

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3733161号

(P3733161)

(45) 発行日 平成18年1月11日(2006.1.11)

(24) 登録日 平成17年10月21日(2005.10.21)

(51) Int. Cl.		F I		
GO6T	7/40	(2006.01)	GO6T	7/40 100A
HO4N	5/278	(2006.01)	GO6T	7/40 B
			HO4N	5/278

請求項の数 13 (全 15 頁)

(21) 出願番号	特願平7-196390	(73) 特許権者	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22) 出願日	平成7年8月1日(1995.8.1)	(74) 代理人	100076428 弁理士 大塚 康德
(65) 公開番号	特開平9-46519	(74) 代理人	100093908 弁理士 松本 研一
(43) 公開日	平成9年2月14日(1997.2.14)	(72) 発明者	砂川 伸一 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
審査請求日	平成14年8月1日(2002.8.1)	(72) 発明者	松林 一弘 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
		審査官	新井 則和

最終頁に続く

(54) 【発明の名称】 画像処理装置および方法

(57) 【特許請求の範囲】

【請求項1】

動画像に含まれる複数のフレーム画像それぞれに対して、各フレーム画像を複数の画素で構成される所定サイズのブロック毎に直交変換することにより空間周波数軸上データを取得する変換手段と、

前記変換手段により得られる現フレーム画像中の注目ブロックの空間周波数軸上データに基づいて、前記現フレーム画像中の注目ブロックの文字らしさを表す第1の情報を生成する第1の生成手段と、

前記現フレーム画像中の注目ブロックにおける空間周波数軸上データと従前に入力された少なくとも一つのフレーム画像中の注目ブロックにおける空間周波数軸上データとの誤差を算出し、当該算出された誤差が所定の閾値以内であるか否かに基づいて、前記現フレーム画像中の注目ブロックが静止状態であるか否かを表す第2の情報を生成する第2の生成手段と、

前記第1の情報により文字らしいと判断され且つ前記第2の情報により静止状態であると判断される前記現フレーム画像中の注目ブロックを、文字領域であるとして識別する識別手段とを備えることを特徴とする画像処理装置。

【請求項2】

前記変換手段で行われる直交変換は、離散コサイン変換、アダマール変換、フーリエ変換のいずれかを含むことを特徴とする請求項1に記載の画像処理装置。

【請求項3】

10

20

前記第2の生成手段において用いられる前記所定の閾値は、文字の微小な揺れや変化を吸収できる程度の値であることを特徴とする請求項1に記載の画像処理装置。

【請求項4】

前記識別手段は、前記第2の情報に基づいて静止状態であると判断された注目ブロックに関して、前記第1の情報に基づいて文字らしいか否か判断し、文字らしいと判断された注目ブロックを文字であると識別することを特徴とする請求項1に記載の画像処理装置。

【請求項5】

前記第2の生成手段で計算される前記誤差は、前記注目ブロックの空間周波数軸上データの各係数と、前記従前に入力された少なくとも1つのフレーム画像における注目ブロックの空間周波数軸上データの各係数との差分の合計であり、

前記第2の生成手段は、当該算出された合計値が前記所定の閾値以内であるか否かに基づいて当該注目ブロックが静止状態であるか否かを示す前記第2の情報を生成することを特徴とする請求項1に記載の画像処理装置。

【請求項6】

前記第2の生成手段は、前記注目ブロックの空間周波数軸上データから、高周波帯域における特徴量を抽出し、該高周波帯域の特徴量を用いて前記誤差を算出し、当該算出された誤差が所定の閾値以内であるか否かに基づいて、前記現フレーム画像中の注目ブロックが静止状態であるか否かを表す前記第2の情報を生成することを特徴とする請求項1に記載の画像処理装置。

【請求項7】

動画像に含まれる複数のフレーム画像それぞれに対して、各フレーム画像を複数の画素で構成される所定サイズのブロック毎に直交変換することにより、前記所定サイズのブロックのそれぞれの空間周波数軸上データを得る変換工程と、

前記変換工程で得られる現フレーム画像中の注目ブロックの空間周波数軸上データに基づいて、前記現フレーム画像中の注目ブロックの文字らしさを現す第1の情報を生成する第1の生成工程と、

前記現フレーム画像中の注目ブロックにおける空間周波数軸上データと従前に入力された少なくとも1つのフレーム画像中の注目ブロックにおける空間周波数軸上データとの誤差を算出し、当該算出された誤差が所定の閾値以内であるか否かに基づいて、前記現フレーム画像中の注目ブロックが静止状態であるか否かを表す第2の情報を生成する第2の生成工程と、

前記第1の情報により文字らしいと判断され且つ前記第2の情報により静止状態であると判断される前記現フレーム画像中の注目ブロックを、文字領域であるとして識別する識別工程とを備えることを特徴とする画像処理方法。

【請求項8】

前記変換工程で行われる直交変換は、離散コサイン変換、アダマール変換、フーリエ変換のいずれかを含むことを特徴とする請求項7に記載の画像処理方法。

【請求項9】

前記第2の生成工程において用いられる前記所定の閾値は、文字の微小な揺れや変化を吸収できる程度の値であることを特徴とする請求項7に記載の画像処理方法。

【請求項10】

前記識別工程では、前記第2の情報に基づいて静止状態であると判断された注目ブロックに関して、前記第1の情報に基づいて文字らしいか否か判断し、文字らしいと判断された注目ブロックを文字であると識別することを特徴とする請求項7に記載の画像処理方法。

【請求項11】

前記第2の生成工程で計算される前記誤差は、前記注目ブロックの空間周波数軸上データの各係数と、前記従前に入力された少なくとも1つのフレーム画像における注目ブロックの空間周波数軸上データの各係数との差分の合計値であり、

前記第2の生成手段は、当該算出された合計値が前記所定の閾値以内であるか否かに基

10

20

30

40

50

づいて当該注目ブロックが静止状態であるか否かを示す前記第2の情報を生成することを特徴とする請求項7に記載の画像処理方法。

【請求項12】

前記第2の生成工程では、前記注目ブロックの空間周波数軸上データから、高周波帯域における特徴量を抽出し、該高周波帯域の特徴量を用いて前記誤差を算出し、当該算出された誤差が所定の閾値以内であるか否かに基づいて、前記現フレーム画像中の注目ブロックが静止状態であるか否かを表す前記第2の情報を生成することを特徴とする請求項7に記載の画像処理方法。

【請求項13】

メモリ媒体から所定のプログラムを読み込むことで、動画像に含まれる複数のフレーム画像それぞれを空間周波数軸上のデータに変換し、当該変換された空間周波数軸上のデータ特性に基づいて該フレーム画像内の注目ブロックの属性を判別する画像処理装置であつて、

前記メモリ媒体は、

動画像に含まれる複数のフレーム画像それぞれに対して、各フレーム画像を複数の画素で構成される所定サイズのブロック毎に直交変換することにより、前記所定サイズのブロックのそれぞれの空間周波数軸上データを得る変換工程の手順コードと、

前記変換工程で得られる現フレーム画像中の注目ブロックの空間周波数軸上データに基づいて、前記現フレーム画像中の注目ブロックの文字らしさを表す第1の情報を生成する第1の生成工程の手順コードと、

前記現フレーム画像中の注目ブロックにおける空間周波数軸上データと従前に入力された少なくとも1つのフレーム画像中の注目ブロックにおける空間周波数軸上データとの誤差を算出し、当該算出された誤差が所定の閾値以内であるか否かに基づいて、前記現フレーム画像中の注目ブロックが静止状態であるか否かを表す第2の情報を生成する第2の生成工程の手順コードと、

前記第1の情報により文字らしいと判断され且つ前記第2の情報により静止状態であると判断される前記現フレーム画像中の注目ブロックを、文字領域であるとして識別する識別工程の手順コードとを備えることを特徴とする画像処理装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、画像データを空間周波数軸上のデータに変換して、領域を判別する画像処理装置及び方法に関するものである

【従来の技術】

従来より、複写機やOCR機器等において、画像を空間周波数軸上のデータに変換し、写真領域や網点画像領域、文字領域を分離する方法が考案されている。

【0002】

画像電子学会研究会予稿93-01-02においては、文字画像と網点画像の周波数特性の違いに着目して像域分離する方法が開示されている。この方法では、まず、画像データを8×8サイズの小ブロックに分割し、離散コサイン変換(DCT変換)を行なう。DCT変換は、JPEG標準などの画像符号化方式に広く利用されており、画像データを周波数軸上のデータに変換するものである。この結果として、ブロックの係数は、その1行1列にブロック全体の直流成分が、列方向は水平方向の周波数、行方向は垂直方向の周波数を表すデータが格納されるようになる。各方向とも、行(列)の番号が増えるに従って、より高い周波数の強さを示している。上記のDCT変換に続いて、ジグザグスキャン処理を行ない、2次元ブロックデータを1次元に変換する。これもJPEG標準で用いられている処理方法であり、図9に示すように、低周波部分から高周波部分へ斜め方向にスキャンを行なう。次のステップとして、式(1)に従って、「ジグザグレート」を計算する。

【0003】

10

20

30

40

$$\text{ZigZag_Rate}[i]=\text{ZigZag}[i]\times 2-\text{ZigZag}[i-1]-\text{ZigZag}[i+1] \quad (i=1\sim 63)$$

式(1)

続いて、ジグザグレートの高周波部分と低周波部分での積算を行ない、それぞれ ZZ_Rate_moji 、 ZZ_Rate_HT とする。

【0004】

すなわち、 $\text{Zigzag_Rate}[i]$ の絶対値の i が小さい部分を加えた値を ZZ_Rate_moji 、逆に絶対値の i が大きい部分を加えた値を ZZ_Rate_HT とする。

【0005】

そして、式(2)の判定条件が成り立つ時は文字画像、式(3)の判定条件が成り立つ時は網点画像と判定する。これは、ジグザグレートについての、文字画像は低周波部分の値が大きく、網点画像は高周波部分の値が大きいという性質を利用したものである。 10

【0006】

$\text{ZZ_Rate_moji} + \text{key}$ 定数1 式(2)

$\text{ZZ_Rate_HT} + \text{key}$ 定数2 式(3)

ここで、定数1、2は実験的に設定し、 key は周囲4ブロックの判定結果を式(4)に従って計算したものをを用いる。つまり、注目ブロックに対して、上ブロック、斜め右上ブロック、左ブロック、2つ左ブロックである。

【0007】

さらに、式(4)中の flag は、判定結果が文字ならば負、網点ならば正の値をとる。 20

【0008】

$$\text{key} = 0.25(\text{flag}(\text{上}) + \text{flag}(\text{左})) + 0.125(\text{flag}(\text{二つ左}) + \text{flag}(\text{斜め右上}))$$

式(4)

次に、画像電子学会誌第20巻5号の「適応的量子化を用いたDCT符号化法」における処理を説明する。

【0009】

これは、文字画像と網点画像を分離して画像圧縮の量子化テーブルを切り換えることで、文字画像の劣化防止と網点画像部分の圧縮率向上を図ることを目的とする。同方式においても、初めに画像データを 8×8 サイズのブロックに分割し、DCT変換を行なう。次に、図10(a)~(e)の領域90~94に含まれる係数の絶対値の和をそれぞれ算出する。そして、91~94の係数和の最大値が90より大きく、且つ、91~94の係数和の最大値が所定の閾値Aより大きいときに、ブロックを網点画像であると判定する。また、図10(f)において、領域95に含まれる係数の絶対値の和が閾値Bより大きく、且つ、網点画像ブロックと判別されなかった場合に、文字画像ブロックであると判定する。 30

【0010】

次に、特開平2-202771号公報の「ファクシミリ装置」における処理を説明する。これには、2値画像領域と中間調画像領域の分離の明確化を目的としている。同提案における像域分離パラメータ決定部では、画像データを 4×4 サイズのブロックに分割し、2次元アダマール変換を行う。像域分離パラメータLは、 Y_{ij} をアダマール変換の係数要素とすると、 40

$$L = Y_{ij}^2 \quad (i + j = 3, 4, 5, 6) \quad \text{式(5)}$$

(ここで Y^X は、YのX乗を示している。)

で計算する。そして、Lの値に従って2値化のスライスレベルを決定する。これは「2値画像領域に仮定した変換結果の方が空間周波数の高域に対してエネルギーが大」であることによる。すなわち、2値画の領域はLが大きな値に、中間調画像の領域ではLが小さな値になることを示している。

【0011】

【発明が解決しようとする課題】

しかしながら、上記技術を動画像に適用すると、各フレーム毎に独立に処理を行うことに 50

なり、時間的に隣接したフレーム間でも抽出される領域が大きく異なり、安定した抽出が行えないという問題があった。これは、各フレーム内の被写体の動きによって、文字領域と類似した特徴の領域が表れることによるものである。

【 0 0 1 2 】

【課題を解決するための手段】

本発明はかかる問題点に鑑みなされたものであり、背景の動きによる影響を除去して、高い抽出率で安定した画像切り出しを行う画像処理装置及び方法を提供しようとするものである。

【 0 0 1 3 】

この課題を解決するため、たとえば本発明の画像処理装置は以下に示す構成を備える。
すなわち、

動画像に含まれる複数のフレーム画像それぞれに対して、各フレーム画像を複数の画素で構成される所定サイズのブロック毎に直交変換することにより空間周波数軸上データを得る変換手段と、

前記変換手段により得られる現フレーム画像中の注目ブロックの空間周波数軸上データに基づいて、前記現フレーム画像中の注目ブロックの文字らしさを表す第1の情報を生成する第1の生成手段と、

前記現フレーム画像中の注目ブロックにおける空間周波数軸上データと従前に入力された少なくとも1つのフレーム画像中の注目ブロックにおける空間周波数軸上データとの誤差を算出し、当該算出された誤差が所定の閾値以内であるか否かに基づいて、前記現フレーム画像中の注目ブロックが静止状態であるか否かを表す第2の情報を生成する第2の生成手段と、

前記第1の情報により文字らしいと判断され且つ前記第2の情報により静止状態であると判断される前記現フレーム画像中の注目ブロックを、文字領域であるとして識別する識別手段とを備える。

【 0 0 1 6 】

また、前記第2の生成手段は、前記注目ブロックの空間周波数軸上データから、高周波帯域における特徴量を抽出し、該高周波帯域の特徴量を用いて前記誤差を算出し、当該算出された誤差が所定の閾値以内であるか否かに基づいて、前記現フレーム画像中の注目ブロックが静止状態であるか否かを表す前記第2の情報を生成することが望ましい。この結果、背景画像が変化する場合であっても文字領域を良好に抽出することが可能になる。

【 0 0 1 7 】

また、更に、注目ブロックに対する前記直交変換後のデータに基づいて所定の帯域における特徴量を抽出する抽出手段を備え、前記識別手段は、当該抽出手段で得られた結果を含めて、注目ブロックの属性を識別することが望ましい。この結果、背景画像が変化する場合であっても文字領域を良好に抽出することが可能になる。

【 0 0 1 8 】

【発明の実施の形態】

以下、添付図面を参照して、本発明に係る好適な実施の形態を詳細に説明する。

【 0 0 1 9 】

<装置構成の説明(図2)>

本実施形態の画像処理装置は、外部インターフェースから画像を入力し、蓄積する手段を設けたパーソナルコンピュータ等である。入力画像から文字領域を切出してデータベースを構築し、それを利用して画像検索を行なう機能を備える。

【 0 0 2 0 】

図2は、本実施形態の画像処理装置の概略構成を示すブロック図である。図中、1は装置全体の制御を行なうCPUであり、メモリ部3に格納されたプログラムに従って演算、I/O制御などの処理を実行する。周辺機器コントローラ2は、CPU1とセットで用いられ、周辺機器を制御するのに必要なI/O(シリアル通信、パラレル通信、リアルタイムクロック、タイマ、割り込み制御、DMA制御等)の制御を行なう。メモリ部3は、CP

10

20

30

40

50

U1の主装置として、DRAM、キャッシュRAM、ROM等のメモリを含み、画像領域検出処理におけるワークエリアも兼ねる。4は、ユーザデータや装置の設定、画像データの記憶を行なうハードディスクドライブである。5はFDD制御部であり、6はフロッピーディスクドライブである。7はキーボード制御部であり、8はキーボードである。表示ユニット9において、10は表示制御部であり、VRAM11から表示データを順次読み出し、階調変換等を行ないながら、液晶パネル12へデータを転送する。また、表示制御部10は、CPU1からVRAMへのアクセスと、VRAMから表示部6へのデータ転送の調停を行なう。本実施形態においては、液晶パネル(LCD)12は、各種ファイルの情報を表示すると共に、画像データを表示する。また、通信ユニット13において、15は通信インターフェースであり、14は通信制御部である。通信規格は、RS-232C、イーサネット等のシリアル通信、セントロニクス、SCSI等のパラレル通信のインターフェースであり、テキスト等の各種データや画像データ等のバイナリーデータの入出力を行なう。また、NTSC方式などのテレビ信号を入力するインターフェースを備える。16は画像データから文字領域を抽出する画像領域検出部(詳細は後述する)である。

10

【0021】

<画像領域検出部の説明(図1,図3)>

まず、図3を用いて、本実施形態における文字切り出しの様子を説明する。本実施形態では、カラーの画像をたとえば不図示のビデオカメラや、通信インターフェース13を介して入力し、その中から文字領域を抽出する処理を行なう。図3(a)は画像の例であり、タイトル文字80と81、写真画像82の3つの要素で構成されている。なお、図中のマ 20
ス目は、後述する小ブロックの境界を表わしており、実際に画像に表示されているものではない。尚、実施形態では、1マスの大きさは、8×8画素とした。図3(b)は文字抽出結果を表わす図であり、文字領域80,81のみが切出されている。本実施形態の処理は、図3(a)の入力画像から、図3(b)の画像を作成することである。

【0022】

次に、図1のブロック図を用いて、本実施形態の画像領域検出部16について説明する。

【0023】

図中、20はブロック分割部であり、画像データを8×8画素の小ブロックに分割する。続く21はブロック演算部であり、ブロック毎に文字領域候補であるかを判定し、候補格納領域26にその結果を格納する。 30

【0024】

ブロック演算部21において、22は先に従来技術の項で説明した像域パラメータ決定部である。本実施形態においては、ブロックサイズが8×8サイズであることと、直交変換として離散コサイン変換を用いる点が異なる。24は、DCT変換部、25は空間高調波係数演算部である。尚、20、24はJPEGやMPEG方式等の標準規格で使用されている技術と同様であり、詳細な原理は省略する。

【0025】

像域パラメータ決定部22によって、文字領域と中間調画像領域を分離するための特徴パラメータLが計算される。しかし、本発明の実施形態の像域パラメータ分離処理では、文字領域の特徴が検出できるものであれば、他の方式であっても構わない。23は、属性判定部であり、像域パラメータ分離部23の結果を閾値と比較して、該当ブロックが文字領域であるかの判定を行う。 40

【0026】

26は、静止ブロック検出部であり、隣接するフレームと比較し、特定位置の交流係数の変化が少ないブロックを検出する。静止ブロック検出部26において、27はフレームメモリであり、前フレームの交流係数を記憶する。また、28はブロック比較部であり、フレームメモリ27の前フレームの交流係数と、DCT変換部24で変換された現フレームの交流係数の比較を行う。

【0027】

<処理内容の説明>

50

次に上記構成における本実施形態の装置における処理内容を説明する。

【0028】

本実施形態では、動画像に合成されたキャプション文字を検出する。特にキャプション文字の背後に、無地の背景が合成されている画像を対象とする。

【0029】

まず、動画の各フレームから文字らしい特徴を持つ候補領域を検出する。同時に、前フレームの周波数係数と比較を行ない、変化の少ない(動きの少ない)領域を検出する。そして、候補領域で且つ変化の少ない部分を文字領域と判定する。これはキャプション文字の表示形態を利用したものである。すなわち、キャプション文字は、ユーザが読み終るまでの十分な時間として、数秒間は同じ位置に静止して表示される、という特徴を有する。また、一方では、背景に表示されている物体等は、物体自身の移動、又はカメラワーク等により動きが生じることが多い。そこで、本実施形態では、静止領域か否かを判定条件に加えることで、画像中の物体が文字領域として判定されるのを防ぐ。

10

【0030】

図1の構成で説明する。先ず、ブロック分割部20では入力画像を小ブロックに分割する。小ブロックデータは、ブロック単位で、DCT変換され、空間高調波係数演算部25で文字らしさを表わす特徴量が算出される。一方で、DCT係数は、静止ブロック検出部26において、前フレームの対応する位置の小ブロックの係数分布と比較が行われる。属性判定部23は、算出した現フレームの小ブロックの特徴量が文字らしさの度合いが高く、且つ、前フレームの同位置の交流係数と比較して変化が小さい場合、当該注目ブロックは文字であると判定し、その結果を出力する。この比較処理の後、現フレームの交流係数をフレームメモリ27に記憶する。

20

【0031】

次に、各構成ブロック毎の動作を説明する。

【0032】

図4は、ブロック比較部28の処理の流れを示すフローチャートであり、同図を用いて静止ブロック検出動作の説明を行なう。本実施形態のブロック比較部28は、前フレーム中の該当するブロックと現フレームの該当ブロックの各空間周波数係数との誤差を計算し、全周波数に渡って積算する。これを閾値と比較し、閾値以下の場合、ブロック内の画像が静止していると判定する。

30

【0033】

図中、ステップS100では、フレームメモリ27の検索を行ない、前フレームの対応する位置のブロックの周波数係数を取り出す。ステップS101では、係数カウンタ i, j や積算値レジスタ等の変数をクリアする。ステップS102では、前フレームの周波数係数と現フレーム中の該当する周波数係数との誤差を計算する。 i 行 j 列の現フレームの周波数係数を A_{ij} 、前フレームの周波数係数を B_{ij} とすると、誤差 E_{ij} は式(6)で計算される。

【0034】

$$E_{ij} = A_{ij} - B_{ij} \quad \text{式(6)}$$

続くステップS103では、 i 行 j 列の係数の誤差を、式(7)に従って積算する。

40

【0035】

$$Err = |E_{ij}| \quad \text{式(7)}$$

尚、この式(7)は、 $Err = Err + |E_{ij}|$ としても表される。

【0036】

ステップS104に処理が進むと、係数カウンタ i, j を更新する。ステップS105では、ブロック内の全係数の比較が終了したか判定し、否定であればステップS102に戻り、上記処理を繰り返す。

【0037】

さて、ステップS105の判定が肯定であれば、ブロック内の全ての係数についての誤差の積算値が Err に格納されていることになる。そこで、ステップS106に進み、積算

50

値 Err が、予め設定してある静止判定閾値 Ta 以上であるか判定する。

【0038】

$Err < Ta$ であれば、注目ブロックと前フレームの該当するブロックの間にはほとんど変化がないと判断できるので、ステップ $S107$ に進み、「静止フラグ」を出力し、本処理を終了する。また、 $Err \geq Ta$ であると判断できた場合には、注目ブロックと前フレーム中の該当するブロックとの間には変化があったことになるから、「静止フラグ」の出力は行なわない。

【0039】

尚、ステップ $S107$ の処理は、属性判定部 23 の出力信号をたとえば“1”にすることを意味し、ステップ $S107$ の処理をスキップすることは、その出力信号を“0”にさせたまま終了することを意味する。また、静止しているか否かの判定に用いた閾値 Ta は、文字領域の微小な揺れや変化を吸収できる程度の値である。

10

【0040】

以上の処理により、ブロック内の物体が静止しているかを検出することができる。また、上記例では、直前の1フレームとの比較を行ったが、過去の複数フレームを使用するようにしても良い。複数のフレームの周波数係数を記憶しておき、前後のフレームと比較するようにしても良い。更に、静止ブロックの判定方法として、本実施形態では各係数の誤差を積算する方法を示したが、他の方法であっても良い。例えば、ブロック内の全係数を積算してから差分をとるようにしても良いし、パターンマッチングによる方法であっても本発明の趣旨と何等異なるものではない。尚、上記処理は、ソフトウェアによるものとして説明したが、これをロジック回路で構成しても良いのは勿論である。

20

【0041】

次に、本実施形態における属性判定部 23 の処理内容を図5のフローチャートに従って説明する。

【0042】

本実施形態の属性判定部 23 は、像域パラメータ決定部 22 の出力する特徴パラメータと、静止ブロック検出部 26 の出力する「静止フラグ」から、当該ブロックが文字であるかの判定を行なう。具体的には次の通りである。

【0043】

ステップ $S120$ では、メモリ部 3 から文字判定閾値 Tb を読み込む。ステップ $S121$ では、像域パラメータ決定部 22 から特徴パラメータ L 、静止ブロック検出部 26 から静止フラグの入力を行なう。ステップ $S122$ では、入力した静止フラグに基づいて注目ブロックが静止状態か否かを判定する。静止状態でないと判定した場合には、注目ブロックは文字領域ではないと判断し、本処理を終了する。

30

【0044】

一方、静止状態にあると判断した場合には、ステップ $S124$ に進み、入力した特徴パラメータ L と、先に読み込んだ閾値 Tb と比較し、ステップ $S125$ でその比較結果に基づいて処理を分岐する。具体的には、特徴パラメータ L が閾値 Tb 以上であると判断した場合には、ステップ $S126$ に進み、注目ブロックは文字領域であると判断し、その旨を出力信号として出力する。また、特徴パラメータ L が閾値 Tb 未満であると判断した場合には、当該注目ブロックは文字領域ではないと判断し、本処理を終える。

40

【0045】

尚、単純にブロック毎の判定結果を論理レベルの異なる信号として出力するようにしてもよいし、文字領域と判定されたブロックの画像をファイルに出力するようにしてもよい。

【0046】

以上の処理により、特徴パラメータが文字判定閾値 Tb 以上であって静止したブロックのみを文字領域として検出することができる。

【0047】

以上説明したように本実施形態によれば、フレーム間で係数分布を比較して領域判定することで、フレーム毎のバラツキのない安定した領域抽出が可能となる。

50

【 0 0 4 8 】

尚、本発明は上記実施形態に限らず、幅広く応用することが可能である。例えば、実施形態では、直行変換の方法としてDCTを使用した例を示したが、フーリエ変換やアダマール変換であっても良い。また、像域パラメータ分離処理は、実施形態では、パラメータLの値を計算して閾値と比較したが、文字領域の特徴を検出する他の方法を用いても勿論構わない。また、属性判定部では、静止ブロック判定によって閾値以上のブロックを除去する例を説明したが、静止ブロックの判定によって閾値を補正するようにしても良い。これにより、静止しているキャプション文字はより確実に検出され易く、また、画像中のほぼ静止している文字領域も検出できるようになる。

【 0 0 4 9 】

また、上記処理は、パーソナルコンピュータ等の情報処理装置に備えられた拡張スロットに画像領域検出部16を装着し、その中で処理するものとして説明したが、情報処理装置のCPU1が十分高速演算できるものであれば、これに限るものではない。

【 0 0 5 0 】

この場合には、対応するプログラムをHDD4やFDD6からメモリ部3内のRAMにロードし、CPU1が実行することになる。尚、動画像の入力は、通信インターフェース15を介して入力されるものとするが、これに限らず、如何なるインターフェースを介して受信しても良い。

【 0 0 5 1 】

< 第2の実施形態 >

前述の第1の実施形態では、ブロック内の全係数を比較して領域判定を行なった。これは、無地背景上にキャプション文字が合成されているものを対象としたからである。しかしながら、変化する背景上にキャプション文字が合成されている場合には、小ブロックの係数分布はフレーム毎に変化することとなり、文字として検出が行なえない。本第2の実施形態はこれに対応して、背景の影響を除去して領域抽出率のさらなる向上を図るものである。以下では、小ブロック内の特定部分の係数のみを取り出し、比較判定する例について説明を行なう。

【 0 0 5 2 】

< 画像領域検出部の説明 (図6) >

まず、図6のブロック図を用いて、本第2の実施形態における画像領域検出部16の構成を説明する。尚、他の構成要素については、第1の実施形態と同様であるので、その説明は省略する。同図においては、静止ブロック検出部26の中に帯域分離部29を付加したことが、前述の第1の実施形態と異なる。帯域分離部29は、DCT係数の中からあらかじめ設定された部分の係数のみを取り出す。そして、フレームメモリ27に記憶すると共に、ブロック比較部28の係数データの入力とする。

【 0 0 5 3 】

< 帯域分離部の説明 (図7, 図8) >

以下、本実施形態における帯域分離部29の動作を説明する。

【 0 0 5 4 】

まず、図7は、ブロック内における帯域分割の様子を示す図である。同図において、8×8サイズの小ブロックは対角線で2分割されており、70はDCT変換後の低周波領域、71は高周波領域の交流係数である。

【 0 0 5 5 】

次に、図8は画像の種類による各周波数領域の分布状況を表す図である。図中、80は小ブロックに含まれる画像の種類、81は低周波領域の係数振幅の割合、82は高周波領域の割合を示す。81と82は領域内の係数の絶対値を積算し、ブロック全体に対する割合を求めたものである。従って、これらの合算は丸め誤差を除けば100になる。

【 0 0 5 6 】

また、83は文字(キャプション文字)画像ブロック、84は中間調画像ブロック、85は文字と中間調の混在画像ブロックの平均的な係数分布状況である。これを見ると、画像

10

20

30

40

50

種 8 3 は低周波と高周波領域が半々であり、画像種 8 4 は大部分が低周波側に集中している。また、画像種 8 5 は全域に渡って分布しているが、やや低周波側に片寄っている。

【 0 0 5 7 】

ここで、本実施形態での目的は、無地、又は、動きのある背景上に合成されているキャプション文字を検出することにあつた。これらは、無地背景の文字は画像種 8 3、動く背景上の文字は画像種 8 5 と対応しており、本実施形態は画像種 8 3 と 8 5 のブロックを検出できればよい。そして、静止ブロック検出部 2 6 は、これら両ブロックの文字特徴を示す部分が静止していることを検出する。

【 0 0 5 8 】

これについて検討すると、低周波領域 8 1 については、画像種 8 3 ~ 8 5 で分布が集中している。一方で、高周波領域 8 2 では、文字を含んでいる画像 8 3 と 8 5 の分布割合は多く、中間調画像のみの画像種 8 4 の分布割合は非常に少ない。従って、本実施形態の帯域分離部 2 9 では、高周波領域 7 1 の交流係数を取り出して、ブロックの動きを判定するのに使用する。i 行 j 列の交流係数を A_{ij} とすると、帯域分離部 2 9 は式 (8) に該当する交流係数のみを取り出して、フレームメモリ 2 7、ブロック比較部 2 8 に出力する。

【 0 0 5 9 】

$A_{ij} \quad (i + j : 7 \sim 16) \quad \text{式 (8)}$

この係数を入力するブロック比較部 2 8 の処理は、前述の実施形態と同様であり、説明を省略する。

【 0 0 6 0 】

以上の処理を行なうことで、中間調の背景の動きを除去して、文字特徴と示す部分の静止していることを検出することができる。

【 0 0 6 1 】

以上説明したように本第 2 の実施形態によれば、フレーム間で特定部分の係数比較を行って領域判定することで、背景の動きを除去した安定した領域抽出が可能となる。

【 0 0 6 2 】

なお、本発明は上述の実施形態に限られることなく、幅広く応用することができる。上述の帯域分離部の周波数帯域の分割方法は任意のものでよい。例えば、分割数も 3 つ以上にすることができ、分割形状も任意のものでよい。本実施形態では特定帯域の係数のみを取り込んで静止判定する例を説明したが、各帯域で重み付けを行って判定するようにしてもよい。また、属性判定部の判定条件として、隣接ブロックの判定結果を入れることで、更に抽出率を向上させることができる。

【 0 0 6 3 】

< 第 3 の実施形態 >

上記第 1、第 2 の実施形態では、パーソナルコンピュータ等の拡張スロットと呼ばれる部分に、画像領域検出部 1 6 という拡張ボードもしくはカードを装着し実現した。

【 0 0 6 4 】

しかしながら、動画像を取り込むインターフェースを有し、尚且つ、CPU 1 がある程度の処理速度を有している場合には、ソフトウェアでもって実現することも可能である。

【 0 0 6 5 】

たとえば図 2 における FDD 6 に装着されるフロッピーディスク、或いは、HDD 4 内のディスク等の記憶媒体には、図 1 1 に示すようなプログラムモジュールを格納させ、それを適宜メモリ部 3 内の RAM にロードすることで実現できる。

【 0 0 6 6 】

図 1 1 において、1 0 0 は記憶媒体であり、1 0 1 は直交変換を行うモジュール、1 0 2 は直交変換によって得られたデータに基づいて注目ブロックが文字領域である可能性が高いか否かを判断するための第 1 判断モジュール、1 0 3 は前フレームにおける変換後のブロックと現フレームの変換後のブロックを比較し、変化が大きいか小さいかを判断するための第 2 判断モジュール、1 0 4 は第 1 判断モジュール 1 0 2 及び第 2 判断モジュール 1 0 3 の判断結果に基づいて注目ブロックが文字領域であるか否かを識別し、その判断結果

10

20

30

40

50

を上位処理に返す識別モジュールである。

【 0 0 6 7 】

また、本発明は、複数の機器から構成されるシステムに適用しても、1つの機器からなる装置に適用してもよい。また、本発明はシステム或は装置にプログラムを供給することによって達成される場合にも適用できることは言うまでもない。この場合、本発明に係るプログラムを格納した記憶媒体が、本発明を構成することになる。そして、該記憶媒体からプログラムをシステム或は装置に読み出すことによって、そのシステム或は装置が、予め定められたし方で動作する。

【 0 0 6 8 】

【発明の効果】

以上説明したように本発明によれば、背景の動きによる影響を除去して、高い抽出率で安定した画像、特に文字領域を切り出しを行うことが可能になる。

【 0 0 6 9 】

【図面の簡単な説明】

【図1】実施形態における画像領域検出部の構成を示すブロック図である。

【図2】実施形態における画像処理装置の概略構成を示すブロック図である。

【図3】実施形態における文字領域抽出動作の画像例を示す図である。

【図4】図1における静止ブロック検出部の処理内容を示すフローチャートである。

【図5】図1における属性判定部23の処理内容を示すフローチャートである。

【図6】第2の実施形態における画像領域検出部の構成を示すブロック図である。

【図7】図6の帯域分離部で分離する領域を示す図である。

【図8】帯域領域分離による属性判断テーブルを示す図である。

【図9】従来技術におけるジグザグスキャン処理の順序を説明する図である。

【図10】従来技術における係数の積算領域を説明する図である。

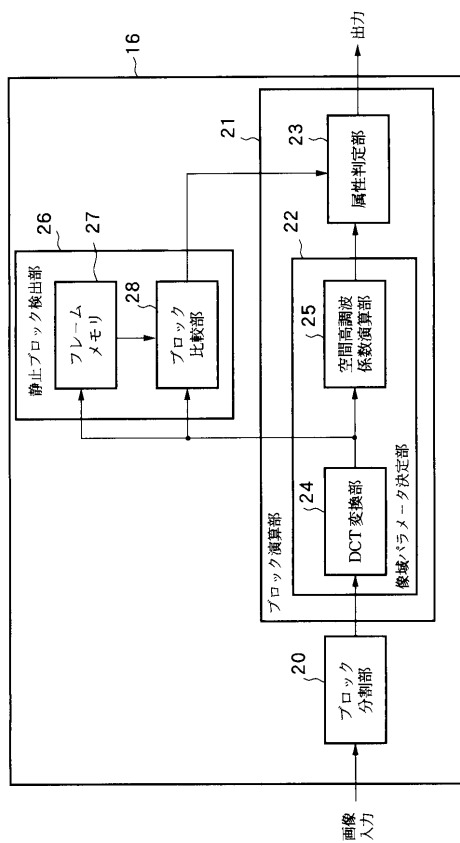
【図11】第3の実施形態におけるプログラムを記憶した記憶媒体の内容を模式的に示す図である。

【符号の説明】

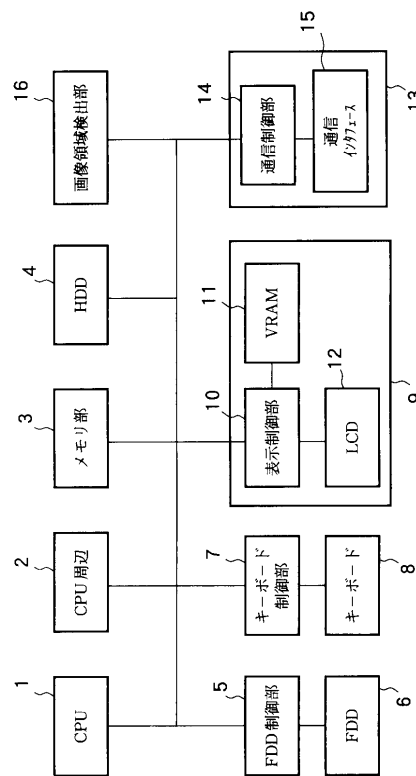
- | | | |
|----|------------------|----|
| 1 | CPU | |
| 2 | 周辺コントローラ | |
| 3 | メモリ部 | 30 |
| 4 | ハードディスクドライブ(HDD) | |
| 5 | FDD制御部 | |
| 6 | FDDドライブ | |
| 7 | キーボード制御部 | |
| 8 | キーボード | |
| 9 | 表示ユニット | |
| 10 | 表示制御部 | |
| 11 | VRAM | |
| 12 | 液晶ディスプレイ | |
| 13 | 通信ユニット | 40 |
| 14 | 通信制御部 | |
| 15 | 通信インターフェース | |
| 16 | 画像領域検出部 | |
| 20 | ブロック分割部 | |
| 21 | ブロック演算部 | |
| 22 | 像域パラメータ決定部 | |
| 23 | 属性判定部 | |
| 24 | DCT変換部 | |
| 25 | 空間高調波係数演算部 | |
| 26 | 静止ブロック検出部 | 50 |

- 27 フレームメモリ
- 28 ブロック比較部

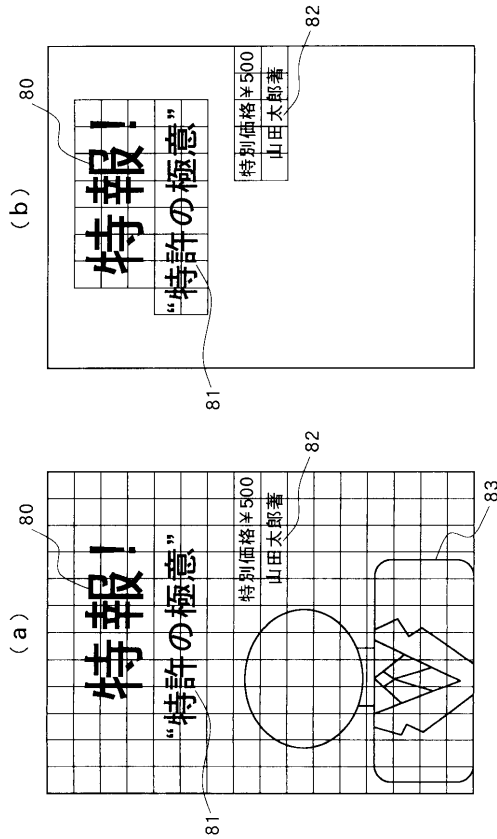
【図1】



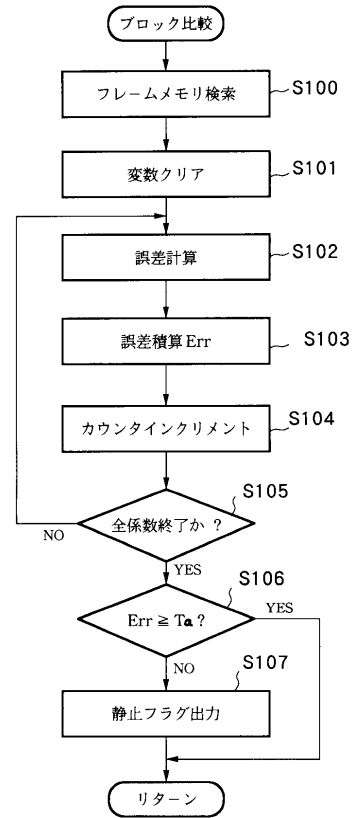
【図2】



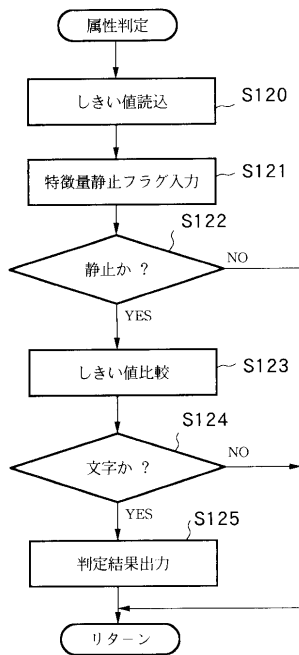
【 図 3 】



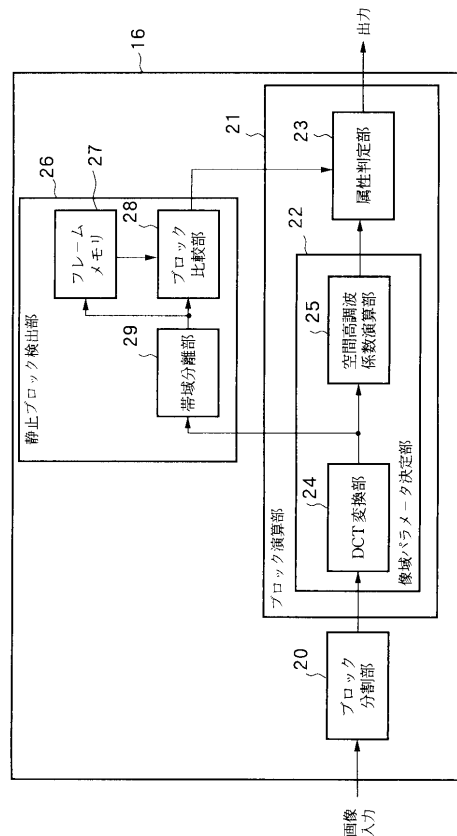
【 図 4 】



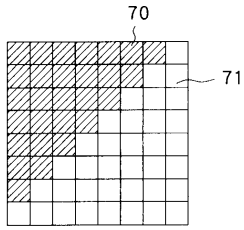
【 図 5 】



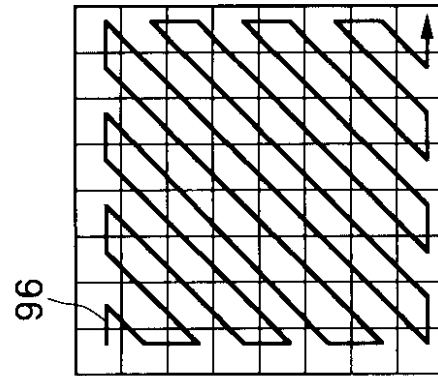
【 図 6 】



【 図 7 】



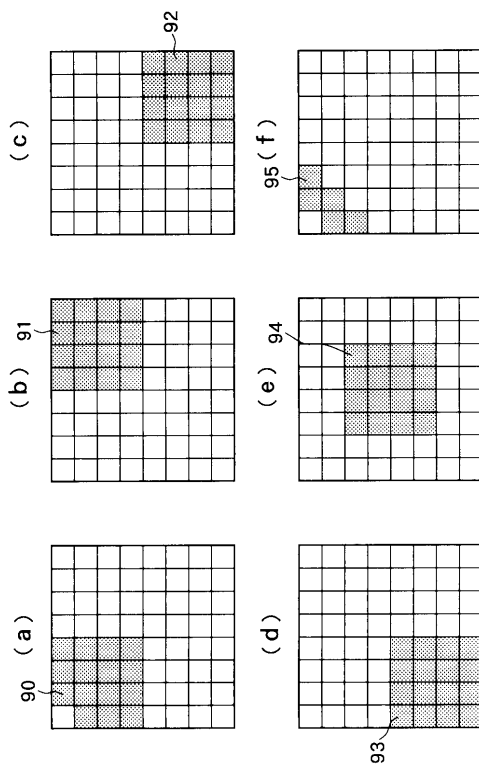
【 図 9 】



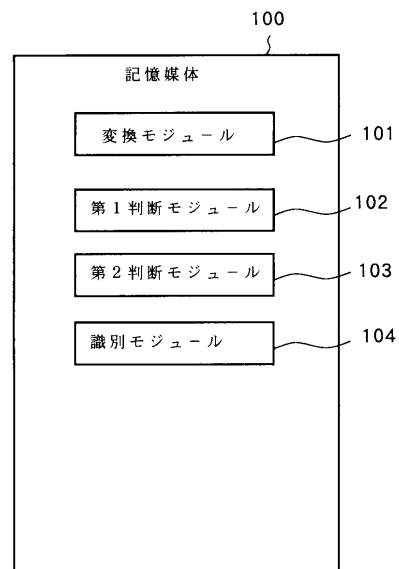
【 図 8 】

	80	81	82
	領域属性	低周波割合	高周波割合
83	文字	50	50
84	中間調	95	5
85	文字+中間調	70	30

【 図 1 0 】



【 図 1 1 】



フロントページの続き

- (56)参考文献 特開平07 - 059086 (JP, A)
特開平07 - 192003 (JP, A)
特開平07 - 074966 (JP, A)
特開昭58 - 201182 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G06T 7/00~7/60

H04N 5/278