference frames including a base layer reference frame and at least a portion of an associated enhancement layer reference frame; and

means (78) for predicting frame residuals from the extended base layer reference frames.

5

32. An apparatus for decoding a compressed video according to claim 31, further comprising scalable decoding means (84) for decoding the frame residuals, the scalable decoding means selected from the group consisting of DCT based decoding means or wavelet based decoding means.

10

33. An apparatus for decoding a compressed video according to claim 32, further comprising:

means for generating enhancement layer frames from the frame residuals; and

means for generating (90) an enhanced video from the base layer frames and

15 the enhancement layer frames.

34. An apparatus for decoding a compressed video according to claim 31, wherein the frame residuals include B frame residuals.

20 35. An apparatus for decoding a compressed video according to claim 34, wherein the frame residuals further include P frame residuals.

36. An apparatus for decoding a compressed video according to claim 31, wherein the frame residuals include P frame residuals.

WO 03/017672

PCT/IB02/02924

1/8

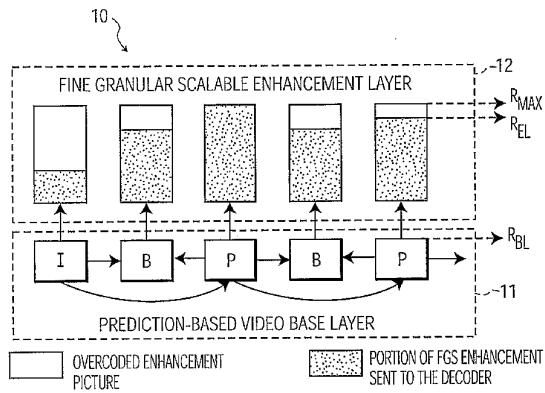


FIG. 1
(PRIOR ART)

WO 03/017672

PCT/IB02/02924

2/8

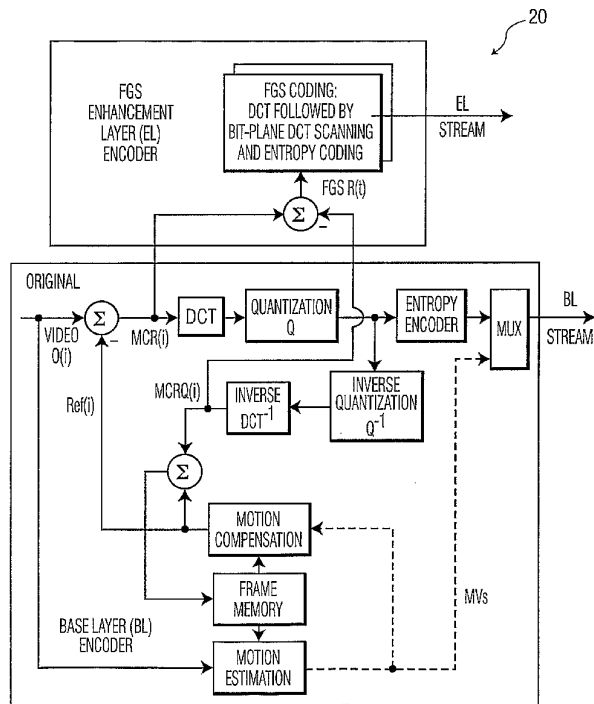


FIG. 2
(PRIOR ART)

WO 03/017672

PCT/IB02/02924

3/8

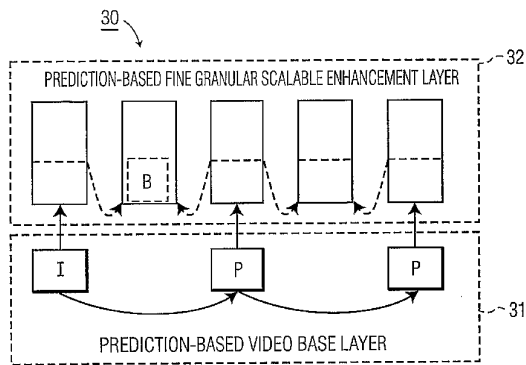


FIG. 3A

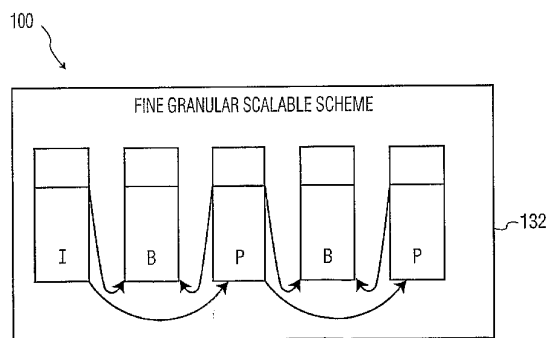
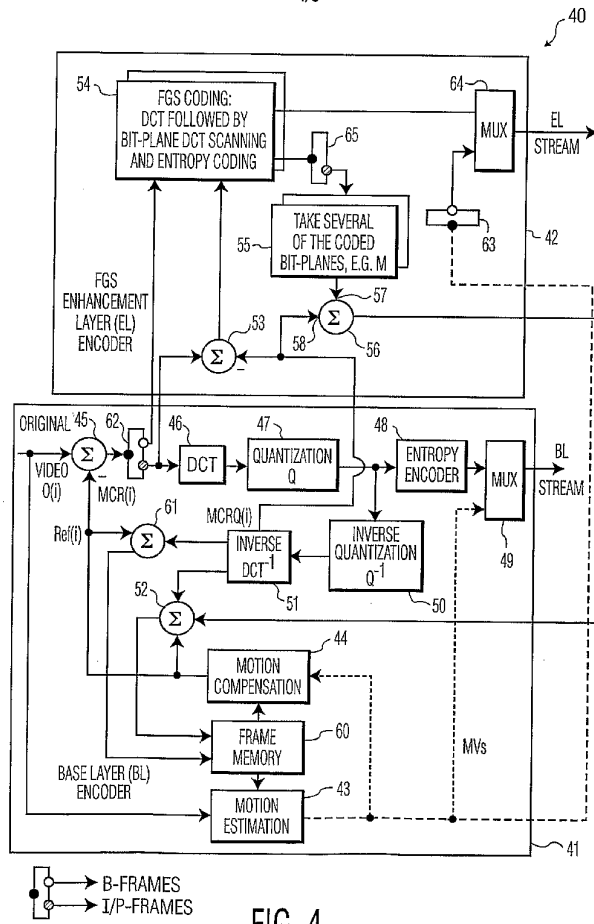


FIG. 3B

WO 03/017672

PCT/IB02/02924

4/8



WO 03/017672

PCT/IB02/02924

5/8

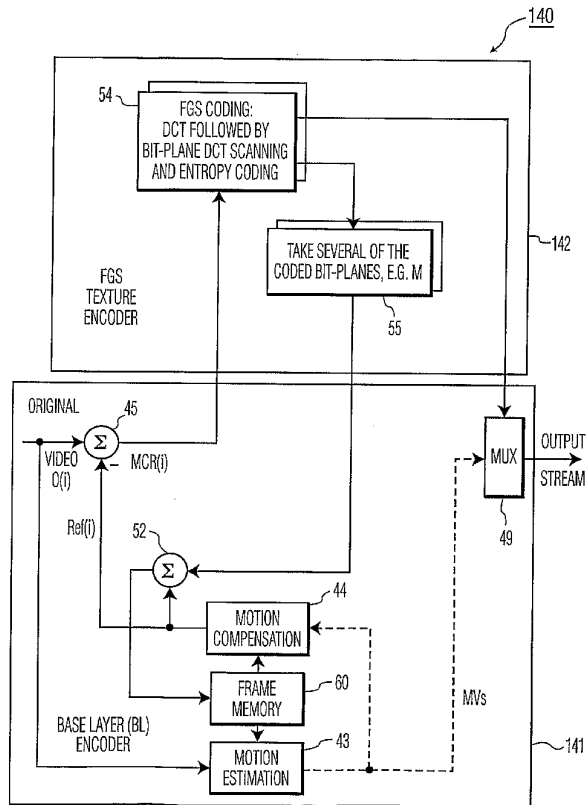


FIG. 5

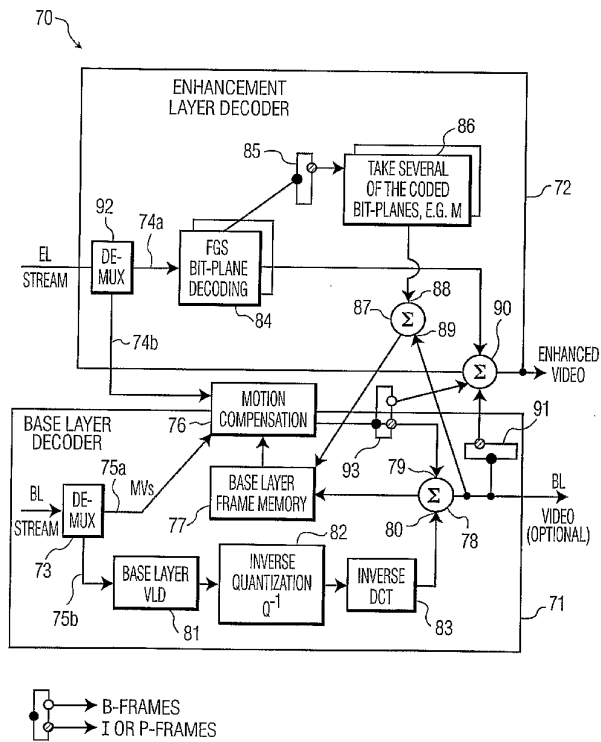
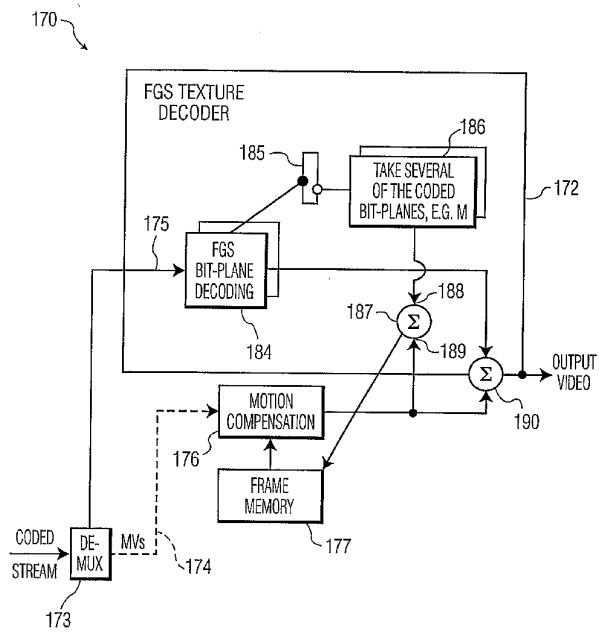


FIG. 6

WO 03/017672

PCT/IB02/02924

7/8



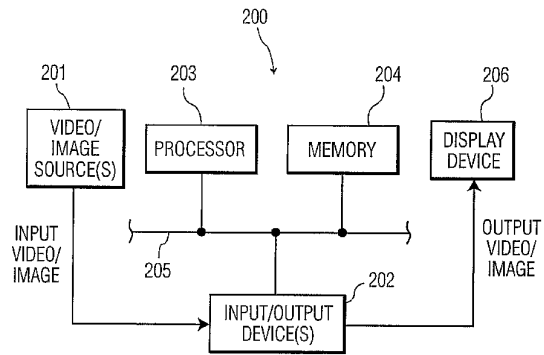


FIG. 8

【国際公開パンフレット（コレクトバージョン）】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau(43) International Publication Date
27 February 2003 (27.02.2003)

PCT

(10) International Publication Number
WO 2003/017672 A3(51) International Patent Classification⁷: H04N 7/50,
7/26, 7/46(74) Agent: LANDOUSY, Christian; Philips Intellectual
Property & Standards, Prof. Holstlaan 6, NL-5656 AA
Eindhoven (NL).(21) International Application Number:
PCT/IB2002/002924

(81) Designated States (national): CN, JP, KR.

(22) International Filing Date: 11 July 2002 (11.07.2002)

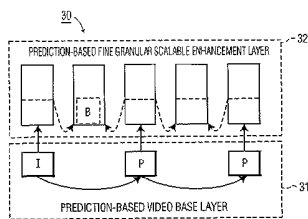
(84) Designated States (regional): European patent (AT, BE,
BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT,
LU, MC, NL, PT, SE, SK, TR).

(25) Filing Language: English

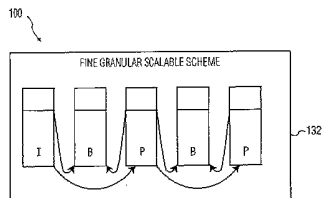
(26) Publication Language: English

Published:
— with international search report(30) Priority Data:
09/930,672 15 August 2001 (15.08.2001) US(71) Applicant: KONINKLIJKE PHILIPS ELECTRON-
ICS N.V. [NL/NL]; Groenewoudseweg 1, NL-5621 BA
Eindhoven (NL).(88) Date of publication of the international search report:
22 April 2004(72) Inventor: VAN DER SCHAAR, Mihaela; Prof. Holst-
laan 6, NL-5656 AA Eindhoven (NL).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: TOTALLY EMBEDDED FGS VIDEO CODING WITH MOTION COMPENSATION



A

(57) Abstract: A scalable video coding scheme
having a single motion compensation loop that
generates bi-directional predicted frames (B
frames) or predicted frames and bi-directional
predicted frames and (P and B frames) coded
entirely with a scalable codec.

B

WO 2003/017672 A3

【国際調査報告】

INTERNATIONAL SEARCH REPORT		Internat. Application No. PCT/ID 02/02924
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 H04N7/50 H04N7/26 H04N7/46		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 H04N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, INSPEC		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 4 723 161 A (KOGA TOSHIO) 2 February 1988 (1988-02-02) column 5, line 25 - line 40 figures 3-6	1-36
X	EP 0 595 403 A (PHILIPS ELECTRONIQUE LAB ; KONINKL PHILIPS ELECTRONICS NV (NL)) 4 May 1994 (1994-05-04) column 2, line 1 - line 23 column 3, line 3 - column 6, line 22 figures 1,2 -/-	1-36
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents: *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art *&* document member of the same patent family		
Date of the actual completion of the international search 28 March 2003		Date of mailing of the international search report 21/05/2003
Name and mailing address of the ISA European Patent Office, P.B. 5618 Patentlaan 2 NL - 2200 HV Rijswijk Tel: (+31-70) 340-2040, Tx 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer Lombardi, G

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT		Internal Application No PCT/JP 02/02924
C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JOHNSON A W ET AL: "FREQUENCY SCALABLE VIDEO CODING USING THE MDCT" PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING (ICASSP). I. IMAGE AND MULTIDIMENSIONAL SIGNAL PROCESSING. ADELAIDE, APR. 19 - 22, 1994, NEW YORK, IEEE, US, vol. 5 CONF. 19, 19 April 1994 (1994-04-19), pages V-477-V-480, XP000533761 ISBN: 0-7803-1776-9 figure 4 -----	1-36
P,X	WO 02 32142 A (KONINKL PHILIPS ELECTRONICS NV) 18 April 2002 (2002-04-18) the whole document -----	1-36
A	LI S ET AL: "EXPERIMENTAL RESULTS WITH PROGRESSIVE FINE GRANULARITY SCALABLE (PFGS) CODING" ISO/IEC JTC1/SC29/WG11 MPEG99/M5742, March 2000 (2000-03), page COMPLETE XP001112953 page 5, paragraph TYPOGRAPHIC.LAST -page 7, paragraph TYPOGRAPHIC.LAST figures 5,6 -----	1-36
A	WO 01 39503 A (KONINKL PHILIPS ELECTRONICS NV) 31 May 2001 (2001-05-31) page 10, line 12 - line 23 figures 7,8A -----	1-36

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT				Internat PCT/ _{ID}	Application No 02/02924
Patent document cited in search report	Publication date	Patent family member(s)		Publication date	
US 4723161	A	02-02-1988	JP	2103366 C	22-10-1996
			JP	7118807 B	18-12-1995
			JP	61214883 A	24-09-1986
			JP	2103367 C	22-10-1996
			JP	7118808 B	18-12-1995
			JP	61214885 A	24-09-1986
			CA	1251276 A1	14-03-1989
			GB	2173067 A ,B	01-10-1986
EP 0595403	A	04-05-1994	FR	2697393 A1	29-04-1994
			AT	176842 T	15-03-1999
			AU	674884 B2	16-01-1997
			AU	5030893 A	12-05-1994
			CA	2109138 A1	29-04-1994
			CN	1088041 A ,B	15-06-1994
			DE	69323523 D1	25-03-1999
			DE	69323523 T2	02-09-1999
			EP	0595403 A1	04-05-1994
			ES	2130214 T3	01-07-1999
			FI	934714 A	29-04-1994
			JP	6233137 A	19-08-1994
			US	5500677 A	19-03-1996
WO 0232142	A	18-04-2002	US	2002037048 A1	28-03-2002
			WO	0232142 A2	18-04-2002
			US	2002037047 A1	28-03-2002
			US	2002037046 A1	28-03-2002
WO 0139503	A	31-05-2001	BR	0007657 A	06-11-2001
			CN	1355995 T	26-06-2002
			WO	0139503 A1	31-05-2001
			EP	1151613 A1	07-11-2001
			PL	348970 A1	17-06-2002
			TR	200102123 T1	21-01-2002

フロントページの続き

(74)代理人 100107766

弁理士 伊東 忠重

(72)発明者 ファン デル スハール, ミハエラ

オランダ国, 5 6 5 6 アーアー アインドーフェン, プロフ・ホルストラーン 6

F ターム(参考) 5C059 MA00 MA05 MA14 MA23 MA24 MA31 MC11 MC38 ME01 NN21

PP05 PP06 PP07 SS01 SS08 SS10 SS11 SS20 SS26 UA02

UA05

5J064 AA01 BB01 BB03 BC01 BC02 BC08 BC25 BD01