

(43) 国際公開日
2006年2月16日 (16.02.2006)

PCT

(10) 国際公開番号
WO 2006/016591 A1

- (51) 国際特許分類:
H04N 5/91 (2006.01) G11B 27/00 (2006.01)
G11B 27/034 (2006.01)
- (21) 国際出願番号: PCT/JP2005/014598
- (22) 国際出願日: 2005年8月9日 (09.08.2005)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願2004-233945 2004年8月10日 (10.08.2004) JP
- (71) 出願人 (米国を除く全ての指定国について): ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒1410001 東京都品川区北品川6丁目7番35号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 村林 昇 (MURABAYASHI, Noboru) [JP/JP]; 〒1410001 東京都

品川区北品川6丁目7番35号 ソニー株式会社内 Tokyo (JP). 岡本 裕成 (OKAMOTO, Hiroshige) [JP/JP]; 〒1410001 東京都品川区北品川6丁目7番35号 ソニー株式会社内 Tokyo (JP). 宮本 勝 (MIYAMOTO, Masaru) [JP/JP]; 〒1410001 東京都品川区北品川6丁目7番35号 ソニー株式会社内 Tokyo (JP).

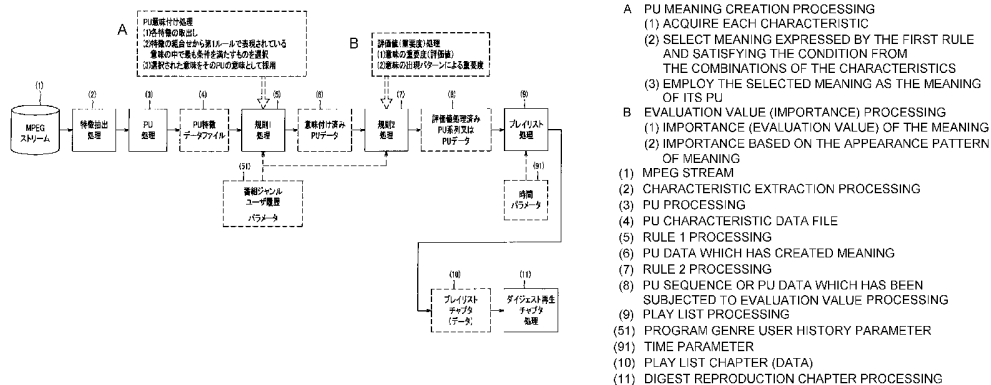
- (74) 代理人: 小池 晃, 外 (KOIKE, Akira et al.); 〒1000011 東京都千代田区千代田一丁目1番7号 大和生命ビル 11階 Tokyo (JP).

- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU,

[続葉有]

(54) Title: INFORMATION SIGNAL PROCESSING METHOD, INFORMATION SIGNAL PROCESSING DEVICE, AND COMPUTER PROGRAM RECORDING MEDIUM

(54) 発明の名称: 情報信号処理方法、情報信号処理装置及びコンピュータプログラム記録媒体



(57) Abstract: There is provided an information signal processing method used when performing a special reproduction operation for digest reproduction of a video/audio information signal of a broadcast program. The method includes: a step for reading data from the video/audio characteristic data of each predetermined section of a predetermined video/audio information signal and additional information data concerning the video/audio information signal or a predetermined recording medium containing these data; a step for creating a predetermined first information meaning for each predetermined section of the video/audio information signal from the read data; a step for connecting a plurality of sections for which the first predetermined meaning has been created so as to perform creation of a predetermined second information meaning or a predetermined evaluation processing for a predetermined section according to the signal pf the first meaning creation processing; and a step for setting a predetermined reproduction section or a predetermined reproduction time according to the signal subjected to the evaluation processing.

(57) 要約: 本発明は、放送番組の画像音声情報信号を要約再生 (ダイジェスト再生) する特殊再生動作を行う場合に用いられる情報信号処理方法であり、所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定の記録媒体からデータを読み込んでその読み込んだデータから上記画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、上記第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は上記第1の意味付け処理の信号に基づく1つの所定区間に対して所定の評価値処理を行い、上記第2の意味付け処理された信号又は上記評価値処理された信号に基づいて所定の再生区間を設定するか又は所定の再生時点を設定する。



SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT,
TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML,
MR, NE, SN, TD, TG).

(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR),

添付公開書類:

— 国際調査報告書

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

明 細 書

情報信号処理方法、情報信号処理装置及びコンピュータプログラム記録媒体

技術分野

- [0001] 本発明は、例えば放送番組における映像信号、音声信号などの画像音声データを MPEG (Moving Picture Export Group) など所定の帯域圧縮処理を行い、光磁気ディスク、ハードディスク(HDD:Hard Disk Drive)、半導体メモリなどの記録媒体に記録し再生する記録再生装置において、所定の要約再生(ダイジェスト再生)処理などの特殊再生動作を行う場合の情報信号処理方法、情報信号処理装置及びコンピュータプログラム記録媒体に関する。

本出願は、日本国において2004年8月10日に出願された日本特許出願番号2004-233945を基礎として優先権を主張するものであり、この出願は参照することにより、本出願に援用される。

背景技術

- [0002] 従来、VTR (Video Tape Recorder) やディスク記録再生装置において、長時間記録した記録内容を時間を短縮して再生してその内容を把握する場合に、音声情報の理解速度を考慮しておおよそ1.5～2倍速程度で再生処理が行われている。

さらに、時間短縮して再生しその要約再生(ダイジェスト再生)を行おうとすると、高速で再生した音声出力では音声の内容が理解できにくいため、無音とし画像出力のみの再生処理を行うようにしている。

行うことが一般的であった。

そこで、記録する放送番組の画像音声データ(画像音声情報信号、画像音声信号、画像音声情報データ)に現れる特徴などに基づいて、所定の特徴データを抽出して、その所定の特徴データを用いてキーフレーム(重要フレーム)と思われるキーフレーム区間の検出を行い、予め決めた所定の規則に従って、所定のキーフレーム区間を順次選択して再生することで、元の記録した放送番組の記録時間よりも短い所定時間内で要約再生(ダイジェスト再生)を行うことがある。

また、記録した画像データの所定区間において、一定時間間隔、例えば3分、5分、10分などの間隔毎に再生位置を示す位置情報データを自動で生成するか、又は、ユーザが手動で所望の位置に位置情報データを生成する、いわゆるチャプタデータ生成を行い、その位置情報データ(チャプタデータ)を利用して、スキップ再生、編集操作、サムネール画像表示を行う。

発明の開示

発明が解決しようとする課題

[0003] ところで、上述した特徴データは、画像信号と音声信号毎に複数種類の特徴について特徴データを抽出することができ、各々の特徴データを例えば、画像音声データの記録の際に抽出処理を行い、その画像音声データとともに特徴データを記録媒体に記録する。

これら記録された特徴データを読み出して所定のルール処理により要約再生(ダイジェスト再生)を行う区間を決定する信号処理を行うことになるが、複数存在する特徴データを各々別々にファイルとして記録媒体に記録したのでは、ファイルの数が多くなり、信号処理の際におけるファイルの取扱いが煩雑になり効率的ではない。

そこで、本発明の目的は、上述の如き従来の実情に鑑み、特徴データを用いる効果的な要約再生(ダイジェスト再生)動作又はチャプタ処理を行うために、特徴データを効率良く処理し、効果的な要約再生(ダイジェスト再生)、チャプタデータを用いる種々の動作を効率良く行うための情報信号処理方法、情報信号処理装置及びコンピュータプログラム記録媒体を提供することにある。

本発明に係る情報信号処理方法は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定の記録媒体からデータを読み込んでその読み込んだデータから画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は第1の意味付け処理の信号に基づく1つの所定区間に対して所定の評価値処理を行い、第2の意味付け処理された信号又は評価値処理された信号に基づいて所定の再生区間を設定

するか又は所定の再生時点を設定するようにしたものである。

本発明に係る他の情報信号処理方法は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定の記録媒体からデータを読み込んでその読み込んだデータから画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は第1の意味付け処理の信号に基づく1つの所定区間に対して所定の評価値処理を行い、第2の意味付け処理又は評価値処理されたデータを所定の記録媒体又は所定データメモリの所定の記録領域に記録を行うようにしたものである。

本発明に係るさらに他の情報信号処理方法は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録されたデータファイルを読み込んでその読み込んだデータから画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、第1の情報的意味付け処理データと画像音声情報信号に対する関連する付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行い、その演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は第1の意味付け処理に基づく信号に対する所定の評価値処理を行うようにしたものである。

本発明に係るさらに他の情報信号処理方法は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録されたデータファイルを読み込んでその読み込んだデータから画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、第1の情報的意味付け処理データと画像音声情報信号に対する関連する付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行い、その演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の

評価値処理を行い、第2の意味付け処理によるデータと評価値処理、又は第1の意味付け処理に基づく信号に対する所定の評価値処理によるデータが記録された所定の記録媒体からそれらデータを読み込む処理を行い、読み込んだデータに基づいて画像音声情報信号の所定区間又は所定時点を設定処理するようにしたものである。

本発明に係るさらに他の情報信号処理方法は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録されたデータファイルを読み込んでその読み込んだデータから画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、第1の情報的意味付け処理によるデータと画像音声情報信号に対する関連した付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行いその演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続する所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は第1の意味付け処理に基づく信号に対する所定の評価値処理を行い、第2の意味付け処理による所定データ又は評価値処理によるデータを用いるか又はそれらデータが記録された所定の記録媒体からデータを読み込んで、読み込んだデータに基づいて画像音声信号の所定区間又は所定時点を設定するようにしたものである。

本発明に係るさらに他の情報信号処理方法は、所定時間長の画像音声情報データから抽出した所定区間毎の特性データとその画像音声情報データに係る付加情報データを検出できる場合にはその付加情報データを用いて画像音声情報データにおける所定区間毎に生成された所定の意味的特徴データとその意味的特徴データに対する所定の評価データを検出する又はその意味的特徴データと評価データが記録された所定のファイルを読み込んで所定の処理により意味的特徴データと評価データを検出し、時間長が所定しきい値以上の場合にはデータ検出からの意味的特徴データと評価データを用いて所定ルールに従って画像音声情報データにおける所定区間を順次再生し時間長が所定しきい値に満たない場合にはルールによらないか又は所定ルール以外のルール処理に応じて時間長以下の時間内で画像

音声情報信号の所定区間を再生するように再生制御を行うようにしたものである。

本発明に係るさらに他の情報信号処理方法は、所定時間長の画像音声情報データから抽出した所定区間毎の特性データとその画像音声情報データに関する付加情報データを検出できる場合にはその付加情報データを用いて画像音声情報データにおける所定区間毎に生成された所定の意味的特徴データとその意味的特徴データに対する所定の評価データを検出する又はその意味的特徴データと評価データが記録された所定のファイルを読み込んで所定の処理により意味的特徴データと評価データを検出し、時間長が所定しきい値以上の場合にはデータ検出からの意味的特徴データと評価データを用いて所定ルールに従って画像音声情報データにおける所定区間又は所定時点を設定し、時間長が所定しきい値に満たない場合にはルールによらないか又は所定ルール以外のルール処理に応じて時間長以下の時間内で画像音声情報信号の所定区間又は所定時点を設定処理するようにしたものである。

本発明に係る情報信号処理装置は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ又はその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された記録媒体又は所定メモリ手段からデータを読み込んで、画像音声情報信号の所定区間毎の所定の第1の情動的意味付け処理を行う第1の情報処理手段と、第1の情報処理手段からの信号を用いて第1の所定の意味付けされた区間を複数接続して所定の第2の情動的意味付け処理とその第2の情動的意味に対する所定の評価値処理又は第1の情報処理手段からの信号に基づく1つの所定区間に対して所定の評価値処理を行う第2の情報処理手段と、第1又は第2の情報処理手段からの信号に基づいて所定の再生区間を設定するか又は所定の再生時点を設定する情報信号処理手段とを備える。

本発明に係る他の情報信号処理装置は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された記録媒体からデータを読み込んで、その読み込んだデータから画像音声情報信号の所定区間毎の所定の第1の情動的意味付け処理を行う第1の情報処理手段と、情報処理手段からの信号を用いて第1の所定の意味付

けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理を行うか又は第1の情報処理手段からの信号に基づく1つの所定区間に対して所定の評価値処理を行う第2の情報処理手段と、第2の情報処理手段からのデータを所定の記録媒体又は所定データメモリの所定の記録領域に記録する記録手段とを備える。

本発明に係るさらに他の情報信号処理装置は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ又はその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定記録媒体から所定データを読み込んでその読み込んだデータから画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行う第1の情報処理手段と、第1の情報処理手段からのデータと画像音声情報信号に対する関連する付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行いその演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続する所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は第1の情報処理手段からの信号に基づく所定信号に対する所定の評価値処理を行う第2の情報処理手段とを備える。

本発明に係るさらに他の情報信号処理装置は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ又はその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定記録媒体から所定データを読み込んで、その読み込んだデータから画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、第1の情報的意味付け処理データと画像音声情報信号に対する関連する付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行い、その演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は第1の意味付け処理に基づく信号に対しての所定の評価値処理を行い、第2の意味付け処理によるデータ又は評価値処理によるデータが記録された所定の記録媒体からそれらデータを読み込むデータ処理手段と、データ処理手段からの信号に応じて画像音声情報信号の所定区間を設定するか又は

所定時点を設定する情報処理手段とを備える。

本発明に係るさらに他の情報信号処理装置は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ又はその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定記録媒体から所定データを読み込んで、画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行う第1の情報処理手段と、第1の情報処理手段からのデータと画像音声情報信号に対する関連する付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行いその演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続する所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は第1の情報処理手段に基づく信号に対しての所定の評価値処理を行う第2の情報処理手段と、第2の情報処理手段からのデータ又はそのデータが記録された所定記録媒体からの所定データに基づいて画像音声情報信号の所定区間又は所定の時点を設定する情報処理手段とを備える。

本発明に係るさらに他の情報信号処理装置は、所定時間長の画像音声情報データから抽出した所定区間毎の特性データとその画像音声情報データに係る付加情報データを検出できる場合にはその付加情報データを用いて画像音声情報データにおける所定区間毎に生成された所定の意味的特徴データとその意味的特徴データに対する所定の評価データを検出するか又はその意味的特徴データと評価データか記録された所定のデータを読み込んで所定の処理により意味的特徴データと評価データを検出するデータ検出手段と、所定動作モードのときに時間長が所定しきい値以上の場合にはデータ検出手段からの意味的特徴データと評価データを用いて所定ルールに従って画像音声情報データにおける所定区間を順次再生し時間長が所定しきい値に満たない場合にはルールによらないか又は所定ルール以外のルール処理に応じて時間長以下の時間内で画像音声情報信号の所定区間を再生するように再生制御を行う再生手段とを備える。

本発明に係るさらに他の情報信号処理装置は、所定時間長の画像音声情報データから抽出した所定区間毎の特性データとその画像音声情報データに係る付加情報データを検出できる場合にはその付加情報データを用いて画像音声情報デ

ータにおける所定区間毎に生成された所定の意味的特徴データとその意味的特徴データに対する所定の評価データを検出するか又はその意味的特徴データと評価データが記録された所定のデータを読み込んで所定の処理により意味的特徴データと評価データを検出するデータ検出手段と、所定動作モードのときに時間長が所定しきい値以上の場合にはデータ検出手段からの意味的特徴データと評価データを用いて所定ルールに従って画像音声情報データにおける所定区間又は所定時点を設定して、時間長が所定しきい値に満たない場合にはルールによらないか又は所定ルール以外のルール処理に応じて時間長以下の時間内で画像音声情報信号の所定区間又は所定時点を設定処理する情報処理手段とを備える。

本発明に係るプログラム記録媒体は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定の記録媒体からデータを読み込んでその読み込んだデータから画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は第1の意味付け処理の信号に基づく1つの所定区間に対して所定の評価値処理を行い、第2の意味付け処理された信号又は評価値処理された信号に基づいて所定の再生区間を設定するか又は所定の再生時点を設定する制御プログラムがコンピュータにより読取実行可能に記録されている。

本発明に係る他のプログラム記録媒体は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定の記録媒体からデータを読み込んでその読み込んだデータから画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は第1の意味付け処理の信号に基づく1つの所定区間に対して所定の評価値処理を行い、第2の意味付け処理又は評価値処理されたデータを所定の記録媒体又は所定データメモリの所定の記録領域に記録を行う制御プログラムがコンピュータにより読取実行可

能に記録されている。

本発明に係るさらに他のプログラム記録媒体は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録されたデータファイルを読み込んで、その読み込んだデータから画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、第1の情報的意味付け処理データと画像音声情報信号に対する関連する付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行い、その演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は第1の意味付け処理に基づく信号に対する所定の評価値処理を行うことを特徴とする制御プログラムがコンピュータにより読取実行可能に記録されている。

本発明に係るさらに他のプログラム記録媒体は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録されたデータファイルを読み込んで、その読み込んだデータから画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、第1の情報的意味付け処理データと画像音声情報信号に対する関連する付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行い、その演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理を行い、第2の意味付け処理によるデータと評価値処理、又は第1の意味付け処理に基づく信号に対する所定の評価値処理によるデータが記録された所定の記録媒体からそれらデータを読み込む処理を行い、読み込んだデータに基づいて画像音声情報信号の所定区間又は所定時点を設定処理することを特徴とする制御プログラムがコンピュータにより読取実行可能に記録されている。

本発明に係るさらに他のプログラム記録媒体は、所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録されたデータファイルを読み込んでその読み込んだ

だデータから画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、第1の情報的意味付け処理によるデータと画像音声情報信号に対する関連した付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行いその演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続する所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は第1の意味付け処理に基づく信号に対する所定の評価値処理を行い、第2の意味付け処理による所定データ又は評価値処理によるデータを用いるか又はそれらデータが記録された所定の記録媒体からデータを読み込んで読み込んだデータに基づいて、画像音声信号の所定区間又は所定時点を設定することを特徴とする制御プログラムがコンピュータにより読取実行可能に記録され記録されている。

本発明に係るさらに他のプログラム記録媒体は、所定時間長の画像音声情報データから抽出した所定区間毎の特性データとその画像音声情報データに係る付加情報データを検出できる場合にはその付加情報データを用いて画像音声情報データにおける所定区間毎に生成された所定の意味的特徴データとその意味的特徴データに対する所定の評価データを検出する又はその意味的特徴データと評価データが記録された所定のファイルを読み込んで所定の処理により意味的特徴データと評価データを検出し、時間長が所定しきい値以上の場合にはデータ検出からの意味的特徴データと評価データを用いて所定ルールに従って画像音声情報データにおける所定区間を順次再生し時間長が所定しきい値に満たない場合にはルールによらないか又は所定ルール以外のルール処理に応じて時間長以下の時間内で画像音声情報信号の所定区間を再生するように再生制御を行うことを特徴とする制御プログラムがコンピュータにより読取実行可能に記録されている。

本発明に係るさらに他のプログラム記録媒体は、所定時間長の画像音声情報データから抽出した所定区間毎の特性データとその画像音声情報データに係る付加情報データを検出できる場合にはその付加情報データを用いて画像音声情報データにおける所定区間毎に生成された所定の意味的特徴データとその意味的特徴データに対する所定の評価データを検出する又はその意味的特徴データと評価デ

ータが記録された所定のファイルを読み込んで所定の処理により意味的特徴データと評価データを検出し、時間長が所定しきい値以上の場合にはデータ検出からの意味的特徴データと評価データを用いて所定ルールに従って画像音声情報データにおける所定区間又は所定時点を設定し、時間長が所定しきい値に満たない場合にはルールによらないか又は所定ルール以外のルール処理に応じて時間長以下の時間内で画像音声情報信号の所定区間又は所定時点を設定処理することを特徴とする制御プログラムがコンピュータにより読取実行可能に記録されている。

本発明により、所定の画像音声特徴データを用いたより効果的なダイジェスト再生処理、チャプタ設定処理などを行うことができる。

また、本発明により、番組毎にその番組の時間長に応じた所定の特徴データを用いる所定の効果的なダイジェスト再生動作を行うことができる。

さらに、ユーザの所望のダイジェスト再生時間に応じて、各々の番組時間長とダイジェスト再生時間の設定操作などを意識せず効果的なダイジェスト再生動作を行うことができる。

さらにまた、本発明により、ユーザは、ある機能が搭載されていない記録再生装置を購入した後でも、その機能が欲しいと思った場合に購入した装置そのものにその機能を容易に動作可能な状態とすることができ、基本的な機能だけを装備した記録再生装置を初期段階で購入し、その後、色々な各ユーザの要望に応じて自分の欲しい機能を購入した装置そのものに容易に後から装備することができるので、効率的にその装置を購入することができる。

本発明によれば、複数種類の異なる画像特徴データと音声特徴データを効率良く各々の特徴データファイル又は1つの特徴データファイルとしてまとめることができる。例えば、画像特徴としてカメラ特徴、テロップ特徴、シーン特徴、カラー特徴など、音声特徴として、無音特徴、音質特徴(例えば、話し声か否か)などといった複数ある特徴データを所定の書式により効率良くデータファイルとして処理を行うことができ、画像音声データとともに所定の記録媒体に記録してファイル管理、信号処理の際のファイル処理なども効率的に行うことができる。

また、本発明によれば、特徴データ毎にファイルを設けないので、記録媒体におい

て特徴データ毎にファイルを設定する場合に比べてそれだけファイルが占有する記録容量が少なくてすむことになる。

本発明のさらに他の目的、本発明によって得られる具体的な利点は、以下において図面を参照して説明される実施に形態から一層明らかにされるであろう。

図面の簡単な説明

[0004] [図1]図1A～図1Gは、本発明を適用した記録再生装置における要約再生、チャプタ処理の動作を示す図である。

[図2]図2は、チャプタ処理による表示の一例を示す図である。

[図3]図3は、記録再生装置における処理プロセスの一例を示すブロック図である。

[図4]図4は、記録再生装置における規則処理を示すブロック図である。

[図5A]図5Aは、記録再生装置における意味付け処理と特徴データの関係の一例を示す図である。

[図5B]図5Bは、記録再生装置における意味付け処理と特徴データの関係の一例を示す図である。

[図6]図6A～図6Cは、記録再生装置における規則ファイル書式の一例を示す図である。

[図7]図7は、記録再生装置における評価値の演算処理方法の一例を示す図である。

[図8]図8A～図8Iは、記録再生装置における時間補正関数の一例を示すグラフである。

[図9]図9は、記録再生装置における時間補正関数の一般型の一例を示すグラフである。

[図10]図10は、記録再生装置におけるビデオデータの構造の一例を示す図である。

[図11]図11は、記録再生装置における再生ユニット間の接続関係の一例の図である。

[図12]図12A、図12Bは、記録再生装置における再生ユニット間の意味付け処理の一例を示す図である。

[図13]図13A、図13Bは、記録再生装置における規則2処理の一例を示す図である。

。

[図14]図14は、記録再生装置における時間補正関数の一例を示すグラフである。

[図15]図15A、図5Bは、記録再生装置における規則ファイルの構成の一例の説明図である。

[図16]図16A～図16Dは、記録再生装置における本発明の処理プロセスの一例を示す図である。

[図17]図17は、本発明を適用した記録再生装置の構成例を示すブロック回路図である。

[図18]図18は、記録再生装置における各種所定データ記録状態の一例を示す図である。

[図19]図19は、記録再生装置における表示の一例を示す図である。

[図20]図20は、本発明を適用した記録再生装置の他の構成例を示すブロック回路図である。

[図21]図21は、記録再生装置における音声系特徴抽出処理系の構成の一例を示すブロック回路図である。

[図22]図22は、記録再生装置における音声系特徴抽出処理系の構成の他の例を示すブロック回路図である。

[図23]図23は、記録再生装置における映像系特徴抽出処理系の構成の一例を示すブロック回路図である。

[図24]図24は、記録再生装置におけるシーンチェンジ処理を示す図である。

[図25]図25は、記録再生装置におけるテロップ、カラー特徴検出領域の一例を示す図である。

[図26]図26は、記録再生装置における類似画像特徴の一例を示す図である。

[図27]図27は、記録再生装置における人物特徴検出領域の一例を示す図である。

[図28]図28は、記録再生装置における人物検出処理の一例を示す図である。

[図29]図29は、記録再生装置における人物検出(人数判定)処理の一例を示す図である。

[図30]図30は、記録再生装置における人数検出処理の一例を示す図である。

[図31]図31は、記録再生装置における人数検出処理の一例を示す図である。

[図32]図32は、記録再生装置における人数検出処理の一例を示す図である。

[図33]図33は、記録再生装置における人数検出処理の一例を示す図である。

[図34]図34A～図34Eは、記録再生装置における再生ユニット処理の一例を示す図である。

[図35]図35A、図35Bは、記録再生装置における再生ユニット処理の一例を示す図である。

[図36]図36は、記録再生装置におけるCM(コマーシャル)検出処理の一例を示す図である。

[図37]図37は、記録再生装置における再生ユニット処理系の構成例を示すブロック図である。

[図38]図38は、記録再生装置における特徴データファイルの構成の一例を示す図である。

[図39]図39は、記録再生装置における特徴データファイルの構成の一例を示す図である。

[図40]図40は、記録再生装置における特徴データファイルの構成の一例の説明図である。

[図41]図41は、記録再生装置における再生ユニットデータの階層構造の一例を示す図である。

[図42]図42は、記録再生装置における再生ユニットデータの階層構造の一例を示す図である。

[図43]図43は、記録再生装置における再生ユニット映像特徴データの構成の一例を示す図である。

[図44]図44A、図44Bは、記録再生装置におけるプレイリスト(要約)データの一例を示す図である。

[図45]図45は、記録再生装置の動作の一例を示すフローチャートである。

[図46]図46は、記録再生装置における記録時間と選択可能要約再生時間の関係の一例を示す図である。

[図47]図47は、記録再生装置における記録時間と自動設定チャプタ数一例を示す図である。

[図48]図48は、記録再生装置の記録動作の一例を示すフローチャートである。

[図49]図49は、記録再生装置の再生動作の一例を示すフローチャートである。

[図50]図50は、記録再生装置の再生動作の他の例を示すフローチャートである。

発明を実施するための最良の形態

[0005] 以下、本発明の実施の形態について、図面を参照して以下の順序で詳細に説明する。なお、本発明は、以下の例に限定されるものではなく、本発明の要旨を逸脱しない範囲で、適宜変更可能であることは言うまでもない。

(1) 本発明を適用したシステムの概要

1. 1 特徴データを用いた要約再生及びチャプタ点設定処理

ここでは、本発明の動作処理概要について説明する。

下記の動作概要に関係する信号処理については、ここでの項目の他に後の項目で詳細に説明する。

下記の説明でプレイリストデータ生成に関し、特別に記述する他に、特別に記述しない場合でもプレイリストデータ生成とチャプタデータを一緒に生成処理するものとしてもよい。

特徴データを用いた要約再生(ダイジェスト再生)及びチャプタ処理の説明図を図1の(A)～(G)に示す。

まず、特徴データを用いた要約再生動作について説明する。

(特徴データを用いた要約再生(ダイジェスト再生)処理)

ここで、図1Aに示すような画像音声データ系列があると想定する。

この画像音声データ系列は、放送番組や映画ソフトその他などがあり、ハードディスク(HDD)や光磁気ディスク、大容量半導体メモリなど所定の記録媒体を用いて、MPEG(Moving Picture Export Group)など所定の帯域圧縮信号処理を用いて記録及び再生処理を行うものとする。

画像音声データ系列において、所定の意味を設定し、シーンチェンジ、音声セグメントなどに応じて所定のビデオ構造(意味的ビデオ構造)に区切った所定区間の概

念図を図1Bに示す。

この所定の意味の設定処理、所定区間の設定処理、ビデオ構造などについては後述する。

ここで、図1Cに示すように、意味毎に区切った各々の所定区間毎に、所定時間内に記録された全区間、所定プログラム区間など、所定の全区間(所定全区間)における各々の区間の所定の評価値を設定する。この評価値が設定された区間を、それぞれ所定評価値区間(評価データ区間)とする。

ここで、「所定時間内に記録された全区間」とは、番組の枠にとらわれず、ある所定時間分の画像音声データがあった場合に、その画像音声データの全区間を示す。

また、「所定プログラム区間」とはある1つの番組の画像音声データがあった場合に、その番組の枠の全区間を示す。

ここで、所定の評価値は所定全区間における所定キーフレーム区間(重要フレーム区間、重要(画像音声)区間)となる場合ほど、高い評価値(評価データ)を設定すると仮定する。

すなわち、評価値が高く設定された区間(重要評価値区間)を再生すれば、その区間にはキーフレーム区間が含まれるので、全区間を再生しなくても概要を把握することができることになる。

図1Cは、所定評価値区間の概要を示すもので、図1Aに示す画像音声データ系列で、f1～f2、f4～f5、f7～f8の各区間が評価値において設定したしきい値Th以上の区間で、図1Dに示すようにA1、A2、A3の各区間を所定の要約再生モード時にスキップ再生することで所定の要約再生(ダイジェスト再生)行うことになる。

(特徴データを用いた自動チャプタ処理)

図1Eは、チャプタ点を設定する場合の概念図であり、先に説明したような、所定キーフレーム区間(重要フレーム区間)の先頭又はその近傍、及び、そのキーフレームの区間の最後に続く(最後に接続する)キーフレーム区間ではない区間の先頭又はその近傍にチャプタ点を設定する。

ここで、例えば、従来用いられているDVD(Digital Versatile Disc)記録再生装置で自動チャプタ機能と言われる所定区間の区切り点を設定することで、その時点を編集

操作の目安にしたり、早送りフォワード再生 (FF再生)、早送り逆再生 (リwind再生、REW再生) などの場合に利用することができる。

従来、上述した自動チャプタ機能として、例えば、5分等間隔、10分等間隔、15分等間隔などのように時間間隔を等間隔とした処理が知られており、このようなチャプタ処理では図1Gに示すように、キーフレームと思われる時点の開始点にはチャプタ点を設定できない場合がある。

また、従来、手動チャプタ処理という、ユーザ自身が、所望するに任意の時点にチャプタ点を設定できる機能が知られているが、この機能は、ユーザ自身が記録した、あるいは記録する番組 (プログラム) を実際に見て設定処理を行うことになるので、ユーザにとっては面倒な操作であり、効率的ではない。

これに対し、本発明の特徴データを用いたチャプタ点設定処理 (所定時点設定処理、所定位置設定処理) では、図1Eに示すように、適切にキーフレーム区間の先頭又はその近傍と、そのキーフレーム区間の最後に接続され、又は最後に続くキーフレーム区間ではない区間の先頭又はその近傍にチャプタ点を自動的に設定処理することができるので、従来のチャプタ処理よりも、より効果的なチャプタ点設定を行うことができる、このチャプタ処理を用いた効果的な編集操作 (編集処理) や、FF再生、REW再生を行うことができる。

ここで、図1Fに示す自動設定したチャプタ点を所定の大きさのサムネール画像として所定の画像モニタに表示させる場合の概念図を図2に示す。

図1Fに示すように、f1、f4、f7が各々、所定キーフレーム区間A1、A2、A3の先頭又はその近傍で、f3、f6、f9が各々A1、A2、A3の区間の後のキーフレーム区間ではない区間B1、B2、B3の先頭又はその近傍であり、ユーザは図2に示すような表示画面を見ることで、例えば、記録再生装置の記録媒体であるハードディスクに記録された放送番組の編集操作において、図1Dに示すキーフレーム区間A1、A2、A3を切り出し、DVD (Digital Versatile Disc) などのディスク記録媒体に記録することなどの処理や、f1、f4、f7の時点にスキップ再生などの操作を想定する。

図1Gに示す従来の所定時点設定点 (チャプタ点、所定位置設定点) の一例を示すように、所定の一定間隔、例えば、5分間隔、10分間隔などの一定間隔又は略一定

間隔で設定点(チャプタ点)が設定処理されるが、図1C、図1Gから分かるように、必ずしもキーフレーム(重要フレーム)に設定されるとは限らない。

このように本発明における特徴データを用いて自動的に所定のチャプタ点(所定設定点、又は所定区切り点)又はセグメント処理を行うことで、より効果的な編集操作やスキップ再生を行うことができる。

1. 2 本発明の処理プロセスの一例

次に、本発明における処理プロセスの一例を図3に示す。

図3に示す処理プロセスでは、MPEG画像音声ストリームデータから、画像系及び音声系の各特徴データを抽出する特徴抽出処理(2)を含んでいる。

ここでは、簡単のため、MPEGストリーム(1)(MPEGデータ)は、所定記録媒体に記録する、又は、所定記録媒体に記録されているデータを想定しているが、例えば、所定の伝送系(有線系又は無線系)において伝送される画像音声データにおいても同様に本発明を適用することができる。

特徴抽出処理(2)は、記録処理と同時に行うことができるが、所定の記録媒体に既に画像音声データが記録されている場合には、その記録媒体から再生して所定の特徴抽出処理を行うこともできる。

ここで、規則処理(ルール処理)について説明する。

この規則処理は、ルールが所定の様式で記述されたルールファイル、又はルールデータを用いて所定の処理が行われる。

ルールファイルは、例えば、番組ジャンルに応じた、特徴データに基づくルールが記述されており、このルールファイルと所定区間の各特徴データが記述されたPU特徴データファイル(再生ユニット特徴データファイル)との演算により、所定プレイリストファイルが生成されることになる。

ここで、説明を分かりやすくするため、便宜上、所定番組ジャンル n に対するルールファイルを $Rf(n)$ 、PU特徴データファイルを Pu 、プレイリストファイルを Df とし、所望の要約時間を t とすると、以下の(1)式のような演算で表現できる。

$$Df = Pu(*)Rf(n)(*)t \cdots (1)$$

ここで、 $(*)$ は所定ファイルのデータを用いた所定の演算子と仮定する。

ルールファイル $Rf(n)$ は、以下で説明するように、例えば、所定の書式で記述され、所定の時間補正関数、意味、意味の重み付け係数(評価値、重要度)などの所定パラメータのデータなどにより構成されている。

(再生ユニット処理)

特徴抽出処理(2)の後、本発明の特徴の1つであるPU処理(3)(再生ユニット処理)を行う。

PU処理(3)において、各特徴データは、PU(再生ユニット)という区切り(4)で所定のデータ(PU特徴データファイル)として所定の記録媒体又はバッファメモリに記録(記憶)される。

(規則1処理)

PU特徴データファイルは、所定の規則1処理(5)によりPUの意味付け処理が行われる。後で説明するが、規則1処理(5)の概要は次の通りである。

(処理1) 各特徴データの取り出し

(処理2) 特徴データの組合せから第1ルールで表現されている意味の中で最も条件を満たすものを選択

(処理3) 選択された意味をそのPUの意味として採用

この規則1処理(5)では、EPG(電子番組ガイド)その他により、番組ジャンル、又は、過去にユーザが視聴した番組ジャンル、時間帯、再生回数、再生時刻、再生日時、その他などのパラメータ、サイド情報などが利用できる場合には、これらパラメータを考慮して所定の処理を行うようにしてもよい。

この処理に関連して行われる時間補正関数の処理については後述する。

(規則2処理)

意味付けされたPU(6)は、所定の規則2処理(7)で所定の評価値処理が行われる。

規則2処理(7)では、次の(処理1)及び(処理2)の重要度についての評価値処理を行う。

(処理1) 意味の重要度

(処理2) 意味の出現パターンによる重要度

所定の評価値処理が行われたPU (8) では、PU 単体、又は、PU が幾つか連結されたPU 群で所定の評価値が付けられている。

ここで、規則1処理 (5)、規則2処理 (7) では、図4に示すように、ルール切替処理系900により、複数の番組ジャンルに応じたルール処理データとして、ジャンルA規則データ、ジャンルB規則データ、ジャンルC規則データ、・・・と幾つかの規則処理用データ (ルール処理データ) を備え、システムコントローラ系20に入力した番組ジャンル情報データに応じて、規則1処理 (5)、規則2処理 (7)、又は、何れか一方のルール処理を切り替える。

また、図4に示すように、個人別にルール処理用データを幾つか設けて切り替える。

この場合は、所定動作モードにおいて、システムコントローラに入力された所定のユーザによる設定処理により、個人1用規則処理データ、個人2用規則処理データ、個人3用規則処理データ、・・・の何れかが、システムコントローラ系20を介して選択処理され、その選択された規則処理データに基づいて所定のルール処理が行われる。

図4に示すような個人別の規則処理データを設けることで、例えば個人別に、通常再生又は特殊再生などの所定再生動作を行い、その再生状態、再生位置などの動作情報、動作位置情報などを所定の個人別規則処理に反映できるように所定メモリ手段に記憶して、それら情報データを所定の学習処理により、個人別規則処理データとして、随時、所定のタイミングでデータを更新処理するなどの動作を行うことにより個人別学習処理には有効な信号処理方法となる。

図4に示すように、ルール切替処理系901により各個人別ルール処理 (規則処理) を切り替える場合も、規則1処理 (5)、規則2処理 (7)、又は、どちらか一方のルール処理を切り替える。

(規則処理の書式)

(規則1処理の場合)

ここで、意味付け処理されたPUは、例えば、ある放送番組を想定した場合に、以下のような英文字と意味を設定して、所定の画像音声特徴データと関連させて記述する。

文字に対する意味付けは、その放送番組においてキーフレーム (重要フレーム、重

要シーン)と想定されるであろうシーン、又は要約再生、チャプタ設定などに有効と想定される所定の記録、再生区間を選択して記述する。

また、ユーザが所望するシーンを記述する。この場合は、所定の調整モードなどで、ユーザが所望するルールを記述できるようにする。

ここで、図5A、図5Bの上段に示すように、ニュース番組、相撲番組の場合の一例について示すと次の表1のようになる。

表1:ニュース(報道)番組の場合の一例

設定文字 (意味文字)	意味
a	アナウンサー出現のニュース項目シーン
b	ニュースの事件、事故などの現場シーン
c	天気予報
d	番組開始シーン
e	番組終了シーン

ここで、図5Aに示す例では、aでアナウンサーのシーンを抜き出すルールを記述しているが、1つの規則処理では、すべての想定されるaのシーン(アナウンサーの出現シーン)を抽出することはできないと思われるため、幾つかの複数の規則に分けて記述するようにする。

図5A中のb, c, d, eなど他の場合についても同様に複数の規則に分ける。

相撲番組の場合では、次の表2のようになる。

表2:相撲番組の場合の一例

設定文字 (意味文字)	意味
a	取組み紹介シーン
b	立会いシーン
c	取組みシーン
d	番組開始シーン
e	勝敗決定シーン

図5Bに示す例においても、aの取組み紹介シーンにおいて、すべての想定されるaのシーン、例えば、対戦に関連する力士、行司、審判員等の全員を抽出することがで

きないものと想定されるので、幾つかの複数の規則に分けて記述するようにする。例えば、各文字に対して複数の規則を設定する。また、場合に応じて、抽出したいシーン(キーフレーム)の想定される規則を分けて記述を行う。

放送番組では、一義的に意味付けできないシーンも想定できる。例えば、定義文字を@とし、次の表3のように設定することもできる。

表3

設定文字 (意味文字)	意味
@	意味なし (意味無定義)

上述のように設定した定義文字(設定文字、意味文字)に対する規則1処理について、ニュース番組の場合を例に具体的に説明する。

図18に示すように、各所定の特徴データが検出される場合に、上述したニュース番組の場合の定義文字a, b, c, d, eに対する各シーンが対応すると仮定する。

ここで、図5A、図5B中○の場合は論理積、△の場合は論理和の所定処理と仮定し、例えば、定義文字a のアナウンサーのシーンでは、音声特徴の属性が話者音声、色特徴の検出領域2又は検出領域3で所定の色が検出され、類似画像情報の頻度1位又は2位が検出され、人物特徴の検出領域1又は検出領域2又は検出領域5で検出され、カメラ特徴は静止の場合と想定できる。

他のb, c, d, eなども図5A、図5Bの各○、△印に応じて、上述のaの場合と同様に、各所定の特徴データと関係付けて、定義文字と特徴データとを関係付けることができる。

上述したように各定義文字と各特徴データは、所定の処理すなわち規則1処理、規則2処理を行うため、所定の書式に従って記述する。

図6Aは、その一例で、ベクトル成分のように想定して記述するものである。

すなわち、図5A、図5Bに示した各特徴データを、例えば、音声特徴の属性とし、属性が話者音声であるときはA1、属性が音楽であるときはA2、属性がその他の場合はA3とする。

映像特徴の色特徴で、領域1はB1、領域2はB2などとする。

以下、同様に、各特徴に対して、B1～B4、C1～C2、D1～D5、E1～E4、F1～F4、G1などが設定できる。

図6Aにおいて、例えば、定義文字aの場合は、次の(2)式のように記述できる。

$$a = 1.0(A1)100 * (1.0(B2)100 + 1.0(B3)100) * (1.0(C1)100 + 1.0(C2)100) * (1.0(D1)100 + 1.0(D2)100 + 1.0(D5)100) * 1.0(F1)100 \dots (2)$$

他の定義文字に対しても、図6Aに示すように記述できる。

なお、ここで、「*」は論理積(AND)、「+」は論理和(OR)と同様の所定演算を表現するものとする。

ここで、例えば、1.0(A1)100の記述について説明する。

上述したように、(A1)は、音声特徴で属性が話者音声の場合を表現している。

(重み付け係数)

1.0(A1)100の1.0は、(A1)に対する重み付け係数で、ここでは、便宜上、0～1.0の範囲を想定している。

重み付け係数は、所定演算を行うための、便宜的な係数なので、重み付け係数は、0～100、又は0～10の範囲で設定(記述)する。

(検出割合係数)

1.0(A1)100の100は、(A1)に対する検出割合係数で、その再生ユニット区間で、100%検出される場合に、1.0(A1)100は、その条件を満たすものとする。

例えば、1.0(A1)50の場合は、その再生ユニット区間で、50%検出される場合に、1.0(A1)100は、その条件を満たすものとする。

この検出割合については、下記の(3)式で説明する。

ここで、検出割合係数は、便宜上、0～100の範囲を想定している。

検出割合係数は、所定演算を行うための、便宜的な係数なので、0～1の範囲で設定することや、0～10の範囲で設定(記述)する。

ここで、この検出割合係数は、その特性がその再生ユニット区間で検出できた割合とすることができる。

例えば、上述の1.0(A1)100では、話者音声は100%検出しなければ、(A1)の

特性を検出したと判定しないとすることができる。

例えば、1. 0 (A1) 50では、50%検出したらその特性を検出したと判定する。すなわち、その所定区間において、所定の特性が検出された割合を係数で表現できる。

(特徴データの検出割合)

そこで、特性の検出の割合について説明する。

処理方法については、図34～図35を参照して後述するが、本発明では、音声セグメント特徴とシーンチェンジ特徴に応じて設定処理される再生ユニット(又はプレイユニット)(PU)という所定の区間を設定する処理概念を導入している。

そこで、例えば、そのPU区間全体に対する所定の特徴データが検出された割合で、上述した各所定の特性の割合を演算する。

例えば、図7において、ある再生ユニットの区間長(フレーム長、時間長など)を f_a とし、ある特徴データPの検出区間を f_0 , f_1 と仮定すると、この場合の特徴データPの検出割合Fは、次の(3)式にて演算処理することができる。

$$F = \Sigma f_i / f_a$$

$$= (f_0 + f_1) / f_a \quad \cdots (3)$$

この(3)式による演算値は、後で説明する評価値処理において用いることになる。

(評価値の演算方法の例)(例1)

評価値(重要度)の演算方法の一例を示す。

各特徴データについて、理想値と検出結果を以下のように処理する。

例えば、 $p = m(M)n$ として次の(処理1)～(処理5)を行う。

(処理1) (3)式を用いて、各所定特徴データの検出割合 s を演算する。

(処理2) 上記検出割合係数 n と比較して、

$$s < n \text{ の場合、 } p = m \times s \quad \cdots (4)$$

$$s \geq n \text{ の場合、 } p = m \times 100 \quad \cdots (5)$$

とする。

(処理3) 上記処理で、各特徴Mにおいて、M1, M2...など同じ属性の特徴の場合で論理和(+)処理の場合は平均処理を行う。

論理積(*)処理の場合は、便宜上、論理積係数 r というような処理概念を導入し、

その平均処理の結果に掛けた値とする。

(処理4) 上述の処理を各特徴データM毎に行い、各演算値の加算処理を行ってその処理結果をその評価値とする。

(処理5) 演算した評価値を比較して評価値が最も大きい場合の意味をその再生ユニットaの意味とする。

上述の評価値処理は、処理方法の一例で、検出された特徴データ、又はその再生ユニット区間において検出された割合などと、設定した「意味」との対応が所定の妥当性を持った処理方法であれば、上記以外の処理方法でもよい。

例えば、上述の(処理3)の処理で論理積処理の場合は、平均化処理や論理積係数を掛けないで、同じ属性の特徴データを加算する処理だけにするなどすることが想定される。

(処理3)の処理の場合で、同じ特徴データで論理積処理の場合は、検出条件が論理和処理の場合と比較して厳しくなるので、検出値を論理和処理の場合よりも大きく取るように処理を行うことができる。

ここで、上述の(2)式の場合について説明する。

例えば、各特徴の検出割合を以下の表4のようにし、検出割合係数、重み係数を一緒に示す。

表4

特徴	検出割合	検出割合係数	重み係数	p
A 1	1 0 0	1 0 0	1 . 0	1 0 0
B 2	8 0	1 0 0	1 . 0	8 0
B 3	8 0	1 0 0	1 . 0	8 0
C 1	1 0 0	1 0 0	1 . 0	1 0 0
C 2	1 0 0	1 0 0	1 . 0	1 0 0
D 1	8 0	1 0 0	1 . 0	8 0
D 2	8 0	1 0 0	1 . 0	8 0
D 5	8 0	1 0 0	1 . 0	8 0
F 1	8 0	1 0 0	1 . 0	8 0

ここで、B2、B3や、C1、C2などのように同じ特徴の種類で、検出属性が異なる場

合や、あるいは検出領域が異なる場合などで、論理和処理(+)の場合は、便宜上、平均処理を求め、(2)式から、評価値hは、次の(6)式にて示される。

$$\begin{aligned} h &= 100 + (80 + 80) / 2 + (100 + 100) / 2 + (80 + 80 + 80) / 3 + 80 \\ &= 100 + 80 + 100 + 80 + 80 \\ &= 440 \quad \dots (6) \end{aligned}$$

又は、特徴データの種類の平均化した値を評価値とすることができ、その場合は、特徴データは、A～Fの5種類なので、次の(7)式に示すような評価値とすることもできる。

$$\begin{aligned} h &= 440 / 5 \\ &= 88 \quad \dots (7) \end{aligned}$$

(属性が同じ特徴データ間の関係が論理積処理の場合)

ここで、(2)式の処理で、同じ属性の特徴データ、例えば、B2、B3が論理積処理の場合、すなわち、(1.0(B2)100 * 1.0(B3)100)のような場合について検討する。

上述の評価値処理の(処理3)から論理積処理係数rという概念を導入し、r(80+80)/2のような処理を想定する。

例えば、r=1.5とすると、

$$\begin{aligned} h &= 100 + 1.5 \times (80 + 80) / 2 + (100 + 100) / 2 + (80 + 80 + 80) / 3 + 80 \\ &= 100 + 120 + 100 + 80 + 80 \\ &= 480 \quad \dots (8) \end{aligned}$$

また、特徴データの種類の5で平均化処理して

$$\begin{aligned} h &= 480 / 5 \\ &= 96 \quad \dots (9) \end{aligned}$$

という、評価値とすることができる。

これは、論理積処理の場合が論理和処理に比較して条件が厳しいので、検出した「意味」の評価値を大きく設定した方がよいとする場合である。

また、場合によっては、r=0.8として、

$$\begin{aligned} h &= 100 + 0.8 \times (80 + 80) / 2 + (100 + 100) / 2 + (80 + 80 + 80) / 3 + 80 \\ &= 100 + 64 + 100 + 80 + 80 \end{aligned}$$

$$=424 \quad \cdots (10)$$

また、特徴データの種類5で平均化処理して

$$h=424/5$$

$$=84.5 \quad \cdots (11)$$

という評価値とすることもできる。

これは、上述の場合とは逆に、論理積処理の場合が論理和処理に比較して条件が厳しいので、評価値を小さく設定した方がよいとする場合である。

(属性の異なる特徴データ間の関係が論理和処理の場合)

ここで、例えば、(2)式で示したように、属性の異なる特徴データは、論理積演算子(*)で表現しているが、論理和演算子(+)の場合もある。

簡単のため、(2)式で第1項目A1、第2項目B2だけを求め、

$$a=1.0(A1)100+1.0(B2)100 \quad (12) \text{式}$$

上述の評価値演算方法(3)で説明したような、便宜上、論理和係数wというような概念に基づいて処理を行う。

この場合、(12)式から、評価値hは、

$$h=(100+80)w \quad \cdots (13)$$

となる。ここで、

$$w=1$$

の場合は、論理積処理の場合で、

$$a=1.0(A1)100*1.0(B2)100 \quad \cdots (14)$$

$$h=100+80$$

$$=180 \quad \cdots (15)$$

となる。

例えば、(8)式の論理和処理の場合には、

$$w=1.5 \quad \cdots (16)$$

として、

$$h=(100+80) \times 1.5$$

$$=270 \quad \cdots (17)$$

と、論理積処理の場合よりも高い評価値となるような処理を行う。

また、

$$w=0.8 \cdots (18)$$

として、

$$\begin{aligned} h &= (100+80) \times 0.8 \\ &= 144 \cdots (19) \end{aligned}$$

のように、論理積処理よりも小さい評価値となるような処理を行う。

評価値処理は、設定した意味と各特徴データ、各種の係数などを結びつけた式の値の評価のために便宜上、導入した概念なので、上記評価式の各係数の範囲、値などは、上述の説明で述べた場合に限らず、小さく、又は大きく設定することもできる。

以下のような評価値の演算により、ルールファイルにより、ルールに記述された再生ユニットの各区間の評価値が決められ、例えば、要約再生モードの場合は、要約再生時間に応じて、評価値の大きいPU区間が選択され、要約時間にできるだけ近くなるように、段々と評価値の小さいPU区間を選択していく。

そして、選択した各PU区間を再生することで、所定の要約再生が実現できる。

(評価値処理の他の処理方法)

上記で述べた各特徴データ n の一項と、所定演算子 $*$ とから $w(M) * k$ とし、各所定特徴データの検出割合 det 重み係数 w 、検出割合係数 k として、評価式の各項の特徴データ n の重み係数を $w(n)$ として、演算関数 P と演算子 $*$ とする。

$P(*k(n), \text{det}(n))$ とし、

$$d(n) = P(*k(n), \text{det}(n)) \cdots (20)$$

とする。

ここで、演算子 $*$ 以下の何れかに該当するものとして、

$d(n)$ は、

(1) $* = (| | >)$ の場合、すなわち $P((| | >)k(n), \text{det}(n))$ で、

if $(k(n) \leq \text{det}(n))$ then $d(n) = 0 \cdots (21)$

else $d(n) = 100 \cdots (22)$

(2) $* = (| | <)$ の場合、すなわち $P((| | <)k(n), \det(n))$ で、

if $(k(n) > \det(n))$ then $d(n) = 0 \cdots (23)$

else $d(n) = 100 \cdots (24)$

となる。

上記(1)、(2)のような処理の場合は、検出 $\det(n)$ と設定検出割合 $k(n)$ に応じて、途中処理値 $d(n)$ を100又は0に処理するので、下記の(3)又は(4)で説明する途中処理値が差分値になる場合に比較して、特徴データを顕著に特徴付けたい場合には有効である。

また、さらに、

(3) $* = (| >)$ の場合、すなわち $P((| >)k(n), \det(n))$ で、

if $(k(n) < \det(n))$ then $d(n) = 0 \cdots (25)$

else $d(n) = | k(n) - \det(n) | \cdots (26)$

(4) $* = (| <)$ の場合、すなわち $P((| <)k(n), \det(n))$ で、

if $(k(n) > \det(n))$ then $d(n) = 0 \cdots (27)$

else $d(n) = | k(n) - \det(n) | \cdots (28)$

であるから、評価値は次の(29)式のようになる。

[数1]

$$\sum^n w(n) \times (100 - d(n)) / \sum^n w(n)$$

$\cdots (29)$

上述の演算子の導入により、例えば、A1、B2の特徴データがあった場合に以下のように記述することができる。

$$a = 1.0(A1)(| | <)100 + 1.0(B2)(| <)100$$

$\cdots (30)$

この場合、例えば、A1特徴の検出割合(実際の検出値)を100、B2特徴の検出割合(実際の検出値)を80と仮定すると、上記(1)、(4)から、評価値 h は、

$$h = (1.0 \times (100 - 0) + 1.0(100 - 80)) / (1.0 + 1.0) = (100 + 20) / 2$$

$$=60 \quad \dots (31)$$

という評価値とすることができる。

上述のように、評価値処理の方法には、幾つかの方法があるが、ここで説明した方法に限らなくともよい。

ここで、図6Aに示す規則1の記述においては、記述するデータの出現パターン(意味)の表現方法の一例で、意味として、a, b, c...などとしたが、その否定としてA, B, C, ...、また、ワイルドカードとして、*などを用いることができる。

(規則2処理の場合)

規則2処理では、上記規則1処理で意味付けされた所定区間である再生ユニット同士の意味の接続を考慮して処理を行うようにする。

また、時間補正関数を用いて時間的な補正、すなわち時間的な重み付け処理を行う。

例えば、上記規則1処理において、意味aの評価値を70、意味bの評価値を80と仮定すると、(ab)の評価値gは、

$$\begin{aligned} g &= 70 + 80 \\ &= 150 \end{aligned}$$

又は、意味数の平均を求め、ここでは、abの2つなので、

$$\begin{aligned} g &= 150 / 2 \\ &= 75 \end{aligned}$$

又は、それぞれの評価値の積を求め、

$$\begin{aligned} g &= 70 \times 80 \\ &= 5600 \end{aligned}$$

例えば、便宜上、最大値を100と仮定して最大値で正規化すると、

$$\begin{aligned} g &= 5600 / 100 \\ &= 56 \end{aligned}$$

とすることができる。

時間補正関数の重み付けは、例えば、上述の(ab)が、ある時点tで検出でき、その評価値がgで、tでの時間補正係数(重み付け係数)をwとすると、gtを最終的な評価

値とする。

時間補正関数は、ルールファイルにおいて規則2の所定記述場所に、所定記述規則に従って、その変化点(変化点座標系の情報データ)を記述する。

規則2処理の一例を図6Bに示す。

(時間補正関数)

まず、時間補正関数について説明する。

この時間補正関数は、ルールファイルにおける所定番組ジャンルにおける要約時間補正を行うために利用することができる。

これは、ユーザによっては、所定の放送番組によっては、放送時間の前半や後半を重点的に再生したいという場合も想定できる。

そこで、記録する番組のジャンル、放送時間、その他などの、その放送番組に応じた種々のパラメータを考慮して、要約再生(ダイジェスト再生)を行う所定の再生区間に対して時間(場合によっては時刻)の重み付けを行うようにすることができる。

すなわち、この重み付けを行う区間は、時間的にそれ以外の重み付けを行わない区間に比較して、要約再生(ダイジェスト再生)を行う場合の再生の重要度を大きく処理することになる。

図8A～図8Iは、上述した時間の重み付けを行うための時間補正関数の一例を示すものである。

図8Aは、フラットな特性で、所定の要約再生区間に対して時間の重み付けを行わない場合である。

図8Bは、所定の区間内において、前半部の方を後半部に比較して、要約再生における重要度としての再生の重みを大きくする重み付けを行っている場合である。

図8Cは、所定の区間内において、後半部の方を前半部に比較して、要約再生における重要度としての再生の重みを大きくする重み付けを行っている場合である。

図8Dは、所定の区間内において、前半部と後半部を中間部に比較して、要約再生における重要度としての再生の重みを大きくする重み付けを行っている場合である。

。

図8Eは、所定の区間内において、中間部を前半部及び後半部に比較して、要約

再生における重要度としての再生の重みを大きくする重み付けを行っている場合である。

図8Fは、図8Dに示す違った形の補正関数を2つ接続したようなもので前半部、前半と中央部の間、中央部、中央部と後半部の間、後半部にそれぞれ重みを付けて、さらに各重み付けを異なったものになっている。

図8Gは、図8Eに示す違った形の補正関数を2つ接続したようなもので前半部、前半と中央部の間、中央部、中央部と後半部の間、後半部にそれぞれ重みを付けて、さらに各重み付けを異なったものになっている。

図8Hは、図8C及び図8Dに示す組合関数で、図8Iは、図8Dと図8Bの組合関数を示している。

図9は、一般的な時間補正関数の様子を示したもので、開始点、変化点、終点の座標をそれぞれ、 $P_0(t_s, s_3)$, $P_1(t_1, s_3)$, $\dots$, $P_e(t_e, s_0)$ としている。

ここで、座標のy成分は、重み付けを表しているので、ここでは、便宜上最大値を100最小値を0とし、0～100の間の値をとるものとし、x座標は、位置情報で、後述する図41～図43に示す「開始終了位置情報」のデータと同じディメンジョンの値、又は開始終了点間の区間に基づく開始点からの割合で、0～100の間で設定して位置を示している。

(再生ユニットの意味と接続関係、判定処理)

上記で説明したが、所定の特徴抽出処理による特徴データからその再生ユニット(PU)における意味設定することができる。

ここで、図10に示すようなビデオデータの構造について説明する。

ある1つのプログラム(番組)kを想定すると、幾つかのシーンm、 $m+1$ 、 $\dots$ に分類することができ、シーンは幾つかのショットに分類することができる。

そして、セグメント(ショット)を構成するのは、1つ1つのフレームとなる。

シーンの切れ目(区切り)は、シーンチェンジになる。

セグメント(又はショット、又は画像セグメント。以下同様)は、シーン毎に、そのシーンに応じた類似画像のまとまり、又は、類似した画像(映像)特性のまとまりとすることができる。

セグメントやシーンなどは、そのプログラム(番組)の中において、固有な意味の概念を持っている。

そして、各々の意味を持ったセグメント、シーンは、幾つかまとまってその番組を構成しているビデオ構造と捉えることができる。

例えば、野球の番組を想定した場合に、打者の画面が続いていたとすると、打者の類似画像が検出され、その類似特性セグメントに分類できる。そのセグメントは、「打者の画像」という意味(意味の概念)を有することになる。

また、投手の投球する画面が続いていたら、投手の類似画面が検出され、その類似特性に応じてセグメントに分類できる。そのセグメントは、「投手の画像」という意味(意味の概念)を持つことになる。

ここで、投手が投球して、打者が打撃し、その打者が走塁するような場合を想定した場合に、「投手の画像シーン」、「打者の画像シーン」、「打者の走塁の画像シーン」という、各々意味を持った画像シーンのつながりを捉えることができる。

所定プログラム(番組)において、上述したPU毎に、画像特徴データ、音声特徴データが処理され、それら特徴データに応じてそのPUの意味を設定することを想定する。例えば、ニュース番組を想定した場合に、キャスター(アナウンサー)が最初にニュースの項目を読み上げるシーン(ニュース番組のヘッドライン)の場合に、そのシーン(画像)の特徴として、人物特徴が1人～2人、テロップ(tlp特徴)、音声特徴の属性が話者音声、さらに、ニュース番組としたときに、そのニュース番組の中で、ニュースを読み上げるシーンは幾つか存在するので、そのニュース読み上げシーンと類似するシーンは幾つか存在することになり、類似画像特徴すなわち、特定のシーンIDは出現頻度が高くなる。

このように、規則1処理でも説明したように人物特徴、音声特徴、テロップ特徴、類似画像特徴、その他所定の特徴データに応じて、そのPUの意味を設定することができる。

例えば、上述した野球の番組の例のように、所定の意味を持つPUの接続関係が想定される。すなわち、所定の特徴データ又は特性データを持つPU間の所定の接続とすることができる。

上述した所定の意味を持つ、すなわち所定の意味が設定されたPUの接続関係を図11に示す。

図11において、あるプログラム(番組)で所定の意味a～意味dが設定されており、ある区間PU(n)～PU(n+2)で、接続関係は、PU(n)の意味a、PU(n+1)の意味b、PU(n+2)の意味cが最も自然なつながりとなることを示している。

すなわち、この図11に示した例は相撲の場合を想定しており、意味a「取組み紹介シーン」の後は意味b「立会いシーン」が続くことが一番妥当性であり合理的であり、意味b「立会いシーン」の後は意味c「取組みシーン」が続くことが一番妥当性であり合理的である。

そして、接続関係として、abcという意味を定義した文字系列とすることができ、このabcの系列がキーフレームとすれば、あるプログラム(番組)の中でabcを探して、探した所定区間の最初と最後、又は、その近傍などを所定設定点として設定処理を行うことができる。

他の例として、例えば、番組ジャンルが野球の場合は、ある区間で再生ユニットが、各々、「投球」、「打った」、「意味なし」、「得点」という場合に、「意味なし」を除いて、3つの意味、「投球」、「打った」、「得点」を持つと判定されたPUを1つにまとめて、「投球、打った、得点」という所定PUのかたまりを想定することができる。

ここで、「意味なし」のPUは、意味がないと判定されたことにより含めても問題なく、上記4つのPUを1つにまとめて「投球、打った、意味なし、得点」という所定PUのかたまりとすることができる。

ここで、「意味なし」を例に挙げたのは、上述の規則1の処理で所定の特徴データから所定の評価処理で、設定した幾つかの意味の中から、所定の意味付け処理を行う、すなわち、複数の意味から所定の信号処理に基づいて確からしい意味付けが行えない場合も想定できるからである。

「意味なし」の代わりに、「どのような意味でも良い」とすることができる。これは、上記した@と同様の処理である。

あるニュース番組の場合で、aabbという接続、すなわち、「アナウンサーシーン」、「アナウンサーシーン」、「現場シーン」、「現場シーン」という接続が、妥当で合理的で

ある場合を図12Aに示す。

先に説明した相撲番組の場合を図12Bに示す。

図13A、図13Bは、上述の番組ジャンルがニュース番組の場合で、図13Aに示すように、参照パターン(参照文字系列)を上記で説明した「aabb」として、図13Bに示すように、例に挙げる所定の番組記録区間の中で「aabb」の区間を探していき、区間A1、区間A2が「aabb」に一致して、検索できたことを示している。

そして、図13Bに示すように、例えば、探すことができた「aabb」区間の最初の位置p1、p3、最後の位置p2、p4を所定設定位置として設定し、後で説明するプレイリストのチャプタデータ(位置情報データ)として所定の処理を行う。例えば、要約再生モードの場合には、上記設定位置p1～p2、p3～p4を再生するように再生制御処理を行う。

チャプタ設定など、所定時点設定(所定位置設定)処理の場合には、p1、p2、p3、p4の各時点、又は、それら各点の所定の近傍の位置をその設定位置として所定の処理を行う。

このように、所定の特徴データから所定のPUの意味を持つと判定し、その所定の意味を持つとそのPUに設定し、それら意味が判定して設定されたPUから意味の接続関係を想定して、所定の意味に応じた所定数のPUの接続や所定数のPUの集合を想定して処理を行うことができる。

図6Bに示した規則2処理の記述方法の一例では、キーフレーム(重要フレーム)と想定し、検索したい文字系列を(aabb)のようにし、その後、重み付け係数として100を設定している。その後のPs(ts, s4), P1(t1, s4), Pe(te, s3)は、先に説明した時間補正関数であって、この例の場合は、図14に示すように、番組の後半部で徐々に重要度が減少するような関数となっている。この図14に示すような時間補正関数の場合には、番組の前半部を重点的に視聴したいような場合に適している。

ここで、図6Bに示した規則2処理の記述においては、記述するデータの出現パターン(意味)の表現方法の一例で、意味として、a, b, c・・・などとしたが、その否定としてA, B, C, …、また、ワイルドカードとして、*などを用いることもできる。この図6Bに示した規則2処理の記述において、ニュース番組のジャンルの場合の一として、

例えば、(Abb)とした場合、には、Aは、「アナウンサーのシーン」以外、bは、「現場のシーン」ということになり、「アナウンサーのシーン」以外に「現場のシーン」が2つ続く場合を検出することになる。

ここで、評価値の演算方法の一例として、以下のような処理である。例えば、再生ユニット群を(abc)として、上記(1)式により、a、b、cの各検出割合(value)と重み付け係数が以下の表5に示すような場合がある。

表5

意味	v a l u e	重み (重要度)	備考
a	1 . 0	1 0 0	$1 . 0 \times 1 0 0 = 1 0 0$
b	0 . 8	8 0	$0 . 8 \times 8 0 = 6 4$
c	0 . 8	8 0	$0 . 8 \times 8 0 = 6 4$

$$\begin{aligned}
 \text{評価値} &= 100 \times (100 + 64 + 64) / (\text{重みの総和}) \\
 &= 100 \times 228 / (100 + 80 + 80) \\
 &= 100 \times 228 / 260 \\
 &= 88
 \end{aligned}$$

ここで、100を掛けたのは、便宜上、割合(%)を考慮したためであるが、上述したように、評価値のスケールは、所定の評価処理が行えて、かつ所定の演算処理上問題なければよいので、オーバーフローなど処理上問題なければ、割合を考慮しなくてもよい。

(規則2処理における再生ユニット群の他の例)

ここで、規則処理2における、ある「意味」の再生ユニットが複数接続した再生ユニット群を1つの意味群ユニットとし、意味群ユニットが複数接続する場合を説明する。

上記規則1処理では、1つの再生ユニットのみの例を挙げて説明した。それは、特徴データから、検出する「意味」に最も確からしいであろう再生ユニットを見つけるためであった。

これをさらに発展させて、再生ユニット群、すなわち、この規則2処理で行った意味のつながりの再生ユニットを1つのかたまりとして、そのかたまり同士を接続した区間を検出するようにすることができる。

例えば、上記(aabb)をGa1とし、(Ga1Ga1)のような接続とすることができる。この場合に、Ga1の評価値について規則1と類似した処理を行う。この場合の評価値の演算方法として、例えば、各意味の再生ユニットの評価値の和の平均や、各意味の再生ユニットの評価値の積の平均などを求めることができる。

例えば、aの評価値を80、bの評価値を60とした場合に、Ga1の評価値は、加算の場合は、

$$(80+80+60+60)/4=70$$

で、70を評価値とすることができる。

規則3処理の場合 通常は、図15Aに示すように、規則2処理まででよいが、複数の番組対して特徴データを設けた場合に、例えば、番組毎に時間的重み付け処理を行う場合には、さらに規則処理として、図15の(B)に示すように、規則3処理を設ける。

その一例として、ニュース番組(news)とスポーツ番組(sports)に対して重み付けと、時間補正を行う場合の例を図6Cに示す。

図6Cに示す例では、ニュース番組は、100%の重み付けを行い、時間補正関数として開始点Ps(ts, s4)、変化点P1(t1, s4)、終点Pe(te, s3)とする補正を行い、スポーツ番組に対しては、70%の重み付けを行い、時間補正関数として開始点Ps(ts, s4)、変化点P1(t1, s4)、終点Pe(te, s3)とする補正を行う。

図3で説明した処理内容を図16を参照してさらに説明する。

図16Aに示すような、規則1処理により、各種所定の特徴データに基づいて、各シーンは幾つかの意味付け処理が行われる。

ここで、規則2によって意味付けされた各シーンには、図16Bに示すように評価値が所定の処理により設定される。

例えば、要約再生モードの場合では、ユーザの所望する時間t1で再生する場合に、上記評価値の一番高いシーン(画像)から選択していき、できるだけt1に近くなるように評価値の高いシーンから選択して、その選択した区間を再生するように、その位置情報を設定する。

設定した位置情報は所定のデータメモリに記憶し、再生制御を行う際に、位置情報

を読み出して、所定区間の再生を行っていく。

そして、各区間を順次再生する(スキップ再生)することで、所定の要約再生(ダイジェスト再生)を行う。

図16Cに示す例では、全記録時間を例えば60分とし、要約再生を15分で行いたいと仮定して、評価値が70以上のPUを選択して、15分にやや満たない場合に、評価値60の $PU_n + 8$ の区間を選択して、所望の再生時間15分にできるだけ近くなるように処理を行っている。

このように評価値の大きい所定PU区間を選択していき、所定の再生時間にできるだけ近くなるように、PU区間を選択していく。

所望の再生時間 T_m に対して所定の許容範囲 t_c 内に再生時間 T があるように、

$$T_m - t_c < T < T_m + t_c$$

となるように、評価値に基づいて所定のPU区間を選択する。

また、図16Dに示すように、例えば、意味付けされた評価値の高い区間の最初(又はその近傍)、評価値の高い区間の最後(又はその近傍)に所定位置(チャプタ)を設定することで、その区間の編集処理をしたり、スキップ再生の一時停止処理、繰り返し再生処理など、所定の操作を行うことに利用できる。

(2) ブロック構成例

ここでは、簡単のため、記録する画像音声データは、放送番組のデータとし、MPEG (Moving Picture Export Group) による所定の帯域圧縮処理が行われるものとする。なお、その他の帯域圧縮信号処理としてウェーブレット変換、フラクタル解析信号処理その他などを用いるようにしてもよい。例えば、下記の説明で画像データのDCT係数は、ウェーブレット変換の場合には多重解像度解析におけるか解析係数などに相当し同様の信号処理を行うことができる。

2. 1 ブロック構成例1

本発明を適用した記録再生装置30の全体ブロック構成例を図17に示す。

ここでは、簡単のためテレビ放送を受信して、受信した放送番組を記録する例を挙げて説明する。

2. 1. 1 記録信号処理系

この記録再生装置30では、受信アンテナ系1と受信系2により所定の放送番組が受信され、音声信号は音声A/D変換処理系3で所定のサンプリング周波数、所定の量子化ビット数で所定のA/D変換信号処理が行われ、その後音声エンコーダ処理系4に入力される。

音声エンコーダ処理系4では、例えばMPEGオーディオやAC3オーディオ(ドルビーAC3、又はAudio Code number 3)などの所定の帯域圧縮方式で信号処理が行われる。

同様に、映像信号は映像A/D変換処理系8で所定のサンプリング周波数、所定の量子化ビット数で所定のA/D変換信号処理が行われ、その後、画像エンコーダ処理系9に入力される。

画像エンコーダ処理系9は、MPEGビデオやウェーブレット変換などの所定の帯域圧縮方式で信号処理が行われる。

音声エンコーダ処理系4及び画像エンコーダ処理系9で処理された音声データ及び画像データは、多重化処理系5を介して記録処理系6に入力される。

音声信号の特徴抽出を行うため、音声エンコーダ処理系4に入力する信号の一部又は上記所定エンコーダ信号処理における信号処理過程の途中の信号の一部は特徴抽出処理系10に入力される。

図17に示す記録再生装置30では、音声エンコーダ処理系4に入力される信号の一部として、音声エンコーダ処理系4から特徴抽出処理系10に信号が入力されているが、音声エンコーダ処理系4に入力されるとともに特徴抽出処理系10に入力するようにしてもよい。

同様に映像(画像)信号の特徴抽出を行うため、映像エンコーダ処理系9に入力される信号の一部又は上記所定エンコーダ信号処理における信号処理過程の途中の信号の一部が特徴抽出処理系10に入力される。

この図17に示した記録再生装置30では、映像エンコーダ処理系9に入力される信号の一部として、映像エンコーダ処理系9から特徴抽出処理系10に信号が入力されているが、映像エンコーダ処理系9に入力されるとともに特徴抽出処理系10に入力されるようにしてもよい。

記録モードにおいて所定区間毎に逐次特徴データは検出され、所定のエンコーダ処理がなされた画像音声データとともに所定の記録媒体7の所定の記録領域に記録される。

上記特徴データから所定の要約再生(ダイジェスト再生)を行うためのプレイリストデータの生成をプレイリスト処理(9)又はチャプタデータの生成を行うダイジェスト再生チャプタ処理(11)をプレイリスト・チャプタ生成処理系19で所定の信号処理を行う。

ここで、プレイリストデータ、チャプタデータの生成は、以下のような信号処理プロセス(処理a又は処理b)で行うことができる。

(処理a) 特徴データを所定メモリ系又はシステムコントローラ系の所定メモリ領域に所定データ量蓄積した後、所定のプレイリストデータの生成処理、所定のチャプタデータの生成処理を行う。

(処理b) 画像音声データを記録する記録媒体7に所定の特徴抽出処理を行う毎に逐次特徴データを記録し、所定データ量記録した後、そのデータを再生して、所定プレイリストデータ、所定チャプタデータの生成を行う。

(処理a)の場合、例えば、所定時間長 t の放送番組を記録すると、その時間長 t の記録が終了したら、その放送番組におけるすべての所定特徴データが集積されるので、この時点で、時間長 t の中で所定の要約再生時間 t_d に対応するキーフレームがどこになるかを定めるプレイリストデータ生成処理を行うことができる。すなわち、この時間長 t に処理される特徴データをメモリ系、又はシステムコントローラ系の所定メモリ領域に蓄積(記憶又は記録)しておくことになる。

(処理b)の場合は、上記(処理a)の場合と同様で所定時間長 t 記録した後、所定時間 t 記録終了したことを検出(検知)して、所定の特徴データを再生して所定要約再生時間 t_d に応じたプレイリストデータ生成処理を開始することになる。

プレイリストデータ生成処理が終了したら、所定の要約再生の動作を行う準備ができたことになり、このプレイリストデータを用いて所定の要約再生(ダイジェスト再生)が行える。

上記所定の特徴データは、プレイリストデータがすでに生成されているので、もうプレイリストデータを生成しないという場合には消去するように信号処理を行うことがある。

が、プレイリストデータを修正するなど、データの生成を再度行う場合には、特徴データはそのまま記録して残してもよい。

上記特徴データは、システムコントローラ系20を介して、所定区間の特徴データの蓄積の後、プレイリスト・チャプタ生成処理系19で所定の要約再生(ダイジェスト再生)用プレイリストデータを生成する。

上記生成されたプレイリストデータは、記録処理系6において所定の記録処理がなされた後、記録媒体7の所定の記録領域に記録される。

ここで、プレイリストデータは、所定の記録された区間をスキップ再生するための、所定再生区間毎の再生開始点情報と再生終了点情報の対となるデータから構成され、例えば、所定区間毎の再生開始フレーム番号と再生終了フレーム番号のデータ対などからなる。

プレイリストデータは、その記録したプログラムにおける、所定の必要な区間をスキップ再生することで要約再生(ダイジェスト再生)を行う処理のために使用するので、上記のようにフレームデータの他に、タイムコードデータやMPEGにおけるPTS(Presentation Time Stamp)、DTS(Decode Time Stamp)などのタイムスタンプデータでもよい。

プレイリストデータは、上記のように放送番組のような画像音声情報データを記録する記録モード時で所定プログラム記録終了後に所定の生成処理を行う他に、後で説明する再生モードにおいて、特徴データを用いて所定の処理を行うようにしてもよい。

図17において、例えば、既にMPEGなど所定のエンコード処理がなされた画像、音声データを記録する場合には、音声エンコーダ処理系4、画像エンコード処理系9でエンコード処理を行う必要はなく、直接、多重化処理系5に入力し、記録処理系6で記録処理を行い記録媒体に記録することができる。

ここで、直接デジタル画像、音声データが入力して記録されるか、受信系2によりアナログ信号が入力し所定のエンコード処理の後に記録されるかは、システムコントローラ系20で検出することができ、このように入力系統違いに応じて、上記所定の画像、音声特徴データ抽出処理を記録モードのときに自動的に行うか、記録終了後に行

うかを定めるようにする、又はデジタル画像、音声データが入力する場合には、所定のエンコード処理系をデータが通らないことから所定のデータ構造解析処理を行う必要がないので、記録終了後に行うようにすることができる。

記録モードにおいて、上記アナログ入力系かデジタル入力系は、ユーザ入力I/F系21を介してユーザの所定操作によって設定することもできる。

また、図17で、音声エンコーダ処理系4又は音声A/D変換処理系3、映像エンコーダ処理系9又は画像A/D変換処理系8からの信号と、所定エンコード処理されたデジタル画像、音声データを直接システムコントローラ系20で検出することで自動的に検出することもできる。

所定エンコードされたデジタルデータが検出され、音声エンコーダ処理系4又は音声A/D変換処理系3、映像エンコーダ処理系9又は画像A/D変換処理系8でデータが検出されない場合は、所定エンコード処理されたデジタル画像、音声データが入力していると判定できる。

所定のエンコードされたデジタルデータが検出さないで、音声エンコーダ処理系4又は音声A/D変換処理系3、映像エンコーダ処理系9又は画像A/D変換処理系8からのデータがシステムコントローラ系20で検出される場合は、アナログ入力と判定できる。

アナログ入力とエンコード処理されたデジタルデータが両方検出される場合は、例えば、受信系2からのアナログ入力信号を初期設定(デフォルト設定)として所定の記録処理を行うようにしてもよい。

上述の特徴抽出処理は、例えば画像のDCTデータなどを用いるので、所定のエンコード処理がなされる場合には、通常の記録処理のために行うDCT処理を特徴抽出処理として兼用することができる。音声の場合には、所定エンコード処理におけるサブバンド処理データを用いることを考えると、所定のエンコード処理がなされる場合には、通常の記録処理のために行うサブバンド処理を特徴抽出処理として兼用することができる。

上述のように、エンコード処理されたデジタルデータが直接入力する場合は、エンコード処理を行う必要がないので、このデータを解析してDCTなどデータを取り出す

ことが必要になり、処理の負荷が生じることになる。

そこで、必要に応じて記録終了後に特徴抽出処理を行うようにする。その他、記録終了後に、特徴抽出処理を行う場合として、上記アナログ入力の場合でも、信号処理系の負荷の具合に応じて、所定の記録が終了したら自動的に行うようにしてもよい。

例えば、図20に示すように、特徴抽出処理はソフトウェア処理で行うこともできるので、システムコントローラ系の性能によっては記録モードの各所定信号処理と同時に行うことができないので、所定の記録処理が終了してから行うようにする。また、システムコントローラ系20はCPU、DSP(デジタルシグナルプロセッサ)、その他各種プロセッサなどで構成することができるが、性能が高いほど高価なので上述のように処理能力に応じて、特徴抽出処理を記録処理と同時に行うか、終了後に行うかを定めるようにしてもよい。

上述の特徴抽出処理を行う所定記録モード終了後としては、例えば、所定のタイマ記録動作終了後や、通常、ユーザがその装置を動作させていないと想定できる夜中に、所定の特徴抽出処理を行うことがある。このような場合に、例えば、装置が動作している時刻をシステムコントローラ系20内の所定メモリ手段により記憶して、所定の学習処理により、適宜、特徴抽出処理する時刻を自動設定したりすればよい。

また、記録再生など通常の動作させていない時間がシステムコントローラ系20で検出される場合には、その動作させていない間に上記所定の特徴抽出処理を行うことができる。その場合に所定のデータすべてが処理されない場合も想定されるが、処理途中の場所をシステムコントローラ系20内の所定メモリ手段に記憶しておき、装置が記録再生など通常動作していないことを検出して、処理できる時間があると判定されたら、途中の続きから所定の信号処理を行うようにすればよい。

2. 1. 2 再生側処理

(通常再生モード動作)

次に、図17に示した記録再生装置30における再生信号処理について説明する。

まず、通常再生モードの動作について説明する。

ユーザ入力I/F系21により、通常再生モードになると、記録媒体7から所定の画像

音声データ、特徴データなどが記録されている所定のデータが再生され再生処理系12において所定の再生処理が行われる。

再生された所定のデータは、再生データ分離処理系13において所定のデータに分離処理され、音声データは音声デコード処理系14に入力され、記録時に帯域圧縮信号処理された信号処理方式に対応する所定のデコード処理がなされ、その後、音声D/A処理系15に入力されてD/A変換処理された後、音声信号として出力される。

また、所定の分類処理された画像(映像)データは、映像デコード処理系16において記録時に帯域圧縮信号処理された信号処理方式に対応する所定のデコード処理がなされた後、映像D/A処理系17に入力されてD/A変換処理が行われ、映像信号として出力される。

(要約再生(ダイジェスト再生)モード)

要約再生モードを考える場合に、画像音声データとともに特徴データ、プレイリストデータが記録媒体に記録されているかどうかで信号処理方法が異なる。

特徴データとプレイリストデータが記録媒体に記録されているか否かは図18のように整理することができる。

まず、図18A、図18Bの場合に相当するプレイリストデータ(プレイリストデータファイル)、チャプタデータが再生できる場合、すなわち、プレイリストデータ、チャプタデータが所定の記録媒体(データ記録媒体)に記録されており、要約再生モード時に再生、又は、チャプタ表示モード時に所定チャプタ画像がサムネール表示できる場合について説明する。

すなわち、ユーザが要約再生モード又は所定チャプタモードを選択した場合について説明する。

ユーザの入力操作によりユーザ入力I/F系21を介して、システムコントローラ系20に所定の要約再生(ダイジェスト再生)モードの動作を行うコマンドが入力された場合に、再生データ分離処理系13で所定データ分離し特徴データ、また、パラメータデータやプレイリストデータ、チャプタデータなどが記録されている場合には、それぞれ分離された所定の特徴データ、所定のパラメータデータ、所定のプレイリストデータ、

チャプタデータなどがシステムコントローラ系20に入力される。

再生データ分離処理系13で特徴データ、パラメータデータ、プレイリストデータ、チャプタデータが分離できない場合には、上記のそれぞれのデータはシステムコントローラ系20に入力されないで、再生データ分類処理系13とシステムコントローラ系20とにより特徴データ、プレイリストデータ、所定チャプタデータ、パラメータデータなどが所定記録媒体7に記録されているか否かの判定処理を行う。

プレイリストデータは、所定の要約再生を行うために、所定の幾つかの再生区間の再生開始情報データと再生終了情報データから構成されている。

チャプタデータは所定特徴区間の先頭又はその近傍、又はその所定特徴区間の最後又はその近傍、その特徴区間に接続された特徴区間以外の区間の先頭又はその近傍、又はその特徴区間以外の区間の最後又はその近傍の位置情報から構成される。

システムコントローラ系20では、再生検出されたプレイリストデータのスキップ再生開始データ情報、スキップ再生終了データ情報に応じてスキップ再生を行うことで要約再生(ダイジェスト再生)を行う。

また、所定チャプタデータによりチャプタ点又はその近傍における画像を所定サムネール画像として所定表示処理を表示処理系27で行い、所定の画像表示を行う。

次に、図18C、図18Dに示す場合に相当するプレイリストデータ(プレイリストデータファイル)、チャプタデータが再生できない場合、すなわち、プレイリストデータ、チャプタデータが記録媒体又は記憶媒体に記録(記憶)されておらず、要約再生モード時に再生できない場合、所定チャプタモード時にサムネール時点を所定サムネール表示、チャプタ再生など一連のチャプタ関連処理できない場合について説明する。

上述した放送番組などを受信した画像音声データでなく、例えば、記録媒体25をDVDソフトとし記録媒体処理系26、再生処理系12により再生する場合など他の記録媒体からの画像音声データを再生する場合や、特徴抽出していない画像音声データを再生する場合などは、ここで説明する処理に該当する。

プレイリストデータ又はチャプタデータが生成されておらず再生検出できない場合や、再生検出されたプレイリストデータ、チャプタデータを生成し直したい場合は、再

生検出された所定の特徴データとパラメータデータとから要約再生用プレイリストデータ、及び所定チャプタ関連モード用チャプタデータを生成することができる。

図26Cに示す場合、すなわち、記録時に特徴抽出処理が行われており、特徴データが再生できる場合には、図17に示した記録再生装置30における再生処理系12又は再生データ分離処理系13からプレイリストデータ又は上記所定の特徴データがプレイリスト・チャプタ生成処理系19に入力され、所定のプレイリストデータ又は所定のチャプタデータが生成される。

ここで説明する動作の場合に、ユーザが要約再生モードのコマンドを行った場合に、図19に示すようなプレイリストデータがないことを示す所定の表示を表示処理系27で行うようにしてもよい。

生成されたプレイリストデータは、システムコントローラ系20に入力される。システムコントローラ系20は、ユーザ入力による所定の要約再生時間に応じて、プレイリストデータに基づいた所定の再生区間を順次再生(スキップ再生)するように再生制御系18をコントロールしその制御により記録媒体7を再生制御する。

また、生成されたチャプタデータはシステムコントローラ系20に入力される。システムコントローラ系20は、ユーザ入力による所定のチャプタ関連動作モードに応じて、上記チャプタデータに基づいた所定のチャプタ時点の画像サムネール表示、チャプタ点のカットや接続などの編集処理、ユーザ選択したチャプタ点のスキップ再生など、所定チャプタ関連動作が行えるように再生制御系18をコントロールし、その制御により記録媒体7の再生制御を行ったり、システムコントローラ系20を介した表示処理系27の制御などを行う。

上述したように、例えば、DVDなど外部記録媒体を記録媒体25として要約再生する場合にも、上述の場合と同様の信号処理により行うことができ、再生制御系18により記録媒体処理系26を制御し、上述したような所定の要約再生処理を行う。

また、チャプタデータを用いた編集処理(編集操作)、所定のチャプタ点間(又はその近傍)のスキップ再生、チャプタ点(又はその近傍)のサムネール画像表示など、一連の所定チャプタ関連動作を上記と同様の信号処理により行うことができ、再生制御系18により記録媒体処理系26を制御し、上記したような所定信号処理を行う。

さらに、図16Dに示す特徴データが再生できない場合について説明する。

上述の例では特徴データからプレイリストデータ、チャプタデータを生成する場合について説明したが、例えば、他のユーザが記録した外部記録媒体25を記録媒体A26にコピーしたような場合には、特徴データが再生できない場合がある。

記録媒体7に放送番組などの画像音声データは記録されているが、特徴データが記録されておらず、再生できない場合について説明する。

ここで説明する動作の場合に、ユーザが要約再生モード又は上記所定チャプタ関連動作モードのコマンドを行った場合に、図19に示すような特徴データがないことを示す所定の表示を表示処理系27で行うようにしてもよい。

この場合の要約再生モードで記録媒体A7から画像音声データを再生する場合は、再生処理系12で再生されたデータは再生データ分離処理系13に入力され、分離された記録時に所定の帯域圧縮方式で処理されている画像データと音声データは特徴抽出処理系10に入力され、画像特性データであるDCT DC係数、AC係数、動きベクトル(モーションベクトル)など、音声特性データである音声パワー検出その他など各種所定の特性データ検出処理などが行われる。

特徴抽出処理系10では、さらに、上述の各種画像音声特性データと所定のパラメータデータとにより、所定のテロップ特徴データ(テロップ区間判定データ)、人物特徴データその他の画像特徴データ(画像特徴区間判定データ)、及び話者音声特徴データ(話者音声判定データ)、拍手歓声特徴データ(拍手歓声判定データ)その他の音声特徴データ(音声特徴区間判定データ)の各種特徴抽出処理が行われる。

上記各種の画像特徴データ及び音声特徴データは、システムコントローラ系20に入力され、所定番組あるいは、所定の画像音声区間のすべてについて所定の特徴抽出処理が終了したら特徴抽出処理が終了したと判定される。

ここで、特徴抽出処理が終了した場合には、システムコントローラ系20から所定の信号処理が終了したことを示す信号が表示処理系27に入力され、例えば図19に示すような所定の表示を行うようにしてもよい。

次に、特徴データから所定のプレイリストデータ、チャプタデータを生成する処理について説明する。

上述の特徴データは、所定の特徴抽出区間毎にメモリ系11に蓄えられ、上記すべての所定特徴データの処理が終了したらプレイリスト・チャプタ生成処理系19に入力され、所定のプレイリストデータ又はチャプタデータが生成される。

ここで、特徴抽出処理系10から直接プレイリスト・チャプタ生成処理系19に、所定区間の特徴抽出処理データを逐次入力するようにしてもよく、上述したように所定のすべての区間や所定放送番組の特徴抽出処理が終了したら、システムコントローラ系20からの所定の信号により、プレイリスト・チャプタ生成処理系19で上述の如き所定のプレイリストデータ又はチャプタデータ生成処理を行うようにしてもよい。

また、特徴抽出処理系からの上記処理された特徴データは、システムコントローラ系20を介してプレイリスト・チャプタ生成処理系19に入力するように信号処理を行うようにしてもよい。

プレイリスト・チャプタ生成処理系19で、所定のプレイリストデータ又はチャプタデータが生成されると、所定の処理が終了したことを示す信号がシステムコントローラ系20に入力され、所望の要約時間に応じた要約再生又は所定のチャプタデータを用いた所定チャプタ関連動作を行うことができる。

この場合に、図19に示すように、プレイリストデータ又はチャプタデータが生成できたことを示す所定の表示を行い、あるいは、要約再生モード、チャプタ関連の所定動作モードであることなどの表示を表示処理系27に行うようにする。

ユーザが要約再生を行う場合に、例えば記録した放送番組が1時間であったとし、それを30分で再生したいのか、20分で再生したいのかなど、ユーザの所望する要約再生時間は分からないので、記録した放送番組など画像音声データの特徴抽出した全区間の全時間長に応じて、予め幾つかの要約時間に対応するプレイリストデータを生成することを考えることができる。

例えば、特徴抽出する放送番組の記録時間が1時間であったら、40分、30分、20分の要約再生を行う、それぞれのプレイリストデータを生成する。このようにプレイリストデータを生成することで、リモコン22などのユーザ入力で要約時間が選択された場合に、すぐに所定の要約時間に対応した要約再生動作を行うことができる。

記録媒体25を再生する場合については、上記記録媒体A7を再生する場合と同様

で、記録媒体処理系26により記録媒体25を検出し、再生処理系12により再生信号処理し、再生データ分離処理系13において所定の画像音声データを分離する。その後の信号処理は上記の記録媒体7の場合と同様なので省略する。

ここで、上述した一連の処理を実行する制御プログラムは、専用のハードウェアに組み込まれているコンピュータ又は各種のプログラムをインストールすることで、各種の機能を実行させることが可能な、例えば汎用のパーソナルコンピュータなどに記録媒体からインストールされる。

この記録媒体は、制御プログラムが記録されているハードディスクだけではなく、コンピュータとは別に、ユーザにプログラムを提供するために配布される、プログラムが記録されている磁気ディスク、光ディスク、光磁気ディスク若しくは半導体メモリなどよりなるパッケージメディアにより構成される。

2. 2 ブロック構成例2

ここで、図17に示した記録再生装置30の他の例として図20に示す例を挙げて説明する。

放送番組を記録再生処理する信号処理の過程は、図17に示した記録再生装置30と同様なので、信号処理の異なる部分について説明する。

2. 2. 1 記録側信号処理

この図20に示す記録再生装置30Aにおいて上述の記録再生装置30と異なるのは、記録モードにおいて特徴抽出処理を行う一連の信号処理をシステムコントローラ系20においてソフトウェア的に行うことである。

この記録再生装置30Aでは、ネットワーク系24により、所定のソフトウェアをダウンロードして、説明するようなソフトウェア処理による、特徴抽出処理、プレイリスト処理(チャプタ生成処理(再生区間、再生時点位置情報生成処理))などが行われる。

ソフトウェアによるダウンロードでは、例えば、本発明の処理が搭載されていない装置が、最初からあった場合に、時間をおいて後からソフト的に本発明を適用できる利点があり、例えば、製造販売などの時間的に間に合わない場合には、設計、製造側においては、本発明を適用しない簡単な構成のシステムと、後述する本発明を適用したシステムの両方のシステムをユーザに提供できる。

ユーザ側では、本発明を適用しない簡単な構成のシステムを購入した後で、ソフト的な処理で、本発明を適用できるので、後から機能を追加できるなどの利点がある。

また、処理系を修正、改良する場合にもソフトをダウンロードしてアップグレードすることで、対応できる利点がある。

本発明をソフトウェアのダウンロードで装備する場合は、ユーザは所定の操作系(リモコン22など)で、所定のインターネットサイトにネットワーク系24を介して接続し、所定の操作系による操作で本発明のソフトウェアをダウンロードする。

ダウンロードされた本発明のソフトウェアは、システムコントローラ系20で、所定の解凍処理、インストール処理などが行われ、後で説明する、特徴抽出処理、プレイリスト処理、チャプタ処理ほか、本発明の所定の処理機能が装備される。

システムコントローラ系20として、所定性能を備えたマイクロプロセッサ(MPU、又はCPU)を用いることで上述した所定の特徴抽出処理を所定の記録処理と同時に行うことができる。

上述したメモリ系11も、このシステムコントローラ系20内に備えられた所定のデータ記憶メモリを用いることができる。

ここで、上述したように所定の記録処理として、所定の画像音声の帯域圧縮を行う場合に、上述のような所定の性能を備えたMPU又はCPU、又はDSP(デジタル・シグナルプロセッサ)を用いることができ、この帯域圧縮処理を行っている同じMPU又はCPU、又はDSPで上記所定の特徴抽出処理、プレイリスト生成処理などを行うことができる。

2. 2. 2 再生側信号処理

この図20に示す記録再生装置30Aにおいて上述の記録再生装置30と異なるのは、再生モードにおいて、特徴データが検出できず特徴抽出処理を行う場合に、一連の信号処理をシステムコントローラ系20においてソフトウェア的に行うことである。

システムコントローラ系20として、所定性能を備えたマイクロプロセッサ(MPU、又はCPU)を用いることで上述した所定の特徴抽出処理を所定の記録処理と同時に行うことができる。

上述したメモリ系11も、このシステムコントローラ系20内に備えられた所定のデータ

記憶メモリを用いることができる。

(3) 特徴抽出処理

次に、音声系特徴抽出処理及び映像(画像)系特徴抽出処理の各信号処理について説明する。

3.1 音声系特徴抽出処理

音声系特徴抽出処理系では、図21に示すように、MPEGの画像音声ストリームデータがストリーム分離系100に入力され、分離された音声データは音声データデコード系101に入力され所定のデコード処理が行われる。

デコードされた音声データ(音声信号)はレベル処理系102、データカウンタ系103、データバッファ系104に各々入力され、レベル処理系102では、音声データの所定区間の平均パワー(又は平均レベル)Pavを演算するため、データの絶対値化処理を行い、データカウンタ系103で所定サンプルデータ数まで計測されるまで、音声データ積算処理系105で積算処理を行う。

ここで、平均パワーPavは、音声データの値(レベル)をAd(n)として以下の(32)式の演算により求めることができる。

[数2]

$$Pav = \sum_{n=1}^{nd} |Ad(n)| / Sm$$

・・・(32)

平均レベルを演算する所定区間として、例えば、約0.01sec(10msec)～1secが考えられ、例えば、サンプリング周波数FsをFs=48KHzとすると、480～48000サンプルの積算演算を行い、サンプル数Smで平均処理を行って平均レベル(平均パワー)Pavを求める。

音声データ積算処理系105から出力されたデータPavは、判定処理系106に入力され、しきい値設定系107で設定された所定しきい値Athと比較処理され無音判定処理が行われる。

ここで、しきい値設定系107における所定しきい値Athの設定において、Athは固

定値Ath0として設定することが考えられるが、固定値Ath0の他に、所定音声区間の平均レベルに応じた変動しきい値Athmを設定することも考えられる。

変動しきい値Athmとして、例えば、いま処理を考えている区間をnとし、それより前の区間(n-k)の平均レベルPav(n-k)を考え、次の(33)式のようにすることが考えられる。

[数3]

$$Athm = \sum_{k=1}^t Pav(n-k)/m \quad (t \leq m)$$

・・・(33)

例えば、t=2として、

$$Athm = (Pav(n-1) + Pav(n-2))/m \quad \cdots (34)$$

例えば、mは、おおよそ2～20くらいの範囲に設定する。

(その他の音声特徴抽出処理)

データバッファ系104に蓄積された所定音声データは、周波数解析処理系108に入力され、所定の周波数解析処理が行われる。

ここで、周波数解析処理としてFFT(高速フーリエ変換)などが考えられ、データバッファ系104からのデータの所定解析サンプルデータ数は、例えば、512、1024、2048、その他、など2のべき乗の所定サンプル数で所定の解析処理を行う。

周波数解析処理系108からの信号(データ)は、判定処理系109に入力され、所定の判定処理が行われる。

音楽(楽音)の判別処理は、所定周波数帯域のスペクトルピークの継続性から行うことができる。

例えば、特開2002-116784号公報などにはそれらの技術が開示されている。

話者音声の判定では、人の会話音声波形で息継ぎの区間があるので、波形に所定の急峻な対上がり、又は立下り区間が見られ、その所定立ち上がり、又は立下り区間を検出することで所定の信号処理を行うことができる。

この場合に、上記音楽(楽音)信号波形の場合は話者音声の場合に比べて、一般

的に波形の立ち上がり、又は立下り区間が現れる確率は小さいと考えられるので、この楽音(音楽)波形の特性(特徴)も考慮して、総合的に音声信号の属性判定を行うようにする。

上述のような、話者音声信号の波形特徴(波形特性)、音楽(楽音)信号の波形特徴(波形特性)の相違から音声信号の属性判定を行う場合に、波形における時間的な物理特性を検出することになるので、上述したような周波数解析を行ってから所定の判定信号処理を行う方法(周波数領域での信号解析、判定処理)の他に、ベースバンド領域で所定の判定処理を行う方法(時間領域での信号解析、判定処理)を用いることができる。

ここで、音声信号(音声データ)をデコード処理しないで、圧縮帯域のままで信号の属性解析を行う場合の音声系特徴抽出処理系の構成例を図22に示す。

図22に示す音声系特徴抽出処理系では、所定の帯域圧縮信号処理が施されたデータストリーム、例えば、MPEGなどの画像音声データがストリーム分離系100に入力されて画像データと音声データに分離され、音声データはストリームデータ解析系110に入力され、所定のサンプリング周波数、量子化ビット数その他などの信号解析処理が行われ、所定の音声データはサブバンド解析処理系111に入力される。

サブバンド解析処理系111で所定のサブバンド解析処理が行われ所定サブバンド帯域のデータは上記(32)式～(34)式で説明したのと同様の所定信号処理が行われる。

すなわち、音声データ積算処理系105に入力され、データカウント系103で所定のサンプリングデータ数が検出されるまで所定の積算処理が行われ、その後、しきい値設定系107で設定される所定しきい値に基づいて判定処理系106で所定の無音判定処理が行われる。

この無音判定処理では、音声データのスペクトルを考慮して、エネルギーが多く集まっている帯域で、サブバンド帯域としては大よそ3KHz以下の所定データ帯域を用いることができる。

また、周波数解析により楽音(音楽)、話者音声の判定処理が行えることを述べたが、サブバンド解析処理系111の処理により、この信号処理系で所定の周波数解析が

行われることと同様であるので、上述したような所定スペクトルピークの継続性判定処理を行うことで属性判定の信号処理を行うことができる。

この場合、スペクトルピークは、各所定サブバンド帯域の中の最大データ帯域であるので、FFT解析処理の場合と同様の信号処理を行うことができる。

3. 2 画像系特徴

次に映像(画像)系の特徴抽出処理について説明する。

映像系特徴抽出処理系では、図23に示すように、ストリーム分離系で所定の分離処理が行われた画像データは、ストリームデータ解析系200に入力され、レート検出、画素数検出その他など所定のデータ解析が行われ、DCT係数処理系201でDCTのDC係数検出、AC係数検出など所定のDCT演算処理(逆DCT演算処理)が行われ、このDCT係数処理系201の出力に基づいて、シーンチェンジ検出処理系202、色特徴検出処理系203、類似画像検出処理系204、人物検出処理系205及びテロップ検出判定処理系206における各種処理が行われ、動きベクトル処理系208では、所定の動きベクトル検出処理が行われる。

3. 2. 1 シーンチェンジ特徴

シーンチェンジ検出処理系202では、例えば、所定画面領域に分割しその領域毎にDCTのDC係数データのY(輝度データ)、Cb、Cr(色差データ)の平均値を演算してフレーム間差分演算又は、フィールド間差分演算をその領域毎に行い、所定しきい値と比較して、所定のシーンチェンジ検出を行う。

シーンチェンジがない場合は、各領域のフレーム間(又はフィールド)差分データは所定しきい値より小さく、シーンチェンジがあるとしきい値より差分データが大きくなる場合が検出できる。

ここで、画面分割の領域は、例えば、図24に示すように有効画面を16分割するような領域とする。

演算する画面分割の方法は図24の場合に限らず、分割数を多くすることも、少なくすることもできるが、少な過ぎるとシーンチェンジの検出精度が鈍感になり、分割数が多いと精度が鋭過ぎることが考えられるので、およそ256(16×16)以下の範囲の間で適当な所定の分割数を設定する。

3. 2. 2 色(カラー)特徴

色特徴検出処理系203では、DCTのDC係数の、所定領域におけるY、Cb、Crデータの平均値から色特徴を検出することができる。

所定領域としては、例えば、図25に示すような領域とすることができる。この図25では、有効画面を横方向に4分割し検出領域1～検出領域4、縦方向に4分割して検出領域5～検出領域8を設けている。各検出領域には領域IDが付され、各検出領域のデータは領域IDによって識別される。

ここで、場合に応じて横方向だけの検出領域1～4、又は縦方向だけの検出領域5～8を設けるようにする。

また、図25のような領域分割以外にも、 5×5 や、 6×6 といった碁盤状の分割方法とすることもできる。

例えば、放送番組で番組ジャンルとして「相撲」の場合を想定した場合に、図25の検出領域3領域で茶色が検出できる場合は「土俵のシーン」の確率が高いと想定できる。

この色特徴と例えば、音声の属性特徴と組み合わせると、「土俵のシーン」+「音声属性その他(又は話者音声)」から「取組みが開始するシーン」の確率が高い想定できるので、このようなシーン区間がキーフレーム区間と設定できる。

この場合には、取組み開始シーンでは観客の歓声などで音声レベルが大きくなったり、通常の状態とは異なる音声周波数帯域のデータが検出されることになるので、音声レベルや、所定の周波数領域データも特徴データとすることができる。

3. 2. 3 類似シーン(類似画像)特徴

類似画像検出処理系204では、類似シーン(類似画像、類似映像)毎に、所定のID(識別番号、又は識別記号)をその画像(シーン)に付与(付加)(又は、割り当て)する処理で、類似画像(シーン)には同一IDが付与(割り当て)処理される。例えば、特開2002-344872号公報にその技術が開示されている。

この付加(付与)する処理はその画像(シーン)又は画像(シーン)の位置情報(フレーム番号、PTS、記録時刻など)と一対一に対応するメモリにそのIDを記録するもので、その画像を表示又はその画像から再生するなどの動作を行う場合に、その画像(

シーン)の位置情報とIDは一对一に対応しており、画像(シーン)自体とその位置情報も言うまでもなく一对一に対応しているので、例えば、同一IDの画像を表示するなど類似画像分類や同一IDの画像シーンの時点をスキップ再生するなど、IDを利用する色々な所定動作を行うことができる。

この特徴データとしては、上述のシーンIDについて説明したように、検出頻度の1位、2位などの検出出現順位とすることができる。

また、図7に示すような、そのPU区間長に対するその出現順位1位、2位など、検出したIDの検出長の割合とすることもできる。

この特徴抽出処理は、例えば、画面を複数分割(例えば、25分割)して、その各分割した画面領域に対応する領域のDCTの平均DC係数を演算して、その演算した平均DC係数をベクトル成分として、所定ベクトル距離が所定しきい値より小さい所に対応する画像(シーン)を類似画像(類似シーン)とし、類似画像(類似シーン)には、同一の所定ID(シーンID)を割り当てる処理である。

例えば、IDとして、初期値を1とし、上記所定しきい値より小さい画像(シーン)が検出されない場合は、IDの最大値に1を加算したものを新たなIDとして、その画像(シーン)に割り当てる。

本発明における、この特徴データの利用方法として、図5を参照して説明したように、所定区間でのIDの出現頻度を演算して、頻度1位～2位の検出を行うなどの処理方法がある。

これは、例えばニュース番組を想定した場合に、アナウンサーシーンが頻繁に出現する場合や、相撲、野球など、類似シーンの出現が多く見込めるような番組ジャンルでは、有効な処理に用いることができる。すなわち、ニュース番組で出現頻度1位や2位では、出現頻度の高いと想定できるアナウンサーシーンが検出できる確率が高いと考えられる。

図26は、IDの出現頻度の演算方法を説明するための概要を示すもので、例えば、区間f1～f2、f3～f4、f5～f6、f7～f8の4区間で同一IDであるID1が検出されている。すなわち、この区間では、類似したシーンが出現している。

図26に示すように、所定区間同じIDが連続している区間を1つとして数え、そのよう

な区間が幾つあるかを演算する。

類似シーンが出現しなくなると同一IDではなくなるので、IDの連続性、不連続性の数を演算することで所定の頻度を算出することも考えられる。

3. 2. 4 人物特徴

人物検出処理系205では、図27に示すように画面の領域を分割し、各領域における所定の特定色を検出することで、人物が画面に現れているかを判定することができる。

図27に示した例では、有効画面を 2×2 に分割した領域1～4の4つの領域と、画面中央付近の領域5の5つの領域としている。

例えば、ニュース番組では、領域5にアナウンサーの顔が出現する場合の確率が高い。ことが考えられる。

また、フリップ又はテロップとアナウンサーシーンが出現する場合を想定すると、領域1又は領域2にアナウンサーの顔が出現する場合がある。その場合に領域2又は領域1にフリップ又はテロップが出現すると想定できる。

例えば、特定色として白人の肌色を想定すると、以下の条件式により特定色を検出することが実験から分かっている。

$$0.6 < Cb / Cr < 0.9 \sim 0.97 \quad \cdots (35)$$

$$(0 \leq Cb \leq 255, 0 \leq Cr \leq 255) \quad \cdots (36)$$

以下に説明するように、図27に示した領域における特定色の検出による方法と別の方法を用いることができる。

ここでは、簡単のため画面サイズを 720×480 とする。

(処理1)

色信号(Cb, Cr)からの検出条件(白人系肌色条件)DCT係数Cb、CrのDC成分においてマクロブロックは 16×16 で、x方向は $720 / 16 = 45$ で、(0～44)、y方向は $480 / 16 = 30$ で、(0～29)毎に以下の(3. 2. 3)式に示す判定条件でデータポイントを検出する。場合によっては、x、y方向各々 $1/2$ に圧縮してx方向0～22、y方向0～14として処理する。ここで、 $0 \leq Cb \leq 255$ 、 $0 \leq Cr \leq 255$ である。

$$0.6 < Cb / Cr < 0.9 \sim 0.97 \quad \cdots (37)$$

ここで、例えば、8ビットシフト(128倍)して以下の(38)式のような判定条件とすることもできる。

$$77 < (Cb < 8 / Cr) < 115 \sim 124 \quad \dots (38)$$

(処理2)

輝度信号AC係数からの検出条件(人物、顔などの輪郭部検出条件) 上述の(37)式や(38)式にして示される判定条件において、所定しきい値Athより大きいデータをx、y方向毎に検出する。

$$xh(x) > Ath \quad \dots (39)$$

$$yh(y) > Ath \quad \dots (40)$$

場合によっては検出データから共分散処理を行う。

例えば、図28に示すように、●の部分が発見点で例えば以下になる。

$$xh(0) = 0 \quad yh(0) = 0$$

$$xh(1) = 2 \quad yh(1) = 0$$

$$xh(2) = 2 \quad yh(2) = 3$$

$$\dots \quad \dots$$

$$\dots \quad \dots$$

$$\dots \quad \dots$$

(処理3)

検出物の大きさの正当性の検出条件を考え、xh(x)、yh(y)の検出で、所定しきい値Hthより大きい検出データ点

$$xh(x) > Hth \quad \dots (41)$$

$$yh(y) > Hth \quad \dots (42)$$

で、x方向、y方向毎に所定しきい値数 Lthより大きいデータ

$$xl(n) > Lth \quad \dots (43)$$

$$yl(m) > Lth \quad \dots (44)$$

の継続長を検出する。

例えば図28に示す場合では、Hth=2で

$$xh(x) > 2$$

$$y_h(y) > 2$$

の検出点が継続している部分の継続長は、

$$x_l(0) = 6$$

$$x_l(1) = 1$$

$$y_l(0) = 7$$

$$y_l(1) = 2$$

と検出され、例えば、 $L_{th} = 3$ とすると、 $x_l(0)$ 、 $y_l(0)$ がここの処理での検出データとなる。

(処理4)

人物の顔として、形状の正当性の検出条件をみる。検出された、 $x_l(n)$ 、 $y_l(m)$ の各々についてその、差分又は比が所定範囲($0 \sim D_{th}$ 又は $eth1 \sim eth2$)のデータを検出する。

$$|x_l(n) - y_l(m)| < D_{th} \quad \dots (45)$$

又は

$$eth1 < x_l(n) / y_l(m) < eth2 \quad \dots (46)$$

図28の例では、 $x_l(0)$ 、 $y_l(0)$ について演算を行う。

ここで、人物の顔の形状を考え、顔を4角形で近似すると仮定し、縦横比を演算する。

例えば、

$$0.8 < x_l(n) / y_l(m) < 1.5 \quad \dots (47)$$

を検出条件と仮定すると、

$$y_l(0) / x_l(0) = 1.2 \quad \dots (48)$$

で、図28の $x_l(0)$ 、 $y_l(0)$ の領域の物体は人物の顔の確率が高いと判定できる。

ここで、(3. 2. 4)の人物特徴を抽出するようなビットシフト処理を用いることができる。

上述の(処理1)～(処理4)の検出条件の他、以下の(処理5)のような検出データの継続性判定を行うことができる。

(処理5)

検出形状の時間継続性の条件を与える。

(継続性判定 方法5. 1)

場合によっては、上述の(処理1)～(処理4)の検出の時間継続性(検出の安定性)を判定するようにする。

例えば、(48)式からピクチャNでの検出値 $S(N)$ を

$$S(N) = y_l(0) / x_l(0) \cdots (49)$$

とし、 $S(N+1)$ 、 $S(N+2)$ などを検出して継続性の判定を行うようにする。

例えば、

$$0.8 < S(N) < 1.5 \cdots (50)$$

$$0.8 < S(N+1) < 1.5 \cdots (51)$$

$$0.8 < S(N+2) < 1.5 \cdots (52)$$

と3ピクチャ継続した場合に検出したと判定する。

ここで、検出処理を行うピクチャは、Iピクチャを用いることができる。

(継続性判定 方法5. 2)

その他の方法として、上述の(処理1)～(処理3)の検出値の何れか、又は幾つかをピクチャNでの検出データとして、 $N+1$ 、 $N+2$ 、 $N+3$ と継続検出できるかを判定するようにしてもよい。

例えば、フレームNでの検出値を

$$Col(N) = (Cb < 8) / Cr \cdots (53)$$

とし、

$$77 < Col(N) < 115 \cdots (54)$$

$$77 < Col(N+1) < 115 \cdots (55)$$

$$77 < Col(N+2) < 115 \cdots (56)$$

のように、3つのIピクチャを継続して検出できたかを判定し、次の検出処理に移行するようにする。

また、検出された $N \sim (N+2)$ ピクチャのデータの平均値を演算し、条件判定するようにしてもよい。

すなわち、検出された3ピクチャデータの平均値を $AvCol$ とし、

$$\text{AvCol} = (\text{Col}(N) + \text{Col}(N+1) + \text{Col}(N+2)) / 3$$

$$\dots (57)$$

$$77 < \text{AvCol} < 115 \dots (58)$$

を判定処理することができる。

(継続性判定 方法5. 3)

上記(39)式、(40)式により、それらをピクチャNでの検出値 $x_h(N)(x)$ 、 $y_h(N)(y)$ として、 $N+1$ 、 $N+2$ などのピクチャでの検出判定の継続性を見ることができる。

すなわち、

$$x_h(N)(x) > \text{Ath} \dots (59)$$

$$x_h(N+1)(x) > \text{Ath} \dots (60)$$

$$x_h(N+2)(x) > \text{Ath} \dots (61)$$

$$y_h(N)(y) > \text{Ath} \dots (62)$$

$$y_h(N+1)(y) > \text{Ath} \dots (63)$$

$$y_h(N+2)(y) > \text{Ath} \dots (64)$$

のように、3つのピクチャが継続して検出できたかを判定し、次の検出処理に移行するようにする。

また、検出された $N \sim (N+2)$ ピクチャのデータの平均値を演算し、条件判定するようにしてもよい。

すなわち、検出された3ピクチャデータの平均値を Avxh 及び Avyh とし、

$$\text{Avxh} = (x_h(N)(x) + x_h(N+1)(x) + x_h(N+2)(x)) / 3$$

$$\dots (65)$$

$$\text{Avyh} = (y_h(N)(y) + y_h(N+1)(y) + y_h(N+2)(y)) / 3$$

$$\dots (66) \text{ 式}$$

$$\text{Avxh} > \text{Ath} \dots (67)$$

$$\text{Avyh} > \text{Ath} \dots (68)$$

を判定処理するようにしてもよい。

(継続性判定 方法5. 4)

上記(43)式、(44)式により、それらをピクチャNでの検出値 $x_l(N)(x)$ 、 $y_l(N)(y)$

として、 $N+1$ 、 $N+2$ などのピクチャでの検出判定の継続性を見るようにしてもよい。

すなわち、

$$xl(N)(x) > Lth \quad \cdots (69)$$

$$xl(N+1)(x) > Lth \quad \cdots (70)$$

$$xl(N+2)(x) > Lth \quad \cdots (71)$$

$$yl(N)(y) > Lth \quad \cdots (72)$$

$$yl(N+1)(y) > Lth \quad \cdots (73)$$

$$yl(N+2)(y) > Lth \quad \cdots (74)$$

のように、3ピクチャ継続して検出できたかを判定し、次の検出処理に移行するようにしてもよい。

また、検出された $N \sim (N+2)$ ピクチャのデータの平均値を演算し、条件判定するようにしてもよい。

すなわち、検出された3ピクチャデータの平均値を $Avxl$ 及び $Avyl$ とし、

$$Avxl = (xl(N)(x) + xl(N+1)(x) + xl(N+2)(x)) / 3 \\ \cdots (75)$$

$$Avyl = (yl(N)(y) + yl(N+1)(y) + yl(N+2)(y)) / 3 \\ \cdots (76)$$

$$Avxl > Lth \quad \cdots (77)$$

$$Avyl > Lth \quad \cdots (78)$$

を判定処理するようにしてもよい。

(人物数検出の基本的な処理方法の概要)

ここで、人物数の検出判定を説明する。

(人数判定 方法5. 1B)

例えば、図29の場合、 x 方向の所定しきい値以上のデータ $xl(0)$ 、 $xl(1)$ の2個が検出され、 y 方向は $yl(0)$ の1個が検出されていると仮定する。

ここで、 $xl(0)$ と $yl(0)$ とで特定される領域1と、 $xl(1)$ と $yl(0)$ とで特定される領域2のデータ密度を説明する。

領域1について、領域の全データポイント $S1$ は、

$$\begin{aligned} S1 &= x1(0) \times y1(0) \\ &= 20 \quad \dots (79) \end{aligned}$$

所定しきい値より大きいデータ数は、

$$\Sigma xh(x) = 17 \quad \dots (80)$$

データ密度 $\Delta 1$ 、すなわち単位データポイント当たりのデータ数 $\Delta 1$ は、

$$\Delta 1 = 0.85 \quad \dots (81)$$

ここで、領域1がすべてしきい値より大きいデータが検出された場合はデータ密度は $\Delta 1 = 1$ になる。そこで、所定しきい値Mthを設定し、

$$\Delta 1 > Mth \quad \dots (82)$$

を判定する。

同様に領域2について、領域の全データポイントS2は、

$$\begin{aligned} S2 &= x1(1) \times y1(0) \\ &= 25 \quad \dots (83) \end{aligned}$$

となる。所定しきい値より大きいデータ数は、

$$\Sigma xh(x) = 21 \quad \dots (84)$$

となる。データ密度 $\Delta 2$ は、

$$\Delta 2 = 0.84 \quad \dots (85)$$

となる。

ここで、例えばしきい値Mthを

$$Mth = 0.80 \quad \dots (86)$$

と仮定すると、(81)式、(85)式から領域1及び領域2は条件を満たし、人物が検出されたとする確率が高いと判定される。

ここで、x方向について、 $x1(0) + x1(1)$ と $y1(0)$ で特定される領域Stは、全データポイント数が、

$$(x1(0) + x1(1)) \times y1(0) = 45 \quad \dots (87)$$

となる。検出データ数は

$$\begin{aligned} \Sigma xh(x) &= 17 + 21 \\ &= 38 \quad \dots (88) \end{aligned}$$

となる。データ密度は Δ は

$$\Delta = 8.4 \cdots (89)$$

となる。

ここで、領域 St についても

$$\Delta > Mth \cdots (90)$$

であるので、領域1と領域2は、同じ y 方向の位置に人物が検出されると判定される。

(その他の人数検出例1 (領域が重なっている場合))

図30に示す例においては、 x 方向に $x1(0)$ の1個が検出され、 y 方向に $y1(0)$ の1個が検出されている。

$x1(0)$ と $y1(0)$ で特定される領域 R について全データポイント Sr は

$$\begin{aligned} Sr &= x1(0) \times h1(0) \\ &= 90 \end{aligned} \cdots (91)$$

となる。検出データ数は、

$$\sum xh(x) = 44 \cdots (92)$$

データ密度 Δ_r は、

$$\Delta_r = 0.49 \cdots (93)$$

となる。

ここで、

$$\Delta_r < Mth$$

なので、領域 R には、1つの人物は検出されとは判定できない。

データ密度の逆数は、

$$1/\Delta_r = 2.0$$

であり、物体が2つある可能性があるが、図31のようにデータがまばらな状態に存在する場合にもデータ密度は同じになる場合がある。

図30について、 y 方向の分散 σ をみる。

$y_h(y)$ の平均値を y_{hav} 、データ数 m として

$$m = y1(0) \cdots (94)$$

$$\sigma_y = (\sum (y_h(y) - y_{hav})^2) / m$$

$$=2.32 \quad \dots (95)$$

となる。x方向については、平均値を x_{hav} データ数 n として

$$n = x_l(0) \quad \dots (96)$$

$$\sigma_x = (\sum (x_h(x) - x_{hav})^2) / n$$

$$=1.04 \quad \dots (97)$$

となる。

次に、図37について、同様にy方向、x方向の分散は、

$$\sigma_y = 0.99 \quad \dots (98)$$

$$\sigma_x = 0.64 \quad \dots (99)$$

となる。

上述の結果から、図36の方がデータの分散値が大きいことが分かる。

そこで、分散値に対して所定しきい値 B_{th} 、検出物数に応じたしきい値 $d1$, $d2$ を設定し、以下のような条件を判定し、検出物数を検出することができる。

$$\sigma_y > B_{th} \quad \dots (100)$$

$$\sigma_x > B_{th} \quad \dots (101)$$

$$d1 < 1/\Delta < d2 \quad \dots (102)$$

例えば、図36の例では、

$$B_{th} = 2.0 \quad \dots (103)$$

$$d1 = 1.8 \quad \dots (104)$$

$$d2 = 2.3 \quad \dots (105)$$

のように、しきい値を設定して判定することができる。

(その他の検出例2 (領域が対角的に離れている場合))

図32の例では、x方向に $x_l(0)$ 、 $x_l(1)$ の2個が検出され、y方向に $y_l(0)$ 、 $y_l(1)$ の2個が検出されている。

$x_l(0)$ と $y_l(0)$ で特定される領域 R_{00} について全データポイント S_{00} は

$$S_{00} = x_l(0) \times y_l(0)$$

$$=20 \quad \dots (106)$$

検出データ数は、

$$\Sigma x_h(x) = 17 \quad \dots (107)$$

データ密度 Δr は、

$$\Delta 00 = 0.85 \quad \dots (108)$$

ここで、上記(3. 2. 52)から

$$M_{th} = 0.80$$

としたので、

$$\Delta 00 > M_{th} \quad \dots (109)$$

であり、領域R00には、1つの人物が検出される確率が高いと判定される。

次に、 $x_l(0)$ と $(y_l(0) + y_l(1))$ で特定される領域Raについて全データポイントSaは

$$\begin{aligned} S_a &= x_l(0) \times (y_l(0) + y_l(1)) \\ &= 40 \quad \dots (110) \end{aligned}$$

となる。全検出データ数は、(107)式から

$$\Sigma x_h(x) = 17 \quad \dots (111)$$

であり、データ密度 Δa は、

$$\begin{aligned} \Delta a &= 17 / 40 \\ &= 0.43 \quad \dots (112) \end{aligned}$$

となる。これは、しきい値条件を満たしていない。

すなわち、今、 $x_l(0)$ と $(y_l(0) + y_l(1))$ を考えているので、仮に、 Δa が所定しきい値より大きいとすると、2つの人物が検出される確率が高いと判定される。

しかし、(112)式から Δa は、所定しきい値以下なので、 $x_l(0)$ と $(y_l(0) + y_l(1))$ で特定される領域には、2つの人物は検出されるとは判定できず、(109)式とから、1つの人物が検出されると判定することができる。

すなわち、 $x_l(0)$ と $y_l(1)$ と特定される領域には人物が検出される確立は低いと判定される。

同様に、 $x_l(1)$ と $(y_l(0) + y_l(1))$ とで特定される領域Rbについて、全データ数は

$$\Sigma x_h(x) = 17 \quad \dots (113)$$

となる。全データポイント数Sbは、

$$\begin{aligned}
 S_b &= x_l(1) (y_l(0) + y_l(1)) \\
 &= 40 \qquad \qquad \qquad \dots (114)
 \end{aligned}$$

となる。データ密度 Δb は、

$$\begin{aligned}
 \Delta b &= 17 / 40 \\
 &= 0.43 \qquad \qquad \qquad \dots (115)
 \end{aligned}$$

となる。

この(115)式から領域 R_b には2つの人物が検出される確率は低いことになる。

ここで、 $x_l(1)$ と $y_l(0)$ で特定される領域のデータ密度 $\Delta 10$ は、検出データ数が

$$\sum x_h(x) = 17$$

であり、全データポイント数が

$$x_l(1) \times y_l(0) = 20$$

なので、

$$\begin{aligned}
 \Delta 10 &= 17 / 20 \\
 &= 0.85 \qquad \qquad \qquad \dots (116)
 \end{aligned}$$

$x_l(1)$ と $y_l(1)$ とで特定される領域のデータ密度 $\Delta 11$ についても同様に

$$\Delta 11 = 0.85 \qquad \qquad \dots (117)$$

となる。

上記(115)式～(117)式から、領域10か領域11の何れかは人物が検出される確率は低いことになる。

次に、 $(x_l(0) + x_l(1))$ と $y_l(0)$ で特定される領域 R_c のデータ密度について考える。検出データ数は、

$$\sum y_h(y) = 17$$

全データポイント数は

$$(x_l(0) + x_l(1)) \times y_l(0) = 40$$

よって、データ密度 Δc は、

$$\begin{aligned}
 \Delta c &= 17 / 40 \\
 &= 0.43 \qquad \qquad \qquad \dots (118)
 \end{aligned}$$

これは、上述の所定しきい値 M_{th} 以下なので、領域 R_c は2つの人物が検出される

確率は低く、(109)式及び(115)式～(117)式などから、結局、人物が検出されるのは、 $x_l(0)$ と $y_l(0)$ で特定される領域、 $x_l(1)$ と $y_l(1)$ で特定される領域の2つとなる。

以上のような判定処理により、人物数検出を行うことができる。

(その他の人検出処理方法 (方式2))

その他の方法として、 x 方向(0～44)及び y 方向(0～29)に対して順次、所定しきい値条件を満たすかを判定して人検出を行うことができる。

データ系列を $d(x)(y)$ として、上記(37)式、(41)式及び(42)式の条件を満たすデータ系列を検出していき、例えば、

$$d(x_1)(y_1), d(x_2)(y_1)$$

$$d(x_1)(y_2), d(x_2)(y_2)$$

このように、 x 方向と y 方向で連続して検出されたとすれば、検出物の大きさと、位置も同時に分かることになる。

この方法では全データを1つずつ検出し、系列データの連続性の判定を行うなどのため、演算時間が上述の(方式1)に比較して多くかかる。

この手法を行う場合には、例えば、 x 方向及び、 y 方向にデータを1/2に圧縮してデータ処理数を減らしてから行うようにする。

(その他の人検出処理方法(方式3))

上述の(方式2)と類似の他の手法として、人物を4角形で近似し、4角形の大きさを順次変えて、その4角形の領域のデータが所定の条件を満たすかを判定することにより人物の検出を行うようにすることができる。

例えば、図33のように(2×2)、(3×3)、(4×4)の4角形の領域を設定する。

上述のような大きさの異なる4角形の領域を順次小さい4角形から1データずつ移動させ、その領域内のデータが条件を満たすか判定していき、すべての領域で判定が終了したら、次の大きさの4角形についても同様に処理を行う。

すべての大きさの4角形について、処理が終了した時点で、検出領域と、検出個数が分かるが、上述の(方式2)同様処理時間を要することになる。

3. 2. 5 テロップ特徴

テロップ検出判定処理系206では、図25に示すような画面の領域におけるDCTのAC係数の平均値を検出する。

所定領域における画面内で、所定の大きさの文字情報を含むテロップは比較的、輪郭がはっきりしており、図25の何れかの領域にテロップ画像が現れた場合に、所定しきい値以上のAC係数が検出でき、これによりテロップ検出を行うことができる。

上述のようにDCTのAC係数を検出する方法の他に、ベースバンド領域(時間領域の信号)でエッジ検出する方法を用いることができ、例えば、画像の輝度データのフレーム間差分によりエッジを検出するようにする。

また、ウェーブレット変換により、多重解像度解析を行い、所定の高周波成分データを含む所定の多重解析度領域におけるデータを用いて、図25に対応する領域の平均値を演算するようにしてDCTのAC係数を用いる場合と同様の信号を行う。

テロップは、フリップの用に淡色領域とは限らず、例えば、ニュース映像の下部に出現する文字情報であり、出現領域は、おおむね、番組ジャンルにもよるが、画面の下部、上部、又は左側際、右側際などの場合の可能性が大きい。

テロップ特徴と、フリップ特徴をまとめて、文字特徴とすることも考えられる。

3. 2. 6 カメラ特徴

カメラ特徴判定処理系209では、ズーム、パンその他などのカメラ動作に関する特徴で、例えば、特開2003-298981号公報に開示されているように、MPEGの場合には、Pピクチャのモーションベクトル(動きベクトル)を用いて判定することができる。

その他に例えば、特表2002-535894号公報にカメラ特徴に関する技術が開示されている。

(4)再生ユニット(プレイユニット)処理

ここで、要約再生(ダイジェスト再生)は、所定の信号処理による音声系特徴データ、映像系特徴データの各特徴データを用いて所定の信号処理により、所定区間内において幾つかの重要再生区間(キーフレーム区間)を選定(選択)して、その各区間を逐次、スキップ再生するようにすることができる。

スキップ再生を行う場合に、例えば、話者音声区間の途中で、スキップすると、画面上は見えていてあまり違和感がない場合でも、音声の途中で途切れる場合にユーザに

よって聴感上で違和感を生じる場合が想定されるので、所定音声レベル(音量)以下の区間を無音区間と設定し、その区間内の所定の時点スキップ時点の候補とする。

また、映像のシーンチェンジは、放送番組、映画その他の映像再生では、話題の区切りの時点とみられるので、シーンチェンジ点、又はその近傍をスキップ時点の候補とすることもできる。

上述のことから、所定音声信号における所定無音区間と所定映像信号のシーンチェンジ時点又はその近傍の所定時点に応じてスキップ再生時点、スキップ再生区間を捉えることができる。

ここでは、上述のような観点から、((スキップ再生時点間(又はその区間)))を、便宜上、所定の再生単位(以下再生ユニット又はプレイユニットPlay Unit(又はPU))を設定して処理を行う。

このように設定された再生ユニット(PU)における所定の画像系特徴データ、所定の音声系特徴データが所定処理され、それら映像、音声特徴データと要約再生時間に応じて所定の要約再生(ダイジェスト再生)区間が設定され、所定の要約再生モードでスキップ再生を行うことで所定の要約再生が実行される。

また、上述のような要約再生を行うだけでなく、所定の信号処理により設定されたPUの最初(又はその近傍)又は最後(又はその近傍)の時点にチャプタ(又は編集点、又は再生区切り点)を設定することもできる。

すなわち、上述したようなチャプタを設定することで、そのチャプタ点を所定の信号処理によりサムネール表示を行うことや、ユーザがそのサムネール表示を見て編集を行うなどの操作を行うことができる。

次に、再生ユニット(プレイユニット)(PU)の処理の一例について、図34～図35を参照して説明する。

(有音区間の場合(音声信号が所定レベル以上の場合))

図34Aに示す処理法1のように、所定平均レベル以上の有音区間で、その音声区間が10秒～20秒の範囲の場合には、シーンチェンジによらず音声セグメントが15秒に最も近い切れ目(所定の無音検出時点)を再生ユニットの区切りとする。

図34Bに示す処理法2のように、音声区間が連続して20秒より長く、シーンチェンジ区間が20秒以下の場合には、シーンチェンジの切れ目が15秒に最も近いシーンチェンジ検出点を再生ユニットの区切り点とする。

図34Cに示す処理法3のように、音声連続して20秒より長く、シーンチェンジ区間が20秒より長い場合には、音声セグメント、シーンチェンジによらず、再生ユニットが20秒になったらその時点で区切り点とする。

図34Dに示す処理法4のように、音声特徴の属性が10秒～20秒の範囲で変化する場合には、その属性変化点を再生ユニットの区切り点とする。

図34Eに示す処理法5のように、CM(コマーシャル)検出した場合には、CM検出点を再生ユニットの区切り点とする。

ここで、CM検出の方法について図35を参照して説明する。

一般的に放送番組のCMの区間長は所定時間長(通常一般的には、15秒又は30秒又は60秒)であり、CMの区切り点(開始、終了時点)ではシーンチェンジがあるので、上記所定時間長の検出とシーンチェンジ検出をすることで、図36に示すようにCMを検出することが可能である。

(無音区間の場合(音声の平均レベルが所定以下の場合))

図35Aに示す処理法6のように、無音区間(音声平均レベルが所定以下の区間)が20秒より長く、シーンチェンジ検出区間長が20秒以下の場合には、15秒に最も近いシーンチェンジ検出点を再生ユニットの区切り点とする。

図35Bに示す処理法7のように、無音区間が20秒より長く、シーンチェンジ検出区間が20秒より長い場合には、シーンチェンジ検出点によらず、再生ユニットの開始点から20秒の時点で区切り点とする。

上述の何れの再生ユニット処理の説明で、再生ユニットの開始点の初期値は、そのプログラム(放送番組)を記録した開始時点とする。

上述のような再生ユニット処理により、所定の音声特徴、所定の映像特徴(シーンチェンジ特徴)に応じた所定の再生単位を再生することができる。

(再生ユニット生成処理系ブロック構成例)

上記で説明した再生ユニットの生成する処理系と、後で説明する、この再生ユニット

に特徴データを入れ込むユニット化特徴データ処理系のブロック構成例を図37に示す。

要約再生、チャプタ点設定などの所定時点設定処理は、再生ユニットの開始点、終点に設定するので、上記で説明した再生ユニット毎に特徴データを対応付けて処理を行う。

すなわち、所定区間毎に特徴抽出した所定の各特徴データ、音声系特徴データ、映像系特徴データを再生ユニットの区間に基づいて反映させる処理を行う。

ここで、図37に示すユニット化特徴データ処理系のブロック構成例について説明する。

図37に示すブロック構成例では、無音判定情報データが時間計測系301に入力され、上記で説明した再生ユニット処理に基づく所定の間隔(時間長)が計測され、その処理出力が再生ユニット処理系302に入力される。

再生ユニット処理系302は、シーンチェンジ判定情報データとCM検出判定情報データも入力され、再生ユニット処理の各処理方法の説明で行ったような信号処理を行い所定の再生ユニットを生成する。

ここで、CM検出系304は、無音特徴検出情報データとシーンチェンジ特徴情報データ、それにCMが放送されている番組のチャンネルかを判定するチャンネル情報が入力され、図36を参照して説明したような所定の信号処理方法により、CM検出処理を行う。

再生ユニット特徴データ処理系303は、音声属性情報、無音情報などの音声系特徴データと、シーンチェンジ特徴、色特徴、類似画像特徴、人物特徴、テロップ特徴、人物特徴などの各特徴データが入力され、後で説明するように再生ユニットに各特徴データを入れ込む処理を行っている。

(5) PU特徴データ処理

次に、PU特徴データファイル処理について説明する。

PU特徴データファイルには、音声系特徴データ及び映像系(画像系)特徴データがある。

この特徴データ処理は、上述した再生ユニットに特徴抽出した各音声系、映像系特

徴データを入れ込む処理を行った結果のデータ(データファイル)であり、再生ユニット毎に各種特徴データが所定の記録媒体に記録される。

ここで、再生ユニット毎に特徴データを記録する場合には、各特徴データを所定の検出区間通りに検出した各特徴データを所定記録媒体に記録して、その後、上述した再生ユニットの所定区間に応じた特徴データに処理を行う。

特徴データは、音声信号(音声データ)、画像(映像)信号(画像(映像)データ)から、所定の特性データ(特性信号)を取り出して、その取り出した信号(データ)を所定の処理を行うことで、音声、画像の特徴を示す特徴データとすることができるが、ここでは、特別な注意書きをする場合を除き、特性データ(特性信号)から所定の処理を行って特徴を示す信号(データ)も特徴データ(特徴信号)と記述するものとする。

映像(画像)信号は、MPEGストリームから特性データとしてIピクチャにおける輝度信号(Y信号)、色信号(色差信号)(Cb、Cr信号)のDCTのDC係数、B又はPピクチャの動きベクトル(モーションベクトル)データ、また、DCTのAC係数をそれぞれ取り出し、取り出した画面位置情報、所定しきい値、相関演算などから、シーンチェンジ特徴(scen特徴)、カメラ動作特徴(カメラ特徴)(cam特徴)、類似画像特徴(類似シーン特徴又はシーンID特徴)(sid特徴)、テロップ特徴(tlp特徴)、色特徴(カラー特徴)(col特徴)、人物特徴(person特徴)などがある。

音声信号は、特性データ処理として、例えば、約20ms毎に平均レベルが演算処理されこの演算データと所定しきい値とから、所定区間における音声信号の属性(種別)、平均パワー(平均レベル)などの音声特徴(seg特徴)である。

ここでは、音声属性として、話者音声、音楽(楽音)、スポーツ番組などにおける歓声などの音声が想定される。

5.1 特徴データファイルの構成

図38に示す特徴データファイルの構成例1は、上述した音声系特徴データ、シーンチェンジ特徴(scen特徴)、カメラ特徴(cam特徴)、類似シーン特徴(sid特徴)、テロップ特徴(tlp特徴)、色特徴(col特徴)、人物特徴(person特徴)などの映像系特徴データを各々別々の特徴データファイルとする例である。

各々の特徴データファイルはテキスト形式のデータ又は、バイナリ形式のデータで

書き込まれている。

なお、これらの特徴データは、所定の記録媒体に記録するファイルデータとしての他に、通常データとして所定の記録媒体(半導体メモリなど)に一時的に記憶(記録)して、後で説明する要約リストデータ生成や所定設定時点生成(チャプタ点の生成)などの所定の処理のために読み出して用いることも考えられる。以下に説明する図39、図40の場合も同様である。

図39に示す例2は、上述したすべての音声系特徴データをテキスト形式又は、バイナリ形式の1つのファイルとしてまとめ、上述したすべての映像系特徴データをテキスト形式又は、バイナリ形式の1つのファイルとしてまとめた場合の例である。

図40に示す例3は、上記したすべての音声系特徴データ及び、上記したすべての映像系特徴データをテキスト形式又は、バイナリ形式の1つのファイルとしてまとめた場合の例である。

このように、1つのファイルとしてまとめることにより、図38の例1の場合と比較してファイル数が1つだけなので、ファイルとしての扱いが簡単になり、さらにバイナリ形式とすると、データサイズ(ファイルサイズ、ファイル容量)が小さくなり効率的になる。

ここでは、特徴データファイルを図40の例3に示すような場合で特徴データをバイナリ形式で書き込む場合について説明する。

また、図40に示す例3は、図39に示す例2において、すべての音声系特徴データをバイナリ形式で記述したデータと、すべての映像系特徴データをバイナリ形式で記述したデータを一緒にしたものなる。

上述のことから、特徴データファイルにおける以下の説明における音声系特徴データの処理方法(記述方法)は、図39に示すにおける音声系特徴データについて適用することができ、映像系特徴データの処理方法(記述方法)は、図39の例2における映像系特徴データについて適用することができる。

5.2 特徴データの階層構造

再生ユニットを単位とする特徴データの階層構造を図41に示す。

以下に示す例は、所定の処理単位(再生ユニット)における所定の特徴データ処理である。

特徴データは、図41に示すように、特徴データヘッダ情報、プログラム1特徴データ、プログラム2特徴データなどで構成される。

特徴データヘッダ情報は、図42に示すように、プログラム1、プログラム2、などプログラム全体の総記録時間、記録開始、記録終了時刻、プログラム数(番組数)、その他情報などの所定データから構成されている。

次に、プログラム1特徴データを例に、プログラム(番組)の特徴データについて説明する。

図41に示すように、プログラム1特徴データは、プログラム1情報、再生ユニット1情報、再生ユニット2情報などで構成されている。

図42に示すように、プログラム1情報は、プログラム記録時間、プログラム開始、終了時刻、プログラムジャンル(番組ジャンル)、その他情報などの所定データで構成されている。

次に、再生ユニット1情報を例に、再生ユニットのデータ構造について説明する。

図41に示すように再生ユニット1情報は、音声特徴データ、映像特徴データで構成されている。

(音声系特徴データの構成)

図41に示すように、音声系特徴データは、系列番号情報、開始終了位置情報、音声属性情報、特徴データ、その他情報データなどで構成されている。

(映像系特徴データの構成)

図41に示すように、映像系特徴データは、シーンチェンジ特徴、色特徴、類似画像特徴、人物特徴、テロップ特徴、カメラ特徴などの各所定の特徴情報データから構成されている。

以下のシーンチェンジ特徴、色特徴、類似画像特徴、人物特徴、テロップ特徴、カメラ特徴など各特徴データの説明で、すべての所定区間でその各項目の特徴データを所定記録媒体に記録(書込処理)する処理の他に、例えば、所定しきい値以上の特徴データが検出された場合のみ、そのデータとして所定の記録媒体に記録(書込処理)するように所定データ処理を行う。

このように、所定しきい値以上の特徴データが検出された場合のみ、所定のデータ

処理を行う場合には、しきい値より小さい場合には所定の特徴データが書き込まれないので、しきい値以上の特徴データが検出されて所定の記録(書込)処理が行われ、最初から何番目の特徴データ検出かを知る場合には、下記で説明する系列番号情報から知ることができる。

(シーンチェンジ特徴)

図43に示すように、系列番号情報、開始終了位置情報、特徴データ、その他データからなる。

ここで、系列番号情報は、0、1、2、3、・・・とそのプログラム(方法番組)の始めからのシーンチェンジが起きた順番を示す情報である。

開始終了位置情報は、上記各順番のシーンチェンジの開始終了の位置を示す情報データで、フレーム(フィールド)番号、PTS、DTS、時間などの情報データを用いることができる。

(色特徴)

図43に示すように、系列番号情報、検出領域を識別する情報データ、開始終了位置情報データ、特徴データ、その他データなどからなる。

ここで、系列番号情報は、0、1、2、3、・・・とそのプログラム(方法番組)の始めからの色特徴検出の順番を示す情報である。

開始終了位置情報は、上記各順番における色特徴検出で、各領域の特徴検出した開始終了の位置を示す情報データで、フレーム(フィールド)番号、PTS、DTS、時間などの情報データを用いることができる。

特徴データは、例えば、RGB、Y、Cb、Crなどのデータがある。

(類似画像特徴)

図43に示すように、系列番号情報、頻度情報開始終了位置情報、特徴データ、その他データなどからなる。

ここで、系列番号情報は、0、1、2、3、・・・とそのプログラム(方法番組)の始めからの類似画像特徴検出の順番を示す情報である。

特徴データとしては、上述したような有効画面を所定数の領域に分割(例えば25分割)した各分割領域のDCTの平均DC係数などがある。

(人物特徴)

図43に示すように、系列番号情報、検出領域を識別する情報データ、開始終了位置情報データ、特徴データ、その他データなどからなる。

ここで、系列番号情報は、0、1、2、3、・・・とそのプログラム(方法番組)の始めからの類似画像特徴検出の順番を示す情報である。

(テロップ特徴)

図43に示すように、系列番号情報、検出領域を識別する情報データ、開始終了位置情報データ、特徴データ、その他データなどからなる。

ここで、系列番号情報は、0、1、2、3、・・・とそのプログラム(方法番組)の始めからのテロップ特徴検出の順番を示す情報である。

(カメラ特徴)

図43に示すように、系列番号情報、検出領域を識別する情報データ、開始終了位置情報データ、特徴データ、その他データなどからなる。

ここで、系列番号情報は、0、1、2、3、・・・とそのプログラム(方法番組)の始めからのカメラ特徴検出の順番を示す情報である。

ここで、放送番組を記録する場合に、放送番組の所定の記録処理と同時に、ここで説明する特徴抽出処理、特徴データの書込処理(記録処理)を行うことができるが、既に、記録済みの放送番組や、その他映画、ドラマその他画像音声ソフトについて、所定の特徴抽出処理を行い、特徴データファイルを生成することもできる。

プログラム1について、上述のようにPUと特徴データを用いるとき、そのほかのプログラム2、プログラム3などを記録する場合にも、上述したプログラム1の場合と同様にPUと特徴データを用いることができる。

(6) プレイリスト処理(要約再生リスト生成処理)

次に、上記した特徴抽出処理が生成したPUファイル(PU特徴データファイル)から、要約再生(ダイジェスト再生)を行うための要約データ処理に関する説明を行う。

6.1 要約ルール処理

本願で述べる特徴データを用いる要約再生(ダイジェスト再生)では、上述したPUを単位とする所定再生区間をスキップ再生処理することで所望の要約再生(ダイジェ

スト再生)を行う。

6. 2 所定時点設定処理(プレイリストファイル)処理

次にプレイリストファイルについて説明する。

このファイルは、上記した特徴データに応じて意味付けされたPU、又はPUの接合体(PUの集合体、又はPUの連結体)の内どれを選択して再生処理を行うかの所定データの情報が所定の書式に応じて記述されているデータである。

ここで、このデータは特徴抽出の基となった画像音声データが記録された所定の記録媒体に記録する(書込処理)場合の他に、所定のメモリ手段に一時的に記憶する場合も考えられる。

プレイリストファイルの一例を図44A、図44Bに示す。

図44Aに示す例1における(a)の縦のデータ系列は、再生区間の開始位置情報のデータで、フレーム番号、時間(時刻)、ストリーム(圧縮された画像音声データ)からのPTS(プレゼンテーション・タイム・スタンプ)、あるいはDTS(デコード・タイム・スタンプ)などの所定の情報データなどである。

図44Aに示す例1における(b)の縦のデータ系列は、再生区間の終了位置情報のデータで、例1の(a)のデータと対応して、フレーム番号、時間(時刻)、ストリーム(圧縮された画像音声データ)からのPTS(プレゼンテーション・タイム・スタンプ)、あるいはDTS(デコード・タイム・スタンプ)などの所定の情報データなどである。

図44Aに示す例1における(c)の縦のデータ系列は、そのPU(再生ユニット)又は再生ユニット群(PU群)の重要度である。

図44Aに示す例1における(d)縦のデータ系列は、要約ルールで規定された、又は設定された意味の文字データである。

図44Bに示す例2は、すべてのPU区間について意味文字と評価値(重要度)を記述し、再生区間、チャプタ設定などの所定時点を示すために「1」、「0」の識別データを設けた場合の例である。

図44Bに示す例2の(a)(b)で示される開始点、終了点は、次の段のデータと連続的になっているのが分かる。

例えば、図44Bに示す例2において、最初の開始点0終了点229で、次の開始点2

30に連続的につながっている。

図44Bに示す例2における(e)の縦のデータ系列は、要約再生を行うかどうかのフラグ情報データで、「1」の場合は再生を行う場合で、「0」の場合は再生を行わない場合である。

また、「1」の最初の時点、「0」の最初の時点所定時点設定点(チャプタ点)と見ることができる。

(7)動作フローチャート

図45は、本発明の動作フローチャートの一例であり、これについて説明する。

処理を開始すると、まず、最初のステップS1で記録モードか再生モードか判定され、記録モードの場合は記録処理(R)に、また、再生モードの場合はステップS2の処理に移行する。

7.1 再生処理関係動作フローチャート

(再生処理動作フローチャートの一例)

再生モードの場合は、ステップS2で要約再生(ダイジェスト再生)モードか通常再生モードか判定され、通常再生モードの場合は通常再生処理(P)に移行する。

要約再生モードの場合は、ステップS3で所定の特徴データが所定記録媒体に記録されているかの検出処理、又は所定ファイルデータとして記録媒体の所定記録領域に記録されているかの検出処理が判定処理される。

ステップS3で所定の特徴データが検出される場合には、ステップS4で所定のプレイリストデータ(データファイル)が所定記録媒体の所定記録領域に記録されているかが検出され、プレイリストデータ(プレイリストファイル)が検出される場合は、ステップS5で所定プレイリストデータを読出処理する。

ステップS3で所定の特徴データが検出されないと判定される場合には、ステップS8でいま要約再生しようとする画像音声データ(プログラム、放送番組)を読み込んで所定の特徴抽出処理を行い、ステップS9で処理が終了したかが判定され終了しない場合はステップS8に戻り終了するまで処理を行う。

ステップS9で所定の特徴抽出処理が終了したと判定された場合には、ステップS6に移行して所定のプレイリストデータ生成処理が行われる。

ステップS4で所定のプレイリストデータ(ファイル)が検出されないと判定される場合は、ステップS6において所定の記録媒体の所定記録領域に記録され、又は記憶されている所定の特徵データを読込処理して所定のプレイリストデータ(ファイル)を生成処理して所定の記録媒体の所定領域に逐次、あるいは、処理が終了後データを書き込み、ステップS7ですべてのプレイリスト生成処理が終了したかが判定され、終了しない場合はステップS6に戻り処理を繰り返し、S7で所定のプレイリストデータがすべて生成されたと判定された場合は、ステップS5で書き込んだプレイリストデータを読込処理する。

ここで、ステップS6において、逐次生成されたプレイリストデータは上記放送番組などの画像音声情報データが記録されている同じ記録媒体上の所定記録領域に、逐次記録するようにしてもよいし、又は画像音声データが記録されたのとは別の記録媒体、例えば、装着、着脱可能な所定メモリ手段などに情報を書き込むようにしてもよい。

この場合にも、所定プレイリストデータが逐次生成処理されるとともに、逐次データを書き込む(記憶処理する)ようにしてもよいし、所定プレイリストデータがすべて生成処理され、プレイリスト処理が終了してから、生成されたすべてのプレイリストデータをまとめて記録(記憶)処理するようにしてもよい。

また、プレイリストデータは、図46、図47を参照して説明するように、記録時間に応じて、ユーザが複数の要約再生時間を選択できるように、記録時間に応じて、複数のプレイリストデータを生成するようにしてもよい。

ここでは、上述したように、所定PU区間、又は複数のPU区間の接合された所定区間毎に、所定評価値も設定処理されるので、評価値に応じて要約再生時間を操作することができる。

ステップS10で再生時間選択モードになり、ステップS11で、ユーザがすぐ再生時間を選択したか、又は要約再生モード選択した後プレイリストデータの検出処理終了後から所定時間 t_{mod} 内にユーザが再生時間を選択処理したかが判定され、選択されない場合は、S12でユーザにより再生ストップが選択されたかが判定処理される。

ステップS12でユーザにより再生ストップが選択された場合は処理を終了し、再生

ストップでない場合はステップS10に戻り上記所定の処理を繰り返す。

ステップS11で、ユーザが再生時間をすぐ選択した場合、又は上記所定時間の t_{mod} 内で再生時間を選択しない場合はステップS13で要約再生動作処理に移行する。

ここで、ユーザが再生時間を選択した場合はその要約再生時間で、再生時間を選択しないで上記所定時間 t_{mod} 経過した場合は、所定のデフォルト設定再生時間(所期設定再生時間) t_{pb0} が設定される。

ここで、ユーザにより要約再生時間を任意に選択できるようにしてもよいし、記録したプログラム記録時間とプレイリストデータに基づいた、予め設定された再生時間から選択処理できるようにしてもよい。

この場合、例えば、5分、10分、15分、20分、30分などの時間とした場合に、デフォルトの要約再生時間は、記録時間に応じて、例えば、図46のように設定することもできる。

図46に示す例では、所定記録時間以上(T_{recmin})の場合にのみ要約再生モードが設定できるようにして、この所定記録時間 T_{recmin} として、記録時間 T_{rec} が10分未満の場合は、時間が短いので、要約再生は設定されず通常再生のみとしている。

一例として、図46から記録時間 T_{rec} が60分の場合は、ユーザによる選択可能な要約再生時間は、10分、15分、30分、40分となり、デフォルトの設定時間は、30分となる。

図46に示す例では、記録時間 T_{rec} が長くなるほど、ユーザによる選択可能な要約再生時間の選択数が多くなっているが、上記したように、記録時間が短い場合は、スキップ再生処理による要約再生でスキップ処理される総区間が多くなると、それだけ情報が欠落することになり、再生内容が把握できなくなることが考えられるので選択数を少なくし、適切な要約時間の選択が行えるようにし、それに比較して記録時間が長い場合は、情報量が多いので選択数を多くしてユーザによる効果的、有効な動作が行えるようにしている。

このようなユーザによる選択可能な要約再生時間の一覧、デフォルトの再生時間などの情報は、本発明を適用した記録再生装置における所定表示手段又は、その装

置に接続された所定の表示手段、又は装置のリモコン上における液晶などの所定表示画面などに表示することが考えられる。

ここで、プレイリスト生成処理と同時に、チャプタ設定処理を行うこともでき、記録時間に応じて図44に示すように、設定可能なチャプタ数に応じて自動的に所定のチャプタ設定処理が行われる。

例えば、図44から記録時間が1時間の場合は、5～40個のチャプタが設定されるように所定の信号処理が行われる。

ステップS13では要約再生動作が行われるが、上記したように、所定PU区間又は複数のPU区間の接合区間毎に所定評価値が設定されているので、設定時間と評価値に応じてスキップ再生処理が行われ、それにより要約再生が行われる。

すなわち、評価値が高いPU区間から最優先して順次選択され、選択した要約再生時間にできるだけ近くなるように、順次、上記最優先評価値に比較して評価値の小さい区間を選択処理していく。

ステップS14では再生動作を終了するか判定され、終了の場合は処理を終了し、終了しない場合はステップS15で再生している所定プログラム(番組)が終了したか判定され、終了の場合は処理を終了し終了しない場合は、ステップS16に移行し再生時間を変更するか判定する。

ステップS16で再生時間を変更する場合はステップS10に戻り、上記処理を繰り返す、変更しない場合はステップS13に戻り、要約再生動作を繰り返す。

7. 2 記録処理関係動作フローチャート

(記録処理動作フローチャートの一例)

記録モードの場合における動作フローチャートの例を図48に示す。

図45に示したフローチャートのステップS1で記録モードが選択された場合は、図48に示すフローチャートのステップR1でタイマ記録モードか通常記録モードかが判定され、通常記録モードの場合は、ステップR9に移行し通常記録動作を行う。

ステップR9の通常記録動作で所定の記録信号処理に移行して、ステップR10においてMPEGなどの所定エンコード処理される画像音声データ、又はエンコード処理された画像音声データから所定の特徴抽出処理が行われる。

ここで、記録信号処理と特徴抽出信号処理は、同時に行うことができる。

所定エンコード処理される画像音声データについては、所定エンコード処理される途中の画像音声データを用いて所定の特徴抽出処理を行うもので、例えば、画像のDCT処理系からDCT信号処理のDC係数データ、AC係数データなどを取り出すことができ、それら所定のデータを用いて所定信号処理を行うことでシーンチェンジ特徴の検出(カット点特徴の検出)、テロップ特徴の検出など上述した各所定の特徴抽出信号処理を行う。

音声データは、所定の帯域圧縮信号処理における所定サブバンド信号処理において、所定サブバンド帯域におけるデータを用いることで、話者音声、音楽(楽音)判定検出などの信号処理を行うことができる。

楽音判定信号処理については、例えば、所定サブバンド帯域におけるデータの継続性を判定することで判定処理を行うことができる。

また、ベースバンド帯域の画像音声データを用いることもでき、例えば、画像のベースバンド信号を用いて、フレーム(又はフィールド)間差分信号処理によりシーンチェンジ検出処理や、その差分信号によるエッジ検出によりテロップ特徴信号処理など、その他所定の特徴抽出信号処理を行うことができる。

ここで、各画像、音声特徴抽出信号処理された特徴データは、画像音声データが記録される同じ所定記録媒体、又は所定のバッファメモリなどの所定データ記憶手段(データ記録手段)に記録する。

ステップR11で通常記録モード終了か判定され、終了ではない場合はステップR9に戻り、上記動作を繰り返し、終了の場合は、ステップR12に移行しプレイリストデータ生成処理(又はチャプタデータ生成処理)に移行する。

ステップR1でタイマ記録モードの場合は、ステップR2で記録開始、記録終了時刻設定を行い、ステップR3で所定の動作時刻か判定され、所定時刻ではない場合は、ステップR7で動作待機し、ステップR8でユーザによりタイマ動作解除の割り込み処理が行われたか判定され、タイマ動作を継続する場合は、ステップR3に戻り上記動作を繰り返す。

ステップR8でタイマ動作が解除された場合は、図45のステップS1に戻り、最初の

動作モード選択処理を行う。

ステップR3で所定の記録動作時刻になったと判定されたら、記録動作を開始し、上述したステップR9～ステップR11と同様の動作をステップR4～ステップR6で行う。

特徴データは、上述したように、各画像、音声特徴抽出信号処理された特徴データ(特徴抽出データ)は画像音声データが記録される同じ所定記録媒体、又は所定のバッファメモリなどの所定データ記憶手段(データ記録手段)に記録する。ステップR6で記録終了時刻と判定された場合は、ステップR12に移行してプレイリストデータ生成処理又はチャプタデータ生成処理を行う。

ステップR12では、各種の所定特徴抽出処理された特徴データ(特徴抽出処理された所定特徴データを所定の加工処理、所定の信号処理を施したデータ、それらデータを用いて所定判定処理を行ったデータなども含む)を所定記録媒体から読出処理を行い、所定のプレイリストデータ(ファイル)生成処理、チャプタデータ生成処理を行う。

生成されたプレイリストデータ、チャプタデータは、所定記録媒体に記録され、ステップR13で生成処理が終了したか判定処理され、終了しない場合は、ステップR12に戻り上記処理動作を繰り返し、ステップR13で終了したと判定された場合は動作を終了する。

ここで、プレイリストデータ、チャプタデータは、逐次、データの生成処理と同時に所定記録媒体に記録する場合の他に、上記、処理対象にしている所定の放送番組、プログラム、又は所定記録区間に対する所定プレイリストデータ、チャプタデータのすべての生成処理が終了した後に、まとめて所定記録媒体に記録するようにしてもよい。(特徴抽出処理と平行して(同時に)プレイリストデータ(チャプタ)処理を行う場合)

ここで、上述の説明では、所定の放送番組、プログラムなど画像音声情報データの記録処理と同時に所定の特徴抽出処理を行い、特徴抽出処理した各種の特徴データ(特徴抽出データ、又は特徴データを用いて所定の加工、所定の信号処理を施した信号を含む)を所定の記録媒体に記録して、上記所定の放送番組、プログラムが終了した後、記録した特徴データを読み出して、プレイリストデータ(ファイル)、チャ

プタデータなどを生成処理する場合を述べたが、特徴抽出処理と同時に、又は特徴抽出処理と平行してプレイリストデータ(ファイル)、チャプタデータ生成処理を行うようにしてもよい。

7.3 再生ユニット処理関係動作フローチャート

(PU処理で所定データ区間毎に音声セグメント処理とシーンチェンジ処理を行う場合の動作フローチャート)

上述したPU信号処理の場合で、音声セグメント検出点とシーンチェンジ検出点から所定信号処理を行う動作フローチャートの一例を図49に示す。

処理を開始するとステップP1で画像音声情報データが記録されている所定記録媒体から音声データ、及び後で説明するシーンチェンジ検出処理のために画像データの所定サンプルデータ数を読出処理して、ステップP2で読み出したデータをメモリなど所定の記録手段であるデータバッファに記憶処理(書込処理、記録処理)を行っていく。

ステップP3で所定サンプル数のデータがバッファに記録されたと判定された場合はステップP4に移行し、まだ所定サンプルデータが記録されないと判定された場合はステップP2に戻り動作を繰り返す。

ここで、ステップP2～ステップP7ではPU処理のために、所定、音声信号の有音、無音判定処理を考えるので、ステップP2の所定サンプルデータ数としては、大よそ0.1秒くらい～1秒くらいの所定区間の間に相当するデータ数のバッファ処理を行う。

例えば、サンプリング周波数48KHzの場合は、1秒間で48000サンプルデータなので、0.1秒の場合は4800サンプルのデータをバッファに記録する。

ステップP4でバッファから音声データを読出処理し、ステップステップP5で、上述したような所定区間の音声レベルの演算処理を行い、ステップP6で所定レベルと比較処理を行い、所定レベル以上か所定レベル以下かの判定処理を行って、無音検出(無音判定)処理が行われる。

ステップP6でその区間が無音区間と判定された場合は、ステップP7でその情報を所定メモリ(バッファ)に記憶(記録)し、無音でなく有音と判定された場合はステップP8に移行し、ステップP1で読み込んだバッファのデータの音声バッファ処理が終了し

たか判定処理され、終了しない場合はステップP2に戻り上記の処理を繰り返し、終了した場合はステップP9に移行する。

ステップP9では、ステップP8で処理された音声セグメント情報データを読み出し、ステップP10で上記した短い無音区間、有音区間、長い無音区間、有音区間のセグメント処理を行う。

ステップP11では、所定データサンプル数の画像データのDCT処理データを所定のバッファメモリ(所定データ記録手段)に記録処理を行い、ステップP12で所定データ量の記録が終了したかが判定され、所定データ量ではない場合は、ステップP11に戻り上記バッファメモリ系への書込処理を繰り返し、ステップP12で所定データ量の書込処理が終了したと判定された場合は、ステップP13に移行する。

ステップP13では上記所定のバッファメモリ系から記録した(書込処理した)所定のDCTデータを読出処理し、ステップP14において、フレーム間差分などの所定信号処理を行い、所定のシーンチェンジ検出処理を行う。

ステップP15で所定のシーンチェンジがあったか判定処理され、シーンチェンジがあったと判定される場合は、ステップP16で所定のメモリ手段(データ記録手段、データバッファ手段など)にシーンチェンジがあった時点の位置情報データを記憶(書込処理)してステップP17に移行し、ステップP15でシーンチェンジがないと判定された場合はステップP17に移行する。

ステップP17では、所定データバッファ内の所定データ量の上記シーンチェンジ検出処理が終了したか判定処理され、終了しない場合はステップP11に戻り上記信号処理を繰り返し、ステップP17で終了したと判定される場合は、ステップP18に移行する。

ステップP18では所定バッファメモリ手段に記録された(記憶された)シーンチェンジ位置情報を読み出し、ステップP19で所定区間長より短いなど、短過ぎる区間は前後区間と接合するなどの、シーンチェンジ検出区間の補正処理を行う。

ステップP20では上記、所定区間における生成処理された音声セグメント位置情報データ及びシーンチェンジ位置情報データを読み出し、ステップP21で音声セグメント位置、音声セグメント区間長、シーンチェンジ位置、シーンチェンジ区間長などの所

定情報データから、所定のPUの位置情報、区間情報など所定PU情報データを生成処理する。

ステップP22では、ステップP21で処理されたPU情報から、そのPU区間に対応する特徴データ(又は特徴抽出データ、又は特徴データを所定の信号処理を行った信号など)を所定の記録媒体、又は所定のデータバッファに書込処理を行う。

上記したように、これら記録媒体は、いま処理の対象としている放送番組、プログラムなど所定区間の画像音声情報データが記録されているのと同じ所定記録媒体上における所定記録領域の他に、別の所定記録媒体上に記録(記憶、書込処理)を行うことも考えられる。

ステップP23では所定データ量の上記音声セグメント処理、シーンチェンジ処理、PU処理など一連の信号処理が終了したか判定処理され、終了したと判定される場合は処理を終了し、終了していないと判定された場合はステップP1に戻り、上述した処理を繰り返す。

(PU処理ですべての音声セグメント処理を行った後にシーンチェンジ処理を行う場合の動作フローチャート)

ここで、上述の例では、記録した所定放送番組、プログラムなどの、画像音声データの所定区間毎に、逐次、音声データのセグメント処理を行い、その後、画像のシーンチェンジ検出処理を行ったが、上述したように所定区間毎の処理ではなく、今処理の対象としている放送番組、プログラムの所定区間すべての音声セグメント処理が終了した後、すべてのシーンチェンジ検出処理を行い、すべてのシーンチェンジ検出処理が終了した後、所定のPU処理を行うようにすることもできる。

上述したPU信号処理の場合で、音声セグメント検出点とシーンチェンジ検出点から所定信号処理を行う動作フローチャートの他の一例を図50に示す。

処理を開始すると、まず最初のステップT1において図49に示すフローチャートにおけるステップP1～ステップP9で説明したような所定音声セグメント処理を行う。

ここで、音声データは所定バッファメモリに逐次所定データサンプル量のデータを読み込んで行う。

ステップT2で音声セグメント処理を行ったセグメント位置情報のデータを所定メモリ

手段(データ記憶手段、データ記録手段)に記録していき、ステップT3において、いま処理対象となっている放送番組、プログラムなどの所定区間すべての音声データについて所定セグメント処理が終了したか判定され、終了しないと判定された場合はステップT1に戻り上記の処理を繰り返し、終了したと判定された場合はステップT4に移行する。

ステップT4において上記図49のフローチャートにおけるステップP11～ステップP18で説明したような所定シーンチェンジ処理を行う。ここで、画像のDCTデータは所定バッファメモリに逐次所定データサンプル量のデータを読み込んで行う。

ステップT5で所定シーンチェンジ処理を行ったシーンチェンジ位置情報のデータを所定メモリ手段(データ記憶手段、データ記録手段)に記録していき、ステップT6において、いま処理対象となっている放送番組、プログラムなどの所定区間すべての画像のDCTデータについて所定シーンチェンジ処理が終了したか判定され、終了しないと判定された場合はステップT4に戻り上述の処理を繰り返し、終了したと判定された場合はステップT7に移行する。

ステップT7では上記所定メモリ手段から所定音声セグメント位置情報のデータと、所定シーンチェンジ位置情報のデータを読み出し、ステップT8で所定PU処理を行い、ステップT9で、いま処理対象となっている放送番組、プログラムなどの所定区間すべての区間にわたり所定PU処理が終了したか判定され、終了したと判定された場合は処理を終了し、終了しないと判定された場合はT7に戻り上記動作を繰り返す。

なお、本発明は、図面を参照して説明した上述の実施例に限定されるものではなく、添付の請求の範囲及びその主旨を逸脱することなく、様々な変更、置換又はその同等のものを行うことができることは当業者にとって明らかである。

請求の範囲

- [1] 1. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定の記録媒体からデータを読み込んでその読み込んだデータから上記画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、
- 上記第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は上記第1の意味付け処理の信号に基づく1つの所定区間に対して所定の評価値処理を行い、
- 上記第2の意味付け処理された信号又は上記評価値処理された信号に基づいて所定の再生区間を設定するか又は所定の再生時点を設定することを特徴とする情報信号処理方法。
- [2] 2. 所定のデータ入力系により所定の操作により上記情報信号処理を実行するためのソフトウェアを取り込み、上記情報信号処理を実行可能な状態に設定し、所定の操作系により所定の動作モードが設定された場合に上記情報信号処理を実行することを特徴とする請求の範囲第1項記載の情報信号処理方法。
- [3] 3. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定の記録媒体からデータを読み込んでその読み込んだデータから上記画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、
- 上記第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は上記第1の意味付け処理の信号に基づく1つの所定区間に対して所定の評価値処理を行い、
- 上記第2の意味付け処理又は評価値処理されたデータを所定の記録媒体又は所定データメモリの所定の記録領域に記録を行うことを特徴とする情報信号処理方法。
- [4] 4. 所定のデータ入力系により所定の操作により上記情報信号処理を実行するためのソフトウェアを取り込み、上記情報信号処理を実行可能な状態に設定し、
- 所定の操作系により所定の動作モードが設定された場合に上記情報信号処理を

実行することを特徴とする請求の範囲第3項記載の情報信号処理方法。

- [5] 5. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録されたデータファイルを読み込んでその読み込んだデータから上記画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、

上記第1の情報的意味付け処理データと上記画像音声情報信号に対する関連する付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行い、

その演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は上記第1の意味付け処理に基づく信号に対する所定の評価値処理を行う

ことを特徴とする情報信号処理方法。

- [6] 6. 所定のデータ入力系により所定の操作により上記情報信号処理を実行するためのソフトウェアを取り込み、上記情報信号処理を実行可能な状態に設定し、所定の操作系により所定の動作モードが設定された場合に上記情報信号処理を実行することを特徴とする請求の範囲第5項記載の情報信号処理方法。

- [7] 7. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録されたデータファイルを読み込んでその読み込んだデータから上記画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、

上記第1の情報的意味付け処理データと上記画像音声情報信号に対する関連する付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行い、

その演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理を行い、

上記第2の意味付け処理によるデータと上記評価値処理、又は上記第1の意味付け処理に基づく信号に対する所定の評価値処理によるデータが記録された所定の

記録媒体からそれらデータを読み込む処理を行い、

上記読み込んだデータに基づいて上記画像音声情報信号の所定区間又は所定時点を設定処理する

ことを特徴とする情報信号処理方法。

- [8] 8. 所定のデータ入力系により所定の操作により上記情報信号処理を実行するためのソフトウェアを取り込み、上記情報信号処理を実行可能な状態に設定し、所定の操作系により所定の動作モードが設定された場合に上記情報信号処理を実行することを特徴とする請求の範囲第7項記載の情報信号処理方法。

- [9] 9. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録されたデータファイルを読み込んでその読み込んだデータから上記画像音声情報信号の所定区間毎の所定の第1の情動的意味付け処理を行い、

上記第1の情動的意味付け処理によるデータと上記画像音声情報信号に対する関連した付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行いその演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続する所定の第2の情動的意味付け処理とその第2の情動的意味に対する所定の評価値処理又は上記第1の意味付け処理に基づく信号に対する所定の評価値処理を行い、

上記第2の意味付け処理による所定データ又は上記評価値処理によるデータを用いるか又はそれらデータが記録された所定の記録媒体からデータを読み込んで読み込んだデータに基づいて上記画像音声信号の所定区間又は所定時点を設定する

ことを特徴とする情報信号処理方法。

- [10] 10. 所定のデータ入力系により所定の操作により上記情報信号処理を実行するためのソフトウェアを取り込み、上記情報信号処理を実行可能な状態に設定し、所定の操作系により所定の動作モードが設定された場合に上記情報信号処理を実行することを特徴とする請求の範囲第9項記載の情報信号処理方法。

- [11] 11. 所定時間長の画像音声情報データから抽出した所定区間毎の特性データとその画像音声情報データに関係する付加情報データを検出できる場合にはその付加

情報データを用いて上記画像音声情報データにおける所定区間毎に生成された所定の意味的特徴データとその意味的特徴データに対する所定の評価データを検出する又はその意味的特徴データと評価データが記録された所定のファイルを読み込んで所定の処理により上記意味的特徴データと評価データを検出し、

上記時間長が所定しきい値以上の場合には上記データ検出からの上記意味的特徴データと評価データを用いて所定ルールに従って上記画像音声情報データにおける上記所定区間を順次再生し上記時間長が所定しきい値に満たない場合には上記ルールによらないか又は上記所定ルール以外のルール処理に応じて上記時間長以下の時間内で画像音声情報信号の所定区間を再生するように再生制御を行う

ことを特徴とする情報信号処理方法。

- [12] 12. 所定のデータ入力系により所定の操作により上記情報信号処理を実行するためのソフトウェアを取り込み、上記情報信号処理を実行可能な状態に設定し、所定の操作系により所定の動作モードが設定された場合に上記情報信号処理を実行することを特徴とする請求の範囲第11項記載の情報信号処理方法。

- [13] 13. 所定時間長の画像音声情報データから抽出した所定区間毎の特性データとその画像音声情報データに係る付加情報データを検出できる場合にはその付加情報データを用いて上記画像音声情報データにおける所定区間毎に生成された所定の意味的特徴データとその意味的特徴データに対する所定の評価データを検出する又はその意味的特徴データと評価データが記録された所定のファイルを読み込んで所定の処理により上記意味的特徴データと評価データを検出し、

上記時間長が所定しきい値以上の場合には上記データ検出からの上記意味的特徴データと評価データを用いて所定ルールに従って上記画像音声情報データにおける上記所定区間又は所定時点を設定し、

上記時間長が所定しきい値に満たない場合には上記ルールによらないか又は上記所定ルール以外のルール処理に応じて上記時間長以下の時間内で画像音声情報信号の所定区間又は所定時点を設定処理する

ことを特徴とする情報信号処理方法。

- [14] 14. 所定のデータ入力系により所定の操作により上記情報信号処理を実行するため

のソフトウェアを取り込み、上記情報信号処理を実行可能な状態に設定し、所定の操作系により所定の動作モードが設定された場合に上記情報信号処理を実行することを特徴とする請求の範囲第13項記載の情報信号処理方法。

- [15] 15. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ又はその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された記録媒体又は所定メモリ手段からデータを読み込んで、上記画像音声情報信号の所定区間毎の所定の第1の情動的意味付け処理を行う第1の情報処理手段と、

上記第1の情報処理手段からの信号を用いて第1の所定の意味付けされた区間を複数接続して所定の第2の情動的意味付け処理とその第2の情動的意味に対する所定の評価値処理又は上記第1の情報処理手段からの信号に基づく1つの所定区間に対して所定の評価値処理を行う第2の情報処理手段と、

上記第1又は第2の情報処理手段からの信号に基づいて所定の再生区間を設定するか又は所定の再生時点を設定する情報信号処理手段と

を備えたことを特徴とする情報信号処理装置。

- [16] 16. 所定の操作により所定の情報信号処理を実行するためのソフトウェアを取り込むデータ入力系と、上記データ入力系により取り込まれたソフトウェアにより所定の情報信号処理を実行可能な状態に設定する信号処理設定手段とを備え、上記情報処理手段により、上記第1又は第2の情報処理手段からの信号に基づいて所定の再生区間を設定するか又は所定の再生時点を設定することを特徴とする請求の範囲第15項記載の情報信号処理装置。

- [17] 17. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された記録媒体からデータを読み込んでその読み込んだデータから上記画像音声情報信号の所定区間毎の所定の第1の情動的意味付け処理を行う第1の情報処理手段と、

上記情報処理手段からの信号を用いて第1の所定の意味付けされた区間を複数接続して所定の第2の情動的意味付け処理とその第2の情動的意味に対する所定の評価値処理を行うか又は上記第1の情報処理手段からの信号に基づく1つの所定区間に対して所定の評価値処理を行う第2の情報処理手段と、

上記第2の情報処理手段からのデータを所定の記録媒体又は所定データメモリの所定の記録領域に記録する記録手段と

を備えたことを特徴とする情報信号処理装置。

- [18] 18. 所定の操作により所定の情報信号処理を実行するためのソフトウェアを取り込むデータ入力系と、上記データ入力系により取り込まれたソフトウェアにより所定の情報信号処理を実行可能な状態に設定する信号処理設定手段とを備え、上記記録手段により、上記第2の情報処理手段からのデータを所定の記録媒体又は所定データメモリの所定の記録領域に記録することを特徴とする請求の範囲第17項記載の情報信号処理装置。

- [19] 19. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ又はその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定記録媒体から所定データを読み込んでその読み込んだデータから上記画像音声情報信号の所定区間毎の所定の第1の情動的意味付け処理を行う第1の情報処理手段と、

上記第1の情報処理手段からのデータと上記画像音声情報信号に対する関連する付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行いその演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続する所定の第2の情動的意味付け処理とその第2の情動的意味に対する所定の評価値処理又は上記第1の情報処理手段からの信号に基づく所定信号に対する所定の評価値処理を行う第2の情報処理手段と

を備えたことを特徴とする情報信号処理装置。

- [20] 20. 所定の操作により所定の情報信号処理を実行するためのソフトウェアを取り込むデータ入力系と、上記データ入力系により取り込まれたソフトウェアにより所定の情報信号処理を実行可能な状態に設定する信号処理設定手段とを備え、上記第2の情報処理手段により、上記第1の情報処理手段からのデータと上記画像音声情報信号に対する関連する付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行いその演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続する所定の第2の情動的意味付け処理とその第2の情動的意

味に対する所定の評価値処理又は上記第1の情報処理手段からの信号に基づく所定信号に対する所定の評価値処理を行うことを特徴とする請求の範囲第19項記載の情報信号処理装置。

- [21] 21. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ又はその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定記録媒体から所定データを読み込んでその読み込んだデータから上記画像音声情報信号の所定区間毎の所定の第1の情動的意味付け処理を行い、上記第1の情動的意味付け処理データと上記画像音声情報信号に対する関連する付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行い、その演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続して所定の第2の情動的意味付け処理とその第2の情動的意味に対する所定の評価値処理又は上記第1の意味付け処理に基づく信号に対しての所定の評価値処理を行い、上記第2の意味付け処理によるデータ又は上記評価値処理によるデータが記録された所定の記録媒体からそれらデータを読み込むデータ処理手段と、

上記データ処理手段からの信号に応じて上記画像音声情報信号の所定区間を設定するか又は所定時点を設定する情報処理手段と

を備えたことを特徴とする情報信号処理装置。

- [22] 22. 所定の操作により所定の情報信号処理を実行するためのソフトウェアを取り込むデータ入力系と、上記データ入力系により取り込まれたソフトウェアにより所定の情報信号処理を実行可能な状態に設定する信号処理設定手段とを備え、上記情報処理手段により、上記データ処理手段からの信号に応じて上記画像音声情報信号の所定区間を設定するか又は所定時点を設定することを特徴とする請求の範囲第21項記載の情報信号処理装置。

- [23] 23. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ又はその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定記録媒体から所定データを読み込んで、上記画像音声情報信号の所定区間毎の所定の第1の情動的意味付け処理を行う第1の情報処理手段と、

上記第1の情報処理手段からのデータと上記画像音声情報信号に対する関連する

付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行いその演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続する所定の第2の情動的意味付け処理とその第2の情動的意味に対する所定の評価値処理又は上記第1の情報処理手段に基づく信号に対しての所定の評価値処理を行う第2の情報処理手段と、

上記第2の情報処理手段からのデータ又はそのデータが記録された所定記録媒体からの所定データに基づいて上記画像音声情報信号の所定区間又は所定の時点を設定する情報処理手段と

を備えたことを特徴とする情報信号処理装置。

- [24] 24. 所定の操作により所定の情報信号処理を実行するためのソフトウェアを取り込むデータ入力系と、上記データ入力系により取り込まれたソフトウェアにより所定の情報信号処理を実行可能な状態に設定する信号処理設定手段とを備え、上記情報処理手段により、上記第2の情報処理手段からのデータ又はそのデータが記録された所定記録媒体からの所定データに基づいて上記画像音声情報信号の所定区間又は所定の時点を設定することを特徴とする請求の範囲第21項記載の情報信号処理装置。

- [25] 25. 所定時間長の画像音声情報データから抽出した所定区間毎の特性データとその画像音声情報データに関係する付加情報データを検出できる場合にはその付加情報データを用いて上記画像音声情報データにおける所定区間毎に生成された所定の意味的特徴データとその意味的特徴データに対する所定の評価データを検出するか又はその意味的特徴データと評価データが記録された所定のデータを読み込んで所定の処理により上記意味的特徴データと評価データを検出するデータ検出手段と、

所定動作モードのときに上記時間長が所定しきい値以上の場合には上記データ検出手段からの上記意味的特徴データと評価データを用いて所定ルールに従って上記画像音声情報データにおける上記所定区間を順次再生し上記時間長が所定しきい値に満たない場合には上記ルールによらないか又は上記所定ルール以外のルール処理に応じて上記時間長以下の時間内で画像音声情報信号の所定区間を再生

するように再生制御を行う再生手段と

を備えたことを特徴とする情報信号処理装置。

- [26] 26. 所定の操作により所定の情報信号処理を実行するためのソフトウェアを取り込むデータ入力系と、上記データ入力系により取り込まれたソフトウェアにより所定の情報信号処理を実行可能な状態に設定する信号処理設定手段とを備え、上記再生手段により、所定動作モードのときに上記時間長が所定しきい値以上の場合には上記データ検出手段からの上記意味的特徴データと評価データを用いて所定ルールに従って上記画像音声情報データにおける上記所定区間を順次再生し上記時間長が所定しきい値に満たない場合には上記ルールによらないか又は上記所定ルール以外のルール処理に応じて上記時間長以下の時間内で画像音声情報信号の所定区間を再生するように再生制御を行うことを特徴とする請求の範囲第23項記載の情報信号処理装置。

- [27] 27. 所定時間長の画像音声情報データから抽出した所定区間毎の特性データとその画像音声情報データに関する付加情報データを検出できる場合にはその付加情報データを用いて上記画像音声情報データにおける所定区間毎に生成された所定の意味的特徴データとその意味的特徴データに対する所定の評価データを検出するか又はその意味的特徴データと評価データが記録された所定のデータを読み込んで所定の処理により上記意味的特徴データと評価データを検出するデータ検出手段と、

所定動作モードのときに上記時間長が所定しきい値以上の場合には上記データ検出手段からの上記意味的特徴データと評価データを用いて所定ルールに従って上記画像音声情報データにおける上記所定区間又は所定時点を設定して、上記時間長が所定しきい値に満たない場合には上記ルールによらないか又は上記所定ルール以外のルール処理に応じて上記時間長以下の時間内で画像音声情報信号の所定区間又は所定時点を設定処理する情報処理手段と

を備えたことを特徴とする情報信号処理装置。

- [28] 28. 所定の操作により所定の情報信号処理を実行するためのソフトウェアを取り込むデータ入力系と、上記データ入力系により取り込まれたソフトウェアにより所定の情報

信号処理を実行可能な状態に設定する信号処理設定手段とを備え、上記情報処理手段により、所定動作モードのときに上記時間長が所定しきい値以上の場合には上記データ検出手段からの上記意味的特徴データと評価データを用いて所定ルールに従って上記画像音声情報データにおける上記所定区間又は所定時点を設定して、上記時間長が所定しきい値に満たない場合には上記ルールによらないか又は上記所定ルール以外のルール処理に応じて上記時間長以下の時間内で画像音声情報信号の所定区間又は所定時点を設定処理することを特徴とする請求の範囲第27項記載の情報信号処理装置。

- [29] 29. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定の記録媒体からデータを読み込んでその読み込んだデータから上記画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、

上記第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は上記第1の意味付け処理の信号に基づく1つの所定区間に対して所定の評価値処理を行い、

上記第2の意味付け処理された信号又は上記評価値処理された信号に基づいて所定の再生区間を設定するか又は所定の再生時点を設定する制御プログラムがコンピュータにより読取実行可能に記録されたプログラム記録媒体。

- [30] 30. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録された所定の記録媒体からデータを読み込んでその読み込んだデータから上記画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、

上記第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は上記第1の意味付け処理の信号に基づく1つの所定区間に対して所定の評価値処理を行い、

上記第2の意味付け処理又は評価値処理されたデータを所定の記録媒体又は所定データメモリの所定の記録領域に記録を行う制御プログラムがコンピュータにより読取実行可能に記録されたプログラム記録媒体。

- [31] 31. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録されたデータファイルを読み込んでその読み込んだデータから上記画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、
- 上記第1の情報的意味付け処理データと上記画像音声情報信号に対する関連する付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行い、
- その演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は上記第1の意味付け処理に基づく信号に対する所定の評価値処理を行う制御プログラムがコンピュータにより読取実行可能に記録されたプログラム記録媒体。
- [32] 32. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像音声情報信号に関連する付加情報データから又はそれらデータが記録されたデータファイルを読み込んでその読み込んだデータから上記画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、
- 上記第1の情報的意味付け処理データと上記画像音声情報信号に対する関連する付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行い、
- その演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続して所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理を行い、
- 上記第2の意味付け処理によるデータと上記評価値処理、又は上記第1の意味付け処理に基づく信号に対する所定の評価値処理によるデータが記録された所定の記録媒体からそれらデータを読み込む処理を行い、
- 上記読み込んだデータに基づいて上記画像音声情報信号の所定区間又は所定時点を設定処理する制御プログラムがコンピュータにより読取実行可能に記録されたプログラム記録媒体。
- [33] 33. 所定の画像音声情報信号の所定区間毎の画像音声特徴データ及びその画像

音声情報信号に関連する付加情報データから又はそれらデータが記録されたデータファイルを読み込んでその読み込んだデータから上記画像音声情報信号の所定区間毎の所定の第1の情報的意味付け処理を行い、

上記第1の情報的意味付け処理によるデータと上記画像音声情報信号に対する関連した付加情報データが検出できる場合にはその付加情報データとを用いて所定の演算処理を行いその演算処理の結果に応じて第1の所定の意味付けされた区間を複数接続する所定の第2の情報的意味付け処理とその第2の情報的意味に対する所定の評価値処理又は上記第1の意味付け処理に基づく信号に対する所定の評価値処理を行い、

上記第2の意味付け処理による所定データ又は上記評価値処理によるデータを用いるか又はそれらデータが記録された所定の記録媒体からデータを読み込んで読み込んだデータに基づいて

上記画像音声信号の所定区間又は所定時点を設定する制御プログラムがコンピュータにより読取実行可能に記録されたプログラム記録媒体。

- [34] 34. 所定時間長の画像音声情報データから抽出した所定区間毎の特性データとその画像音声情報データに関係する付加情報データを検出できる場合にはその付加情報データを用いて上記画像音声情報データにおける所定区間毎に生成された所定の意味的特徴データとその意味的特徴データに対する所定の評価データを検出する又はその意味的特徴データと評価データが記録された所定のファイルを読み込んで所定の処理により上記意味的特徴データと評価データを検出し、

上記時間長が所定しきい値以上の場合には上記データ検出からの上記意味的特徴データと評価データを用いて所定ルールに従って上記画像音声情報データにおける上記所定区間を順次再生し上記時間長が所定しきい値に満たない場合には上記ルールによらないか又は上記所定ルール以外のルール処理に応じて上記時間長以下の時間内で画像音声情報信号の所定区間を再生するように再生制御を行う制御プログラムがコンピュータにより読取実行可能に記録されたプログラム記録媒体。

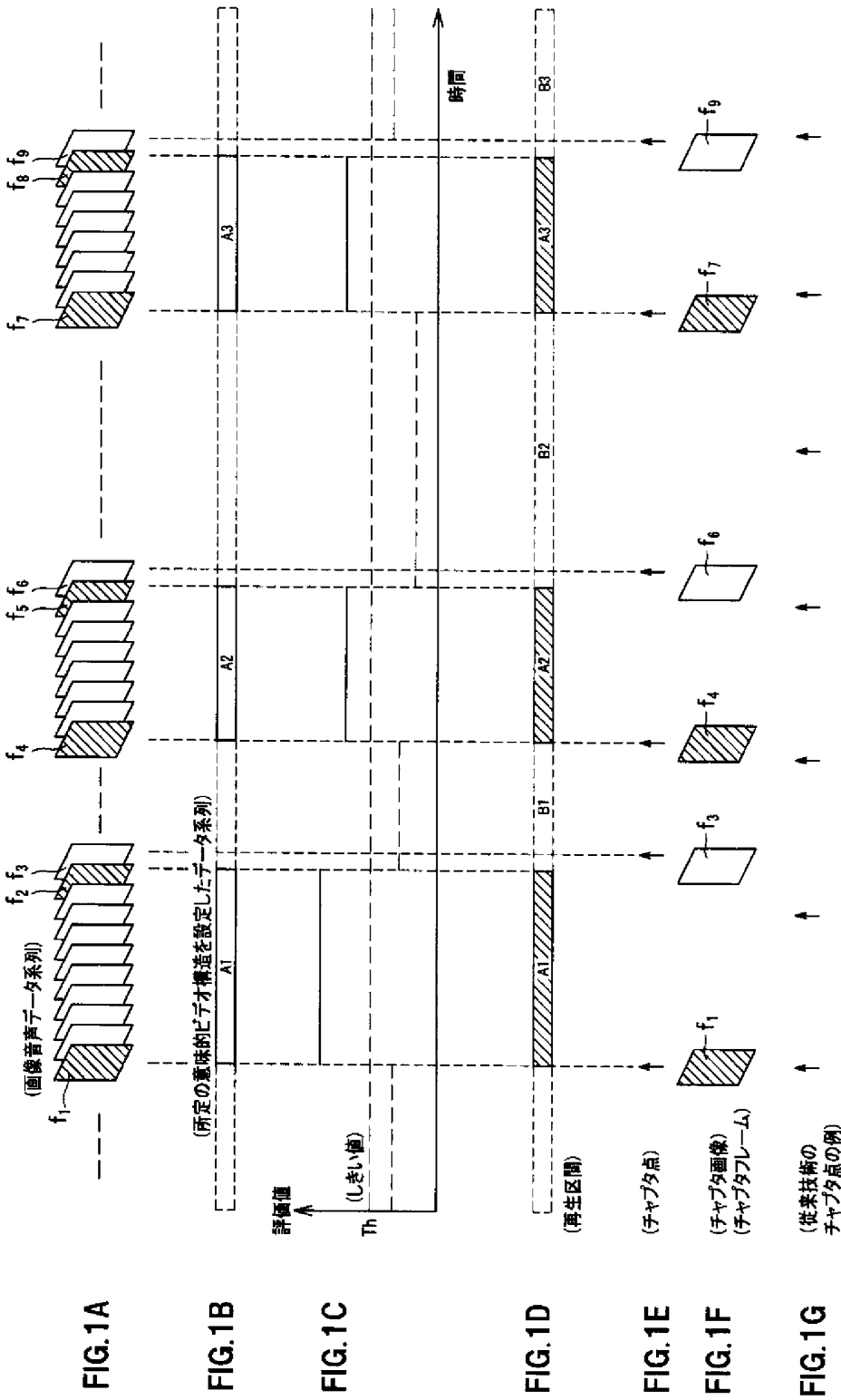
- [35] 35. 所定時間長の画像音声情報データから抽出した所定区間毎の特性データとその画像音声情報データに関係する付加情報データを検出できる場合にはその付加

情報データを用いて上記画像音声情報データにおける所定区間毎に生成された所定の意味的特徴データとその意味的特徴データに対する所定の評価データを検出する又はその意味的特徴データと評価データが記録された所定のファイルを読み込んで所定の処理により上記意味的特徴データと評価データを検出し、

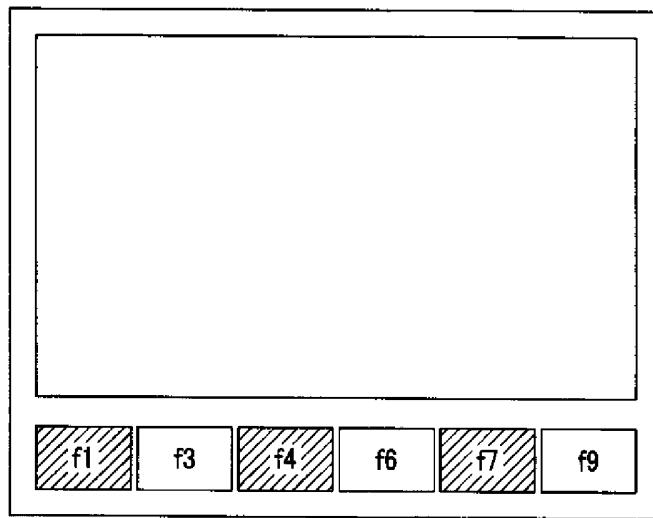
上記時間長が所定しきい値以上の場合には上記データ検出からの上記意味的特徴データと評価データを用いて所定ルールに従って上記画像音声情報データにおける上記所定区間又は所定時点を設定し、

上記時間長が所定しきい値に満たない場合には上記ルールによらないか又は上記所定ルール以外のルール処理に応じて上記時間長以下の時間内で画像音声情報信号の所定区間又は所定時点を設定処理する制御プログラムがコンピュータにより読取実行可能に記録されたプログラム記録媒体。

[図1]



[図2]

**FIG.2**

[図3]

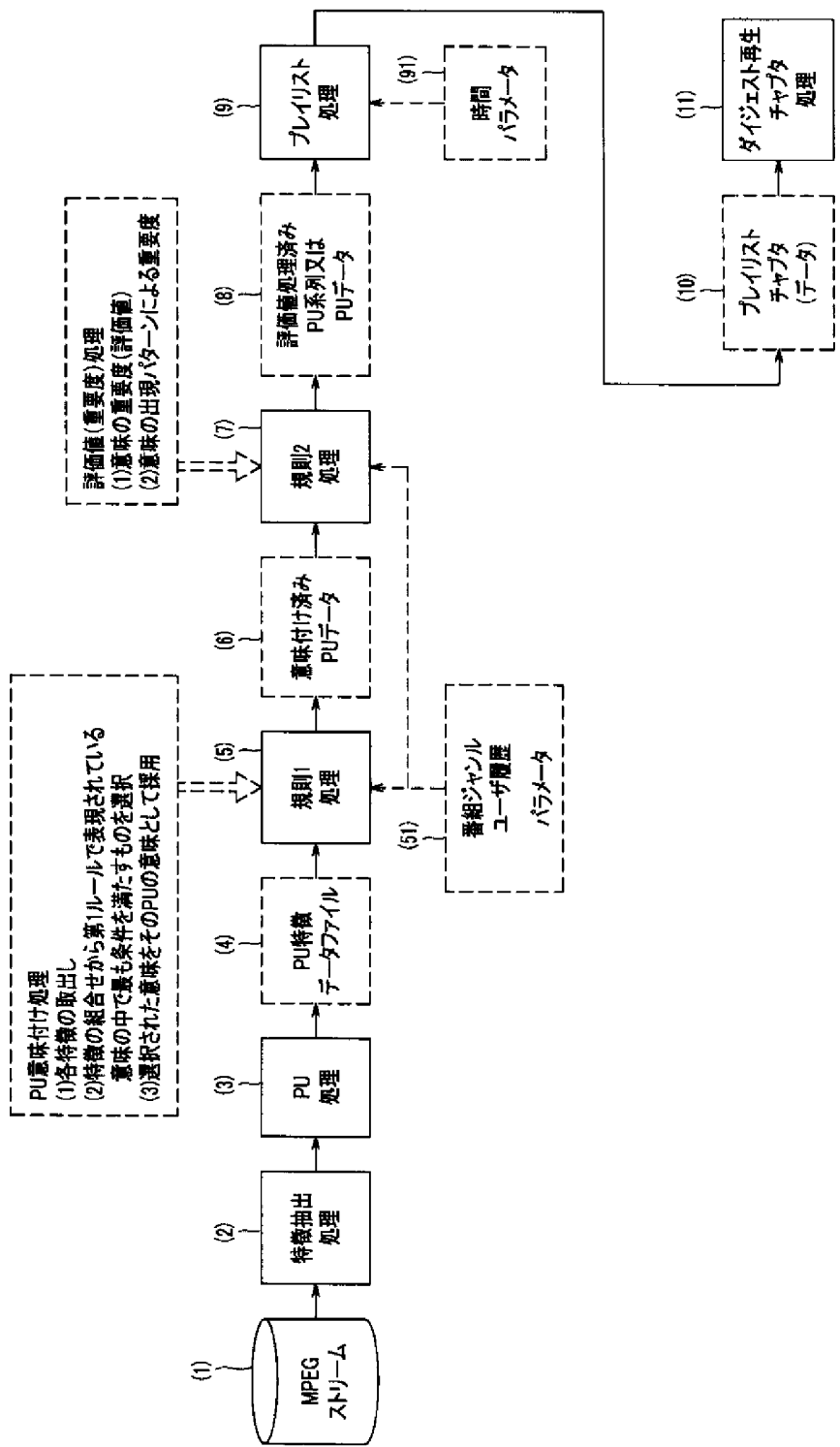


FIG.3

[図4]

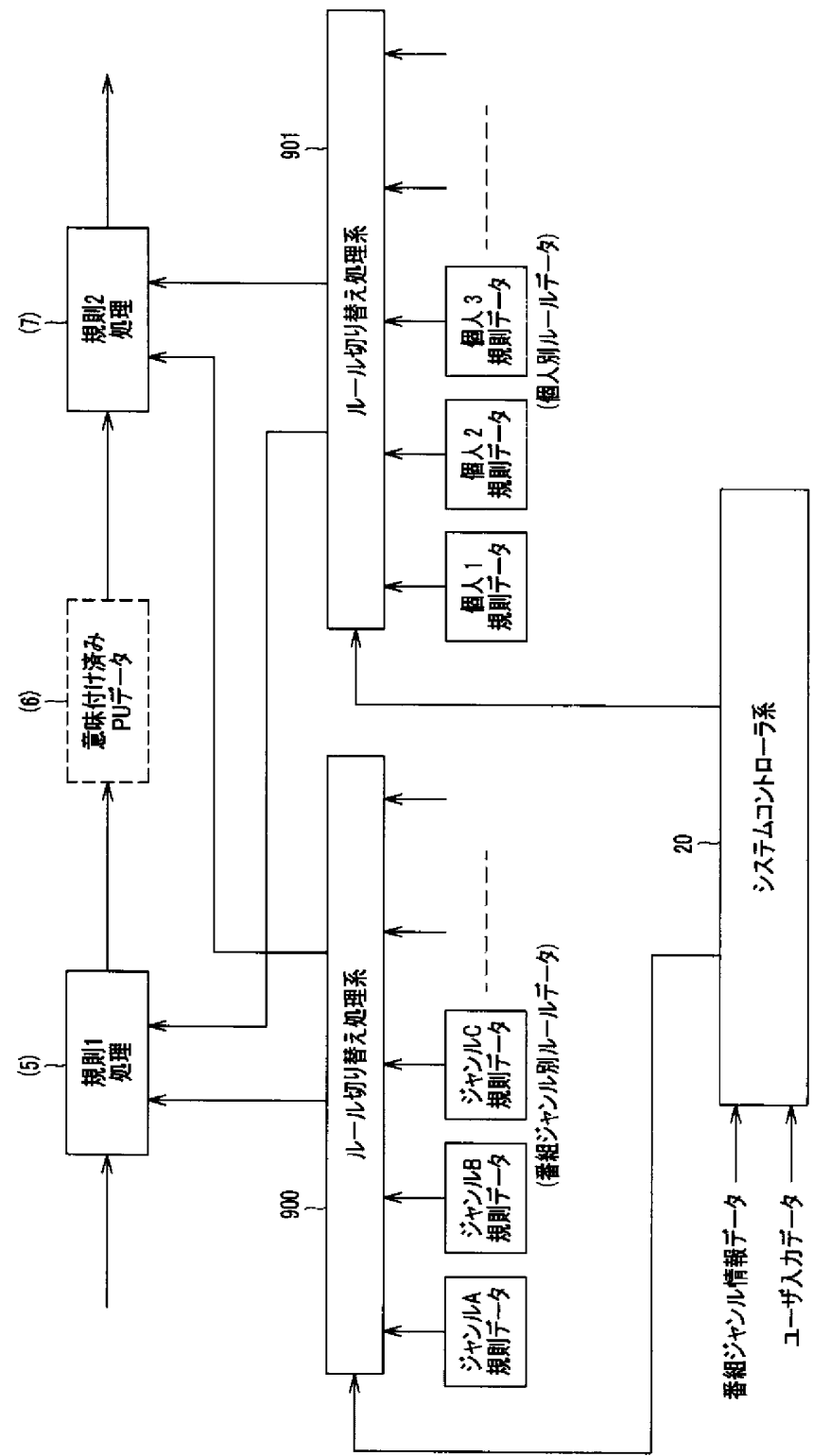


FIG.4

[図5A]

ジャンル: ニュース	定義文字	a	b	c	d	e
	意味定義	アナウンサーシーン	現場シーン	天気予報	番組開始シーン	番組終了シーン
音声属性 情報	A1 話者音声	○	△	○		
	A2 音楽				○	○
	A3 その他		△			
色特徴 (col)情報	B1 領域 1					
	B2 領域 2	△		○		
	B3 領域 3	△		○		
	B4 領域 4					
類似画像 (sid)情報	C1 出現頻度 1位	△				
	C2 出現頻度 2位	△				
人物特徴 (person)情報	D1 領域 1	△				△
	D2 領域 2	△				△
	D3 領域 3					
	D4 領域 4					
	D5 領域 5	△				△
テロップ特徴 (tip)情報	E1 領域 1					
	E2 領域 2				○	○
	E3 領域 3				○	○
	E4 領域 4					
カメラ特徴 (cam)情報	F1 静止	○				
	F2 ズーム		△			
	F3 パン		△			
シーンチェンジ(scn)情報	G1		△		△	△

FIG.5A

[図5B]

ジャンル: スポーツ(相撲)	定義文字	a 取組み紹介シーン	b 立会いシーン	c 取組みシーン	d 番組開始シーン	e 勝敗決定シーン
音声属性 情報	A1	△		△		△
	A2				△	
	A3		△	○	△	○
色特徴 (col)情報	B1					
	B2	△	△	△		△
	B3	△	△	△		△
	B4					
類似画像 (sid)情報	C1	△	△	△		
	C2	△	△	△		
人物特徴 (person)情報	D1		△	△		△
	D2		△	△		△
	D3		△	△		△
	D4		△	△		△
	D5		△	△		△
テロップ特徴 (tip)情報	E1					
	E2	○			○	
	E3	○			○	
	E4					
カメラ特徴 (cam)情報	F1	○				
	F2			△		△
	F3					
シーンチェンジ(scn)情報	G1		△	△	△	△

FIG.5B

[図6]

FIG.6A

規則 1 の例

$$\begin{aligned} a &= 1.0(A1)100 * (1.0(B2)100 + 1.0(B3)100) * (1.0(C1)100 + 1.0(C2)100) * (1.0(D1)100 + 1.0(D2)100 + 1.0(D5)100) * 1.0(F1)100 \\ b &= (1.0(A1)100 + 1.0(A3)100) * (1.0(F2)100 + 1.0(F3)100) * 1.0(G1)100 \\ c &= 1.0(A1)100 * (1.0(B2)100 * 1.0(B3)100) \\ d &= 1.0(A2)100 * (1.0(E2)100 * 1.0(E3)100) * 1.0(G1)50 \\ e &= 1.0(A2)100 * (1.0(D1)100 + 1.0(D2)100 + 1.0(D5)100) * (1.0(E2)100 * 1.0(E3)100) * 1.0(G1)50 \end{aligned}$$

FIG.6B

規則 2 の例

$$\begin{aligned} &\text{news} \\ &(aabb)100 \rightarrow Ps(ts,s4), P1(t1,s4), Pe(te,s3) \\ &(ca)70 \rightarrow Ps(ts,s4), P1(t1,s4), Pe(te,s3) \end{aligned}$$

FIG.6C

規則 3 の例

$$\begin{aligned} &(\text{news})100 \rightarrow Ps(ts,s4), P1(t1,s4), Pe(te,s3) \\ &(\text{sports})70 \rightarrow Ps(ts,s4), P1(t1,s4), Pe(te,s3) \end{aligned}$$

[図7]

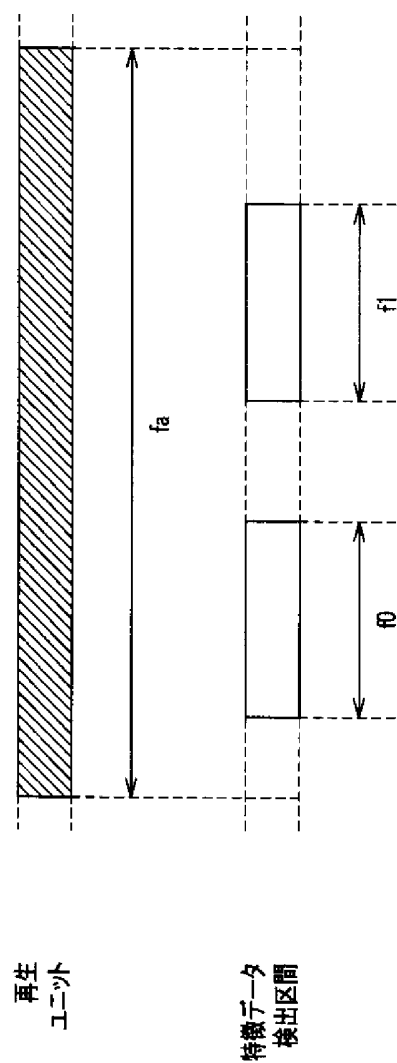


FIG.7

[図8]

FIG.8A

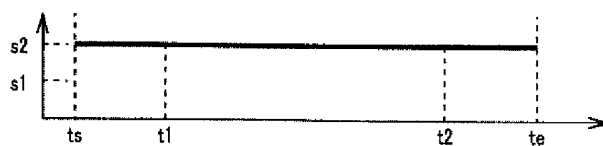


FIG.8B

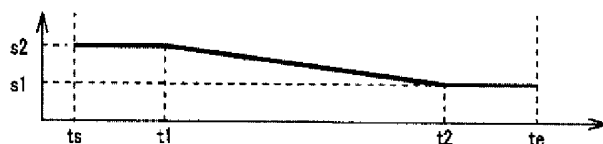


FIG.8C



FIG.8D

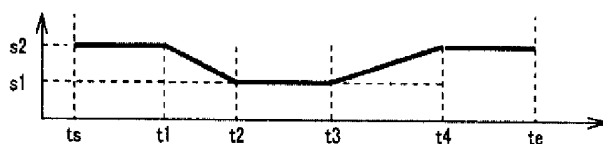


FIG.8E

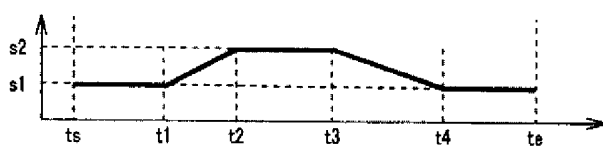


FIG.8F

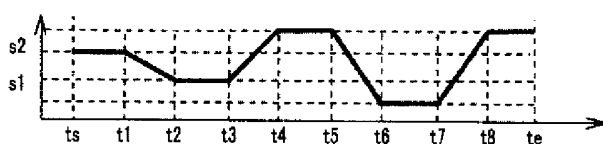


FIG.8G

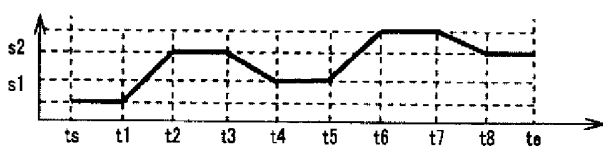


FIG.8H

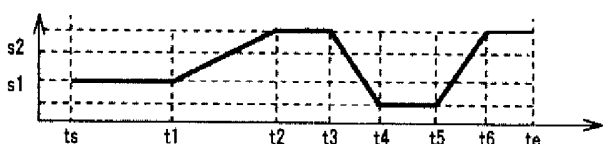
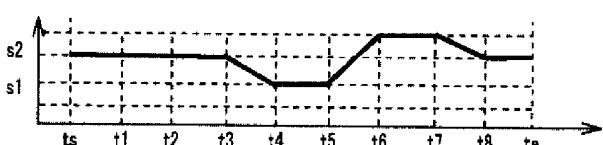


FIG.8I



[図9]

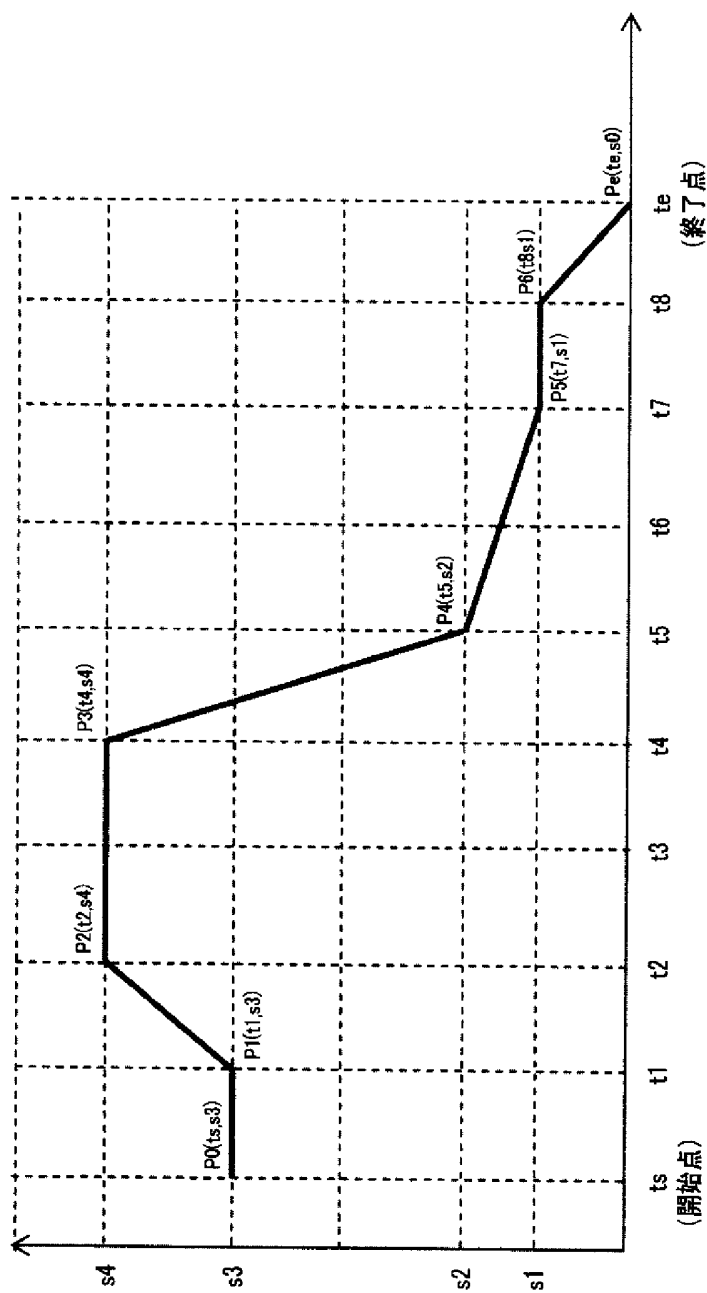


FIG.9

[図10]

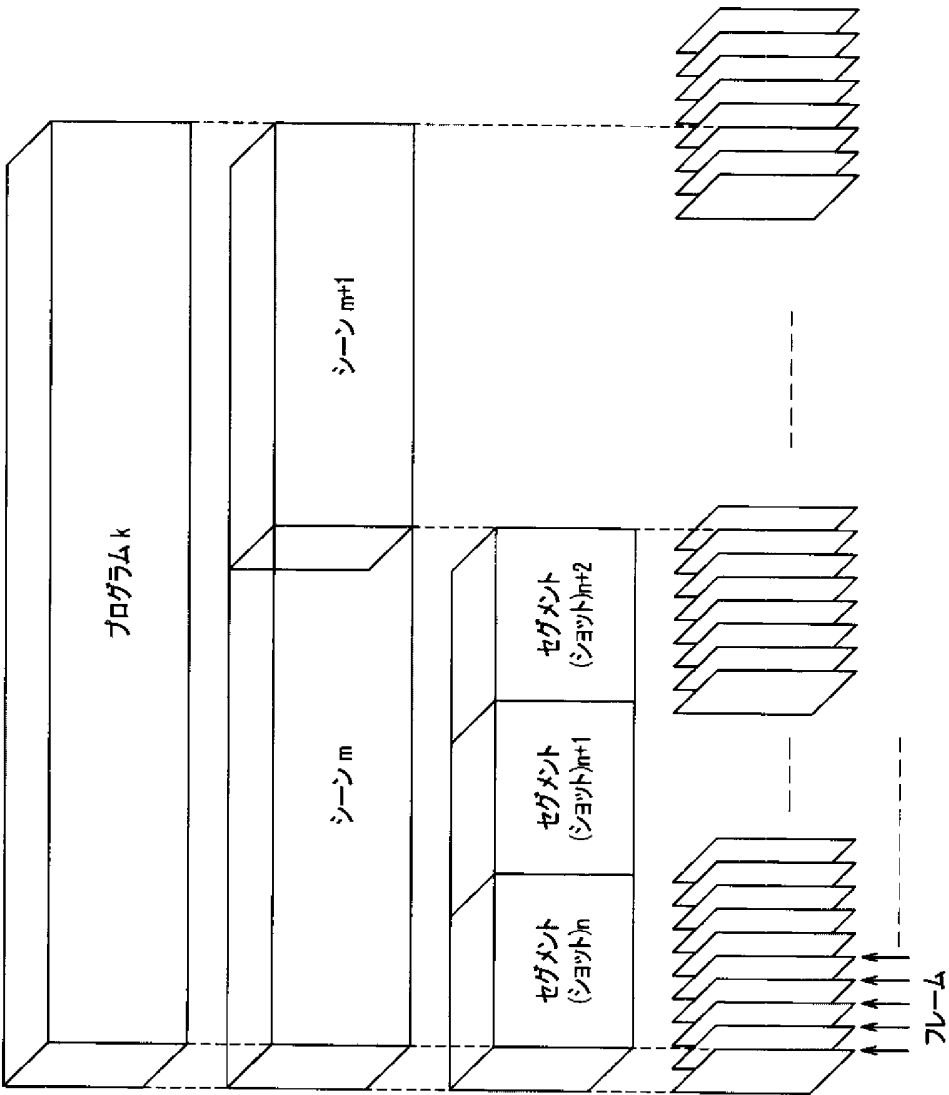


FIG.10

[図11]

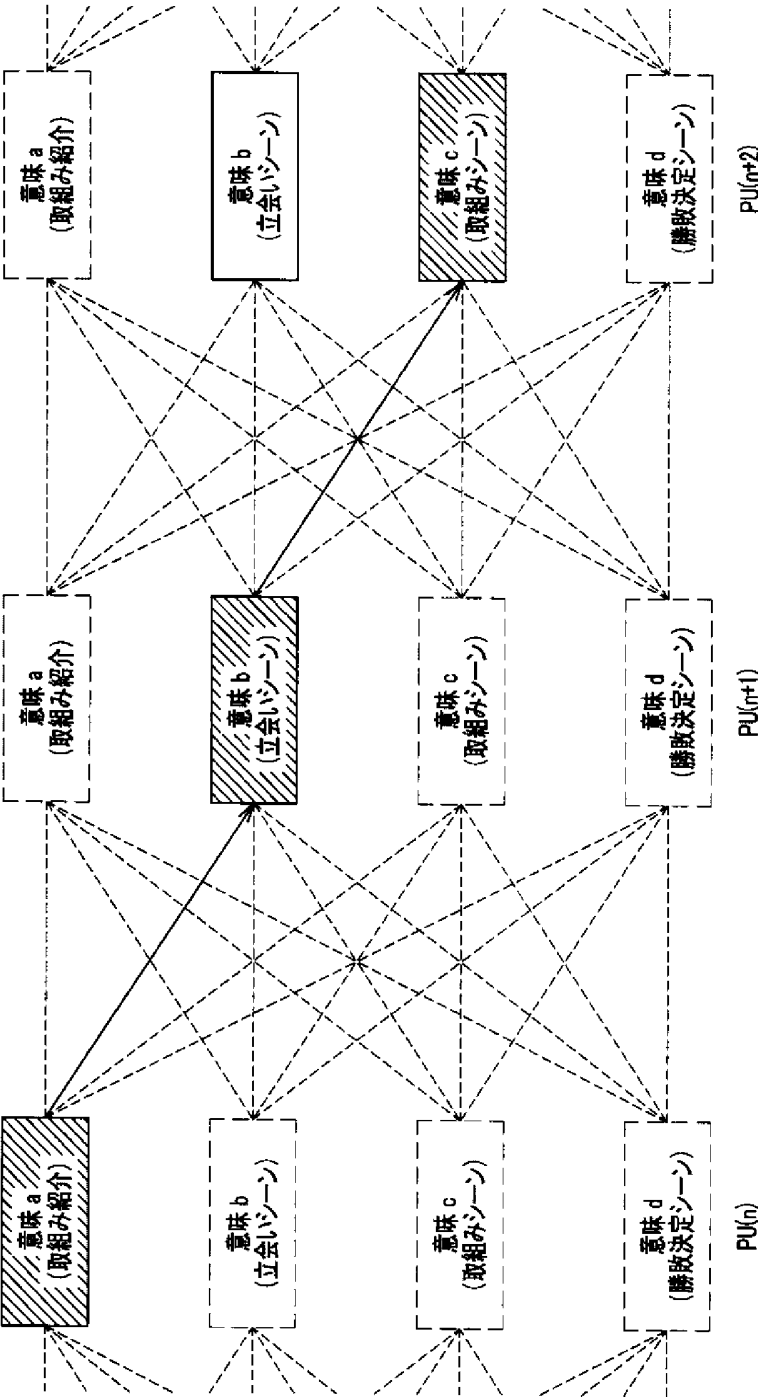


FIG.11

[図12]

FIG.12A

ニュース番組の場合

定義文字による シーン系列	a	a	b	b	a
意味による シーン系列	アナウンサー シーン	アナウンサー シーン	現場 シーン	現場 シーン	アナウンサー シーン

FIG.12B

相撲番組の場合

定義文字による シーン系列	a	a	c	c	d
意味による シーン系列	取り組み 紹介	立会い シーン	取り組み シーン	取り組み シーン	勝敗決定 シーン

[図13]

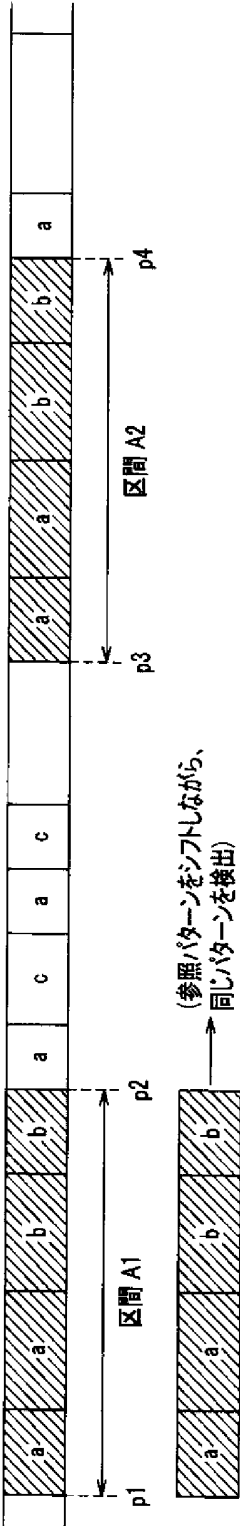
FIG.13A

参照パターン

a	a	b	b
アナウンサー シーン	アナウンサー シーン	現場 シーン	現場 シーン

FIG.13B

プログラム(番組)全系列



[図14]

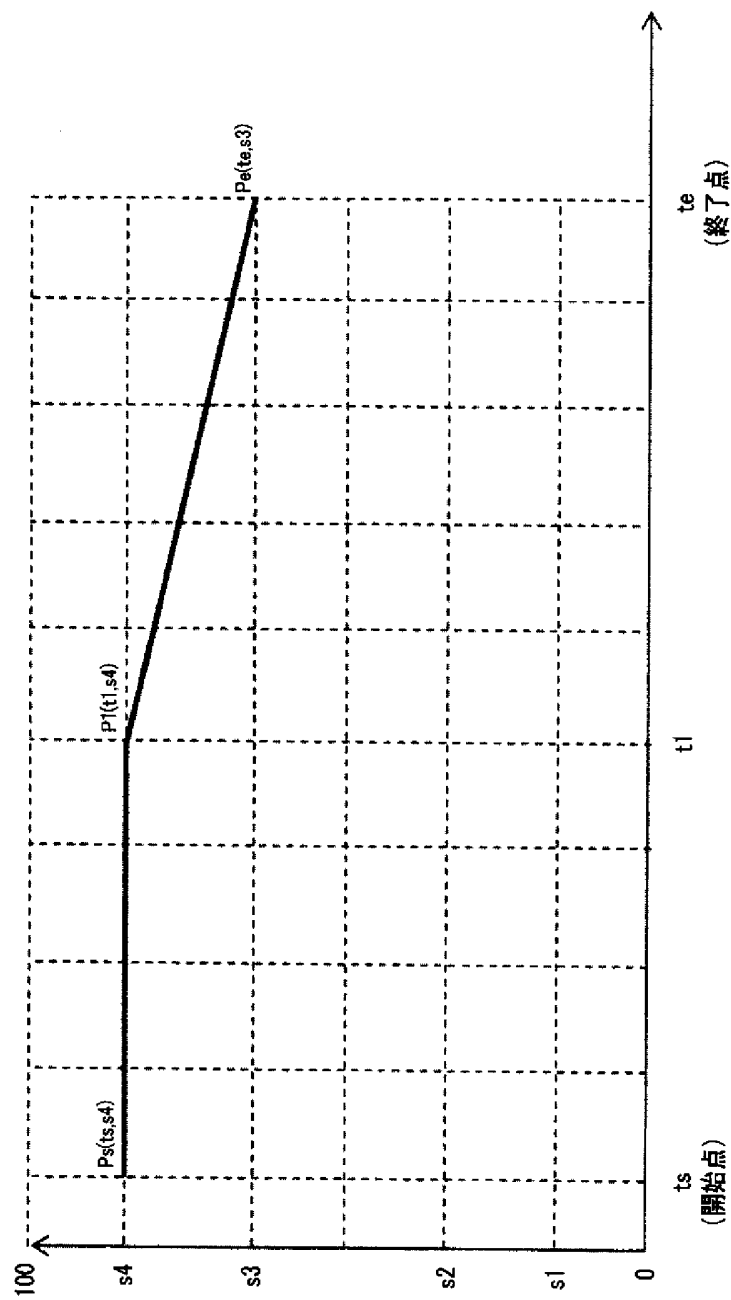


FIG.14

[図15]

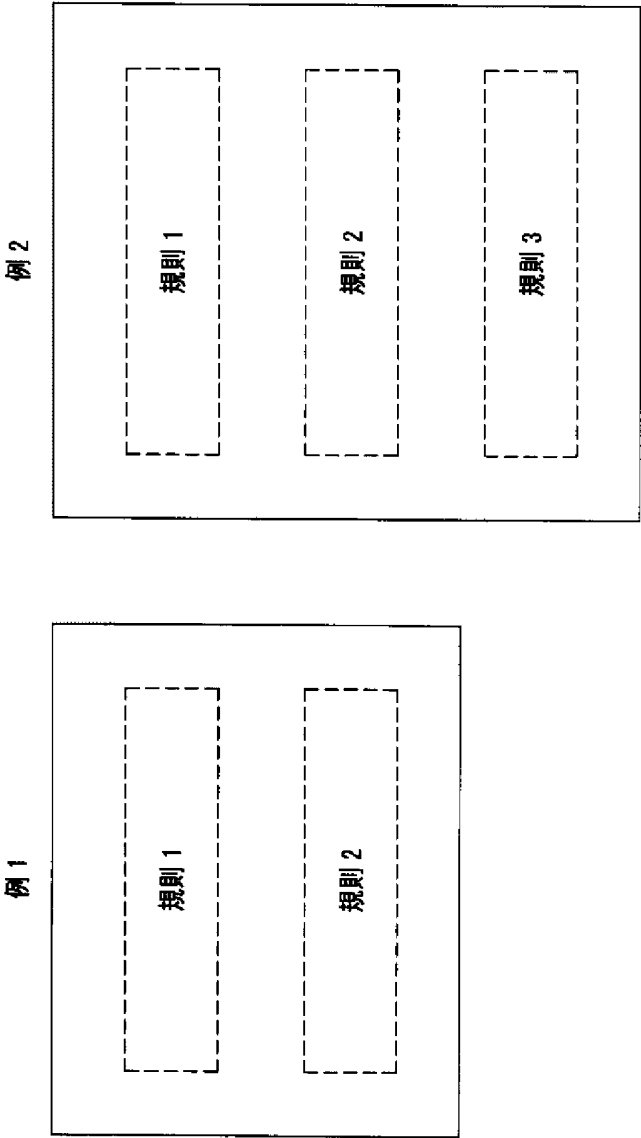


FIG.15A

FIG.15B

[図16]



FIG.16A

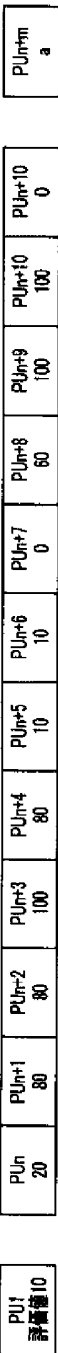
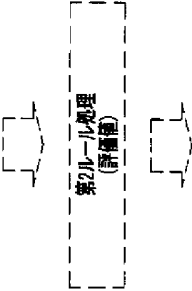


FIG.16B

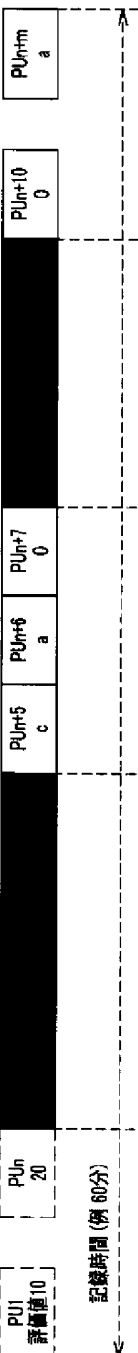
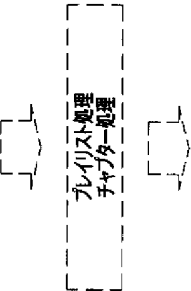


FIG.16C

FIG.16D

所定時点設定
(チャプター点)

[図17]

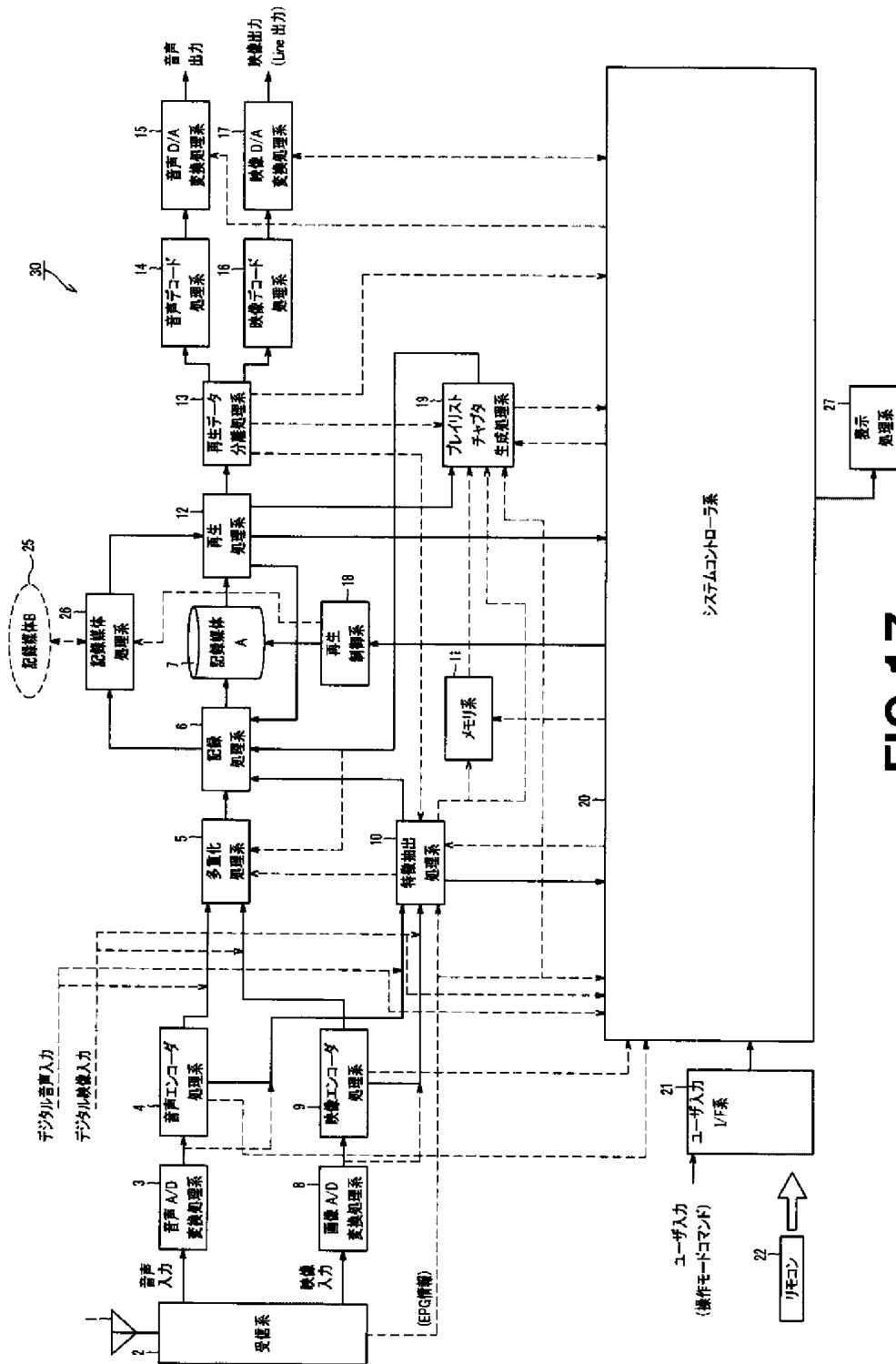


FIG.17

[図18]

	(a)	(b)	(c)	(d)
特徴抽出データ	記録されている	記録されていない	記録されている	記録されていない
プレイリストデータ チャプタデータ	記録されている	記録されている	記録されていない	記録されていない

FIG.18

[図19]

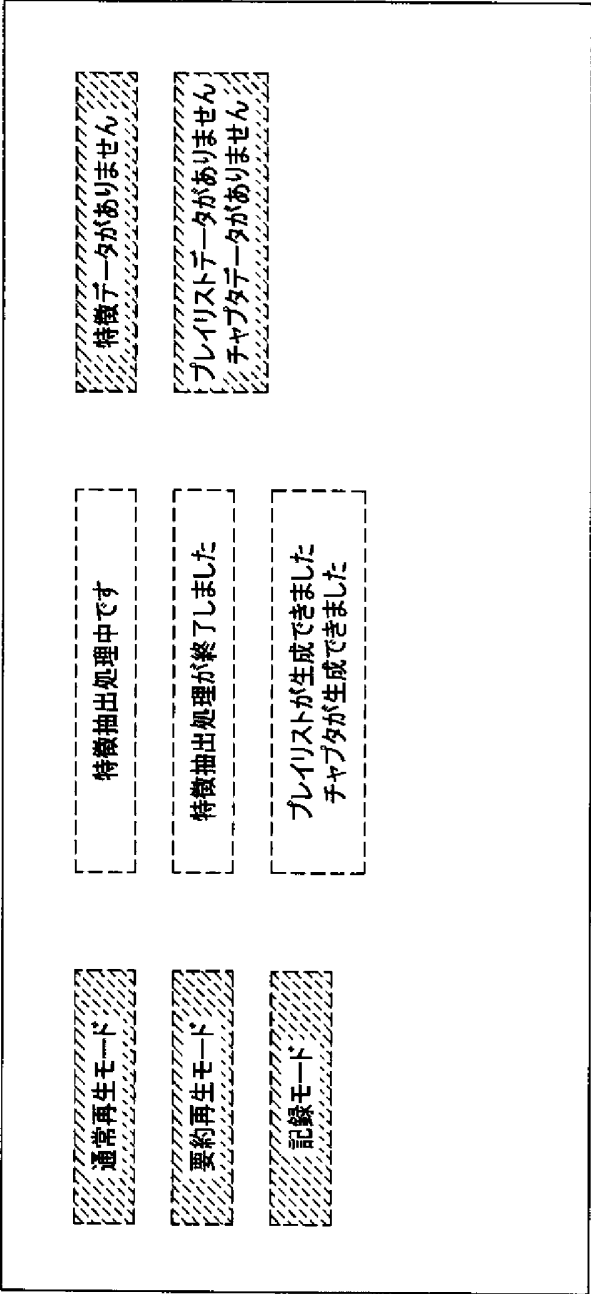


FIG.19

[図20]

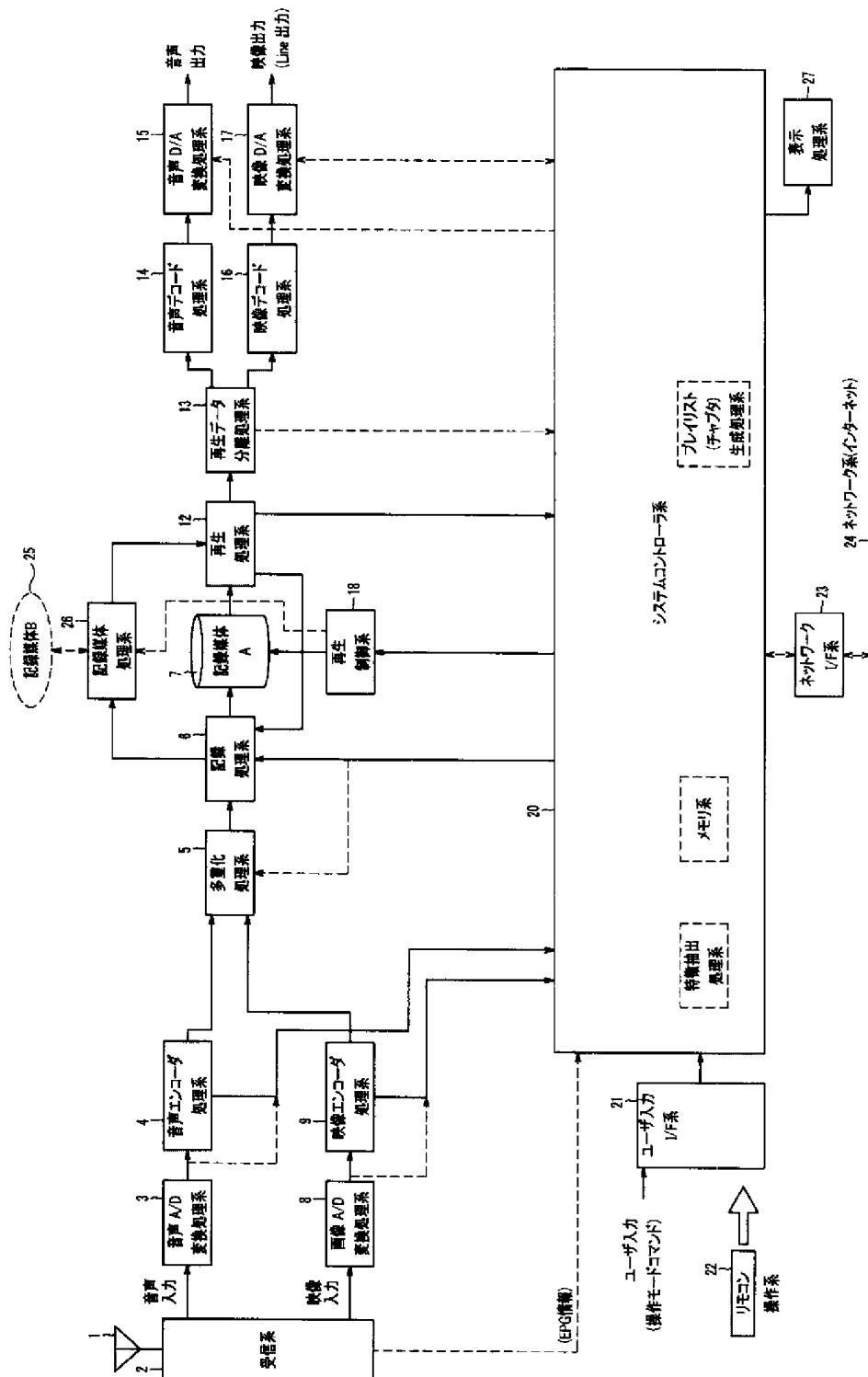


FIG.20

[図21]

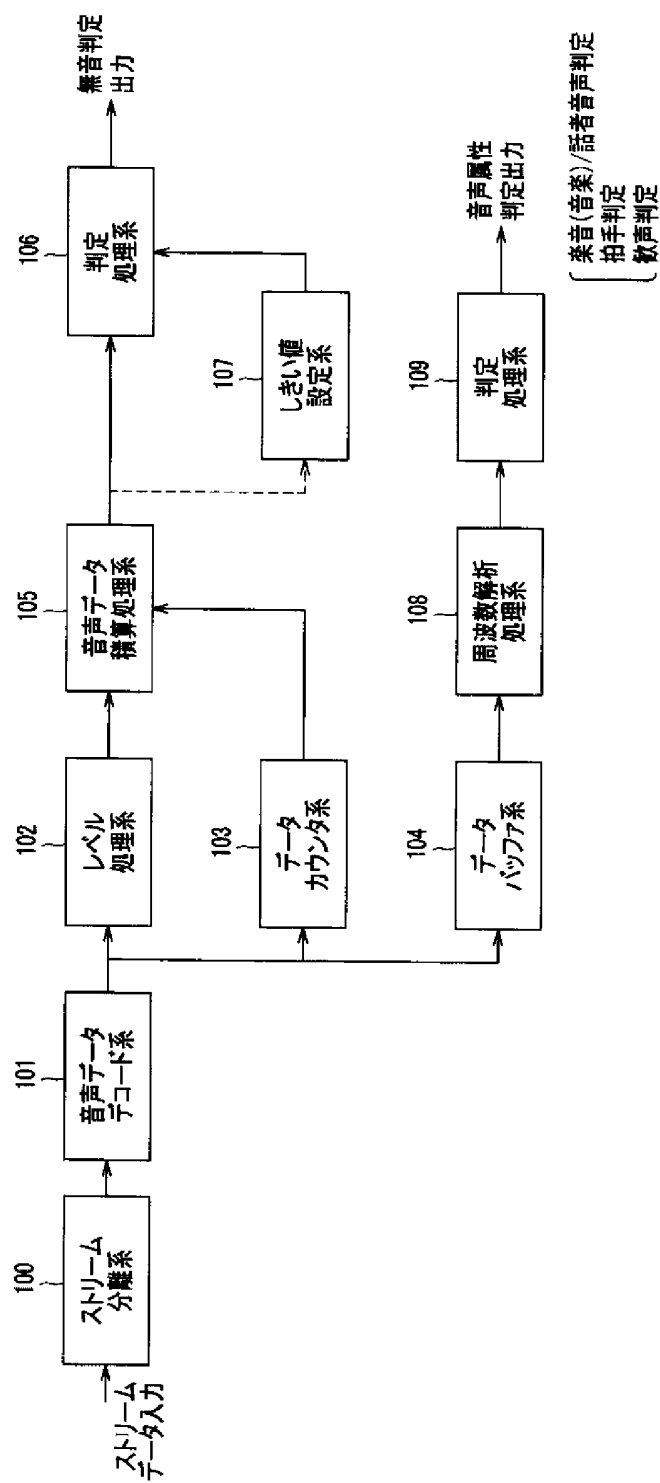


FIG.21

[図22]

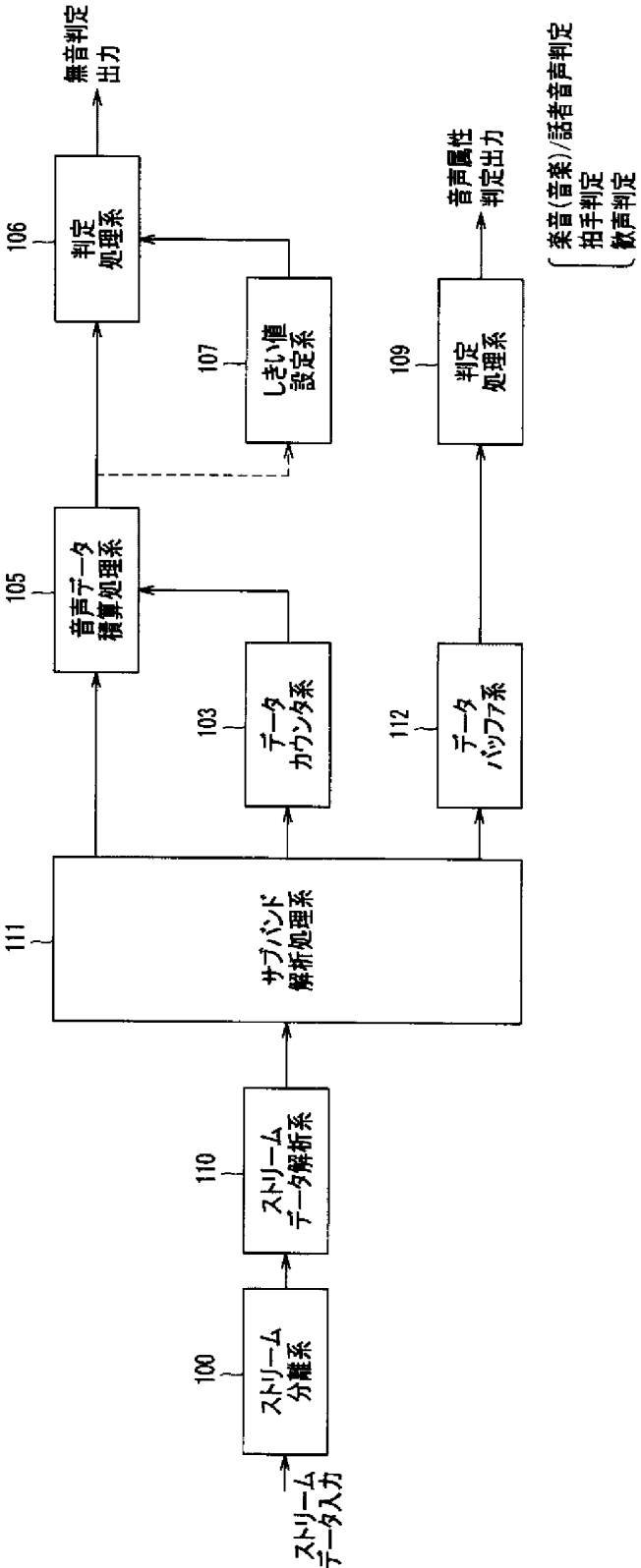


FIG.22

[図23]

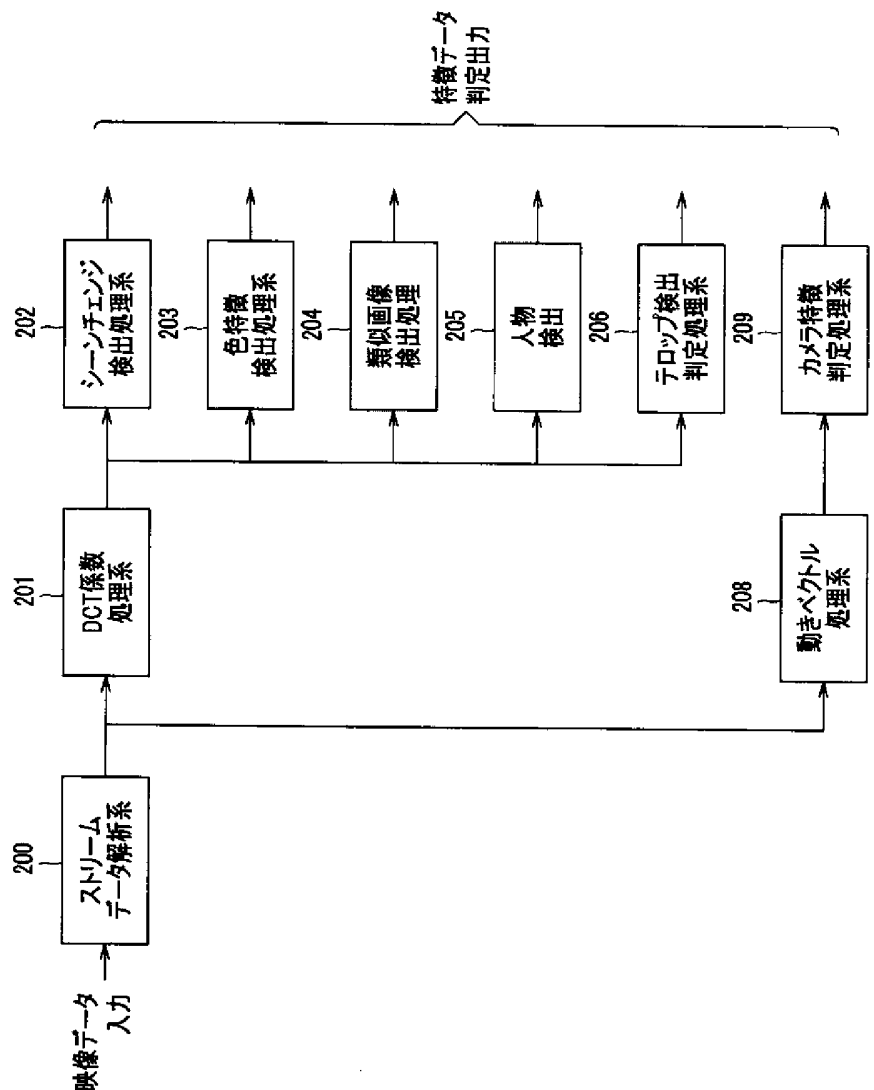


FIG.23

[図24]

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

FIG.24

[図25]

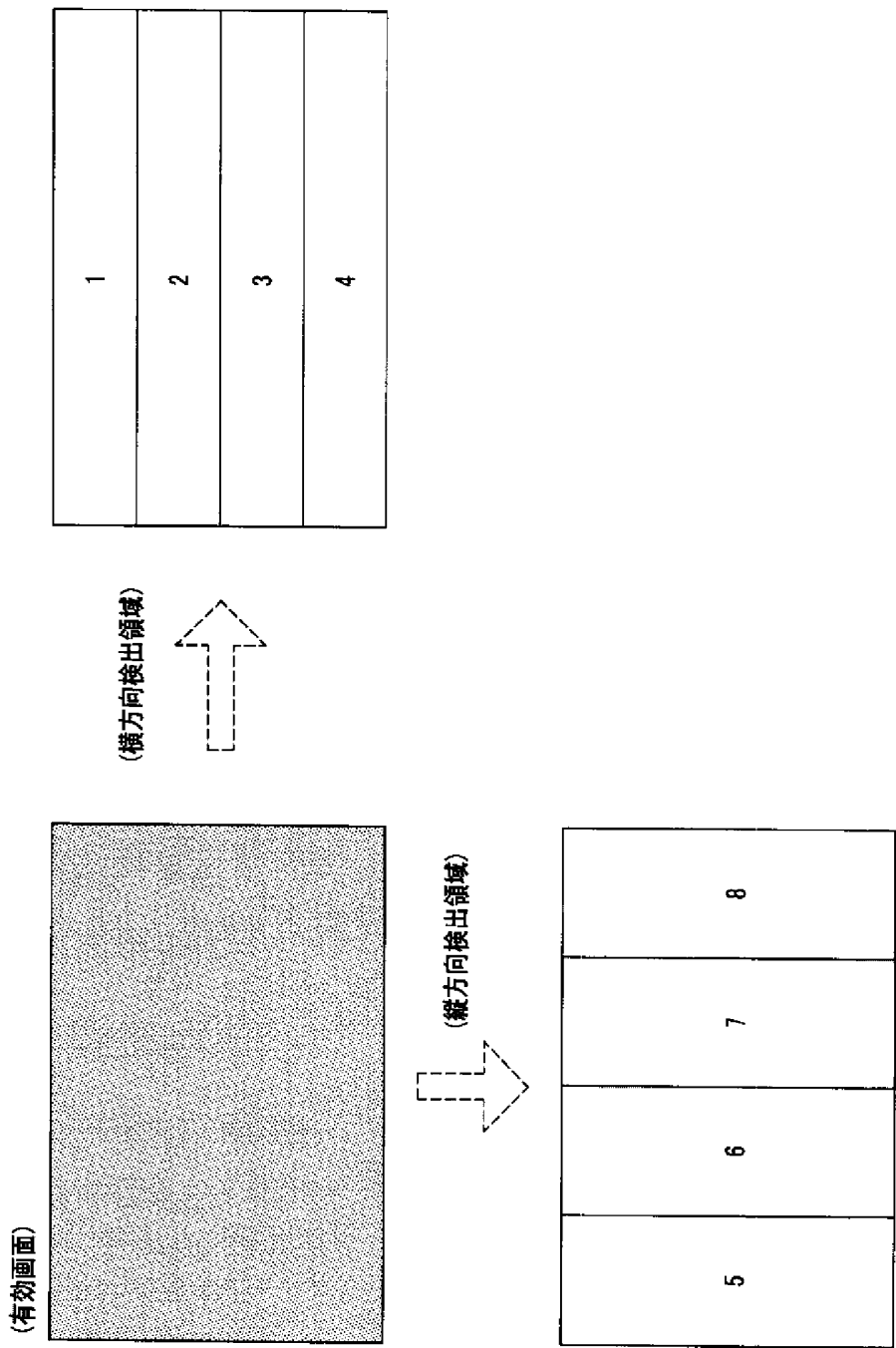


FIG.25

[図26]

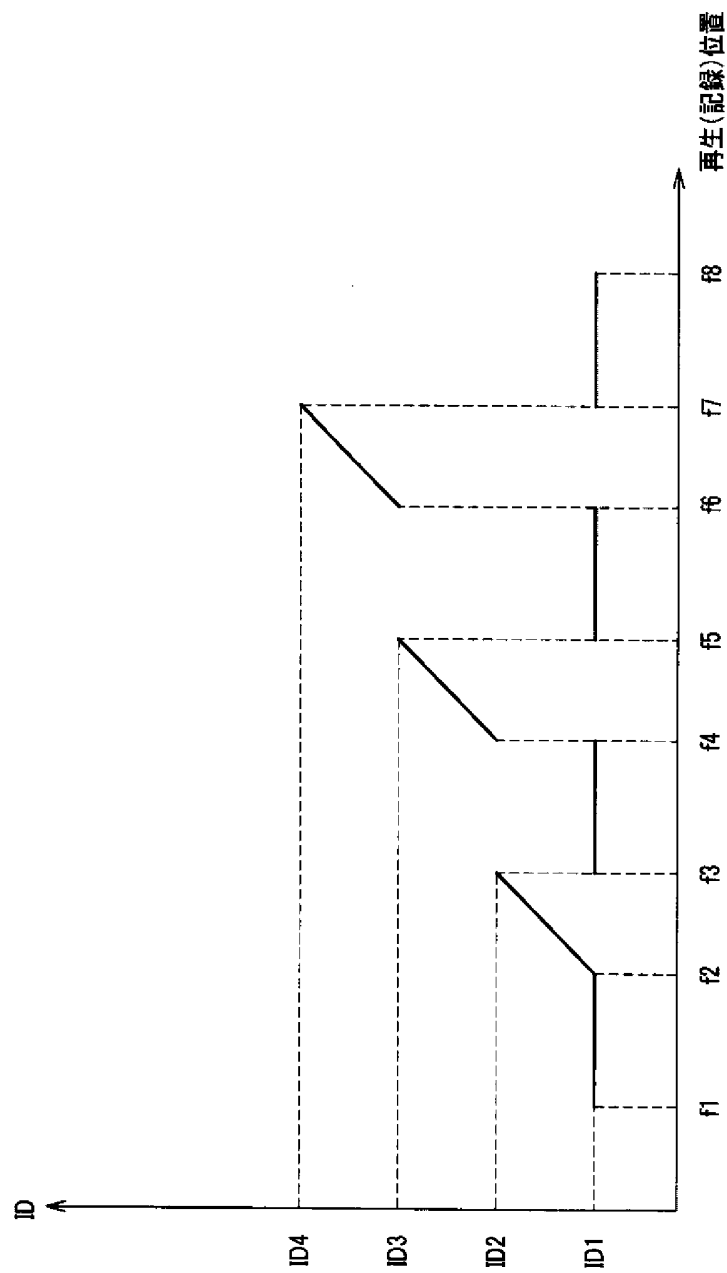


FIG.26

[図27]

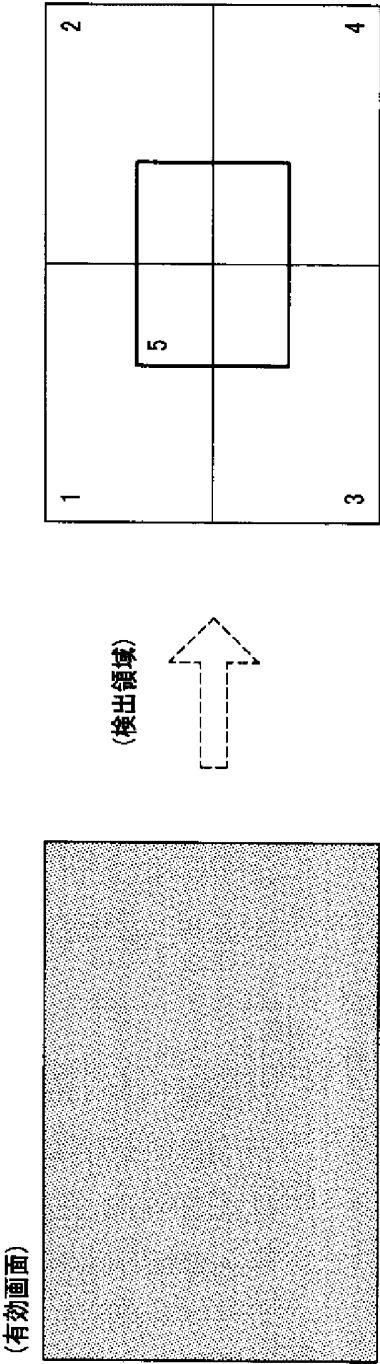


FIG.27

[図28]

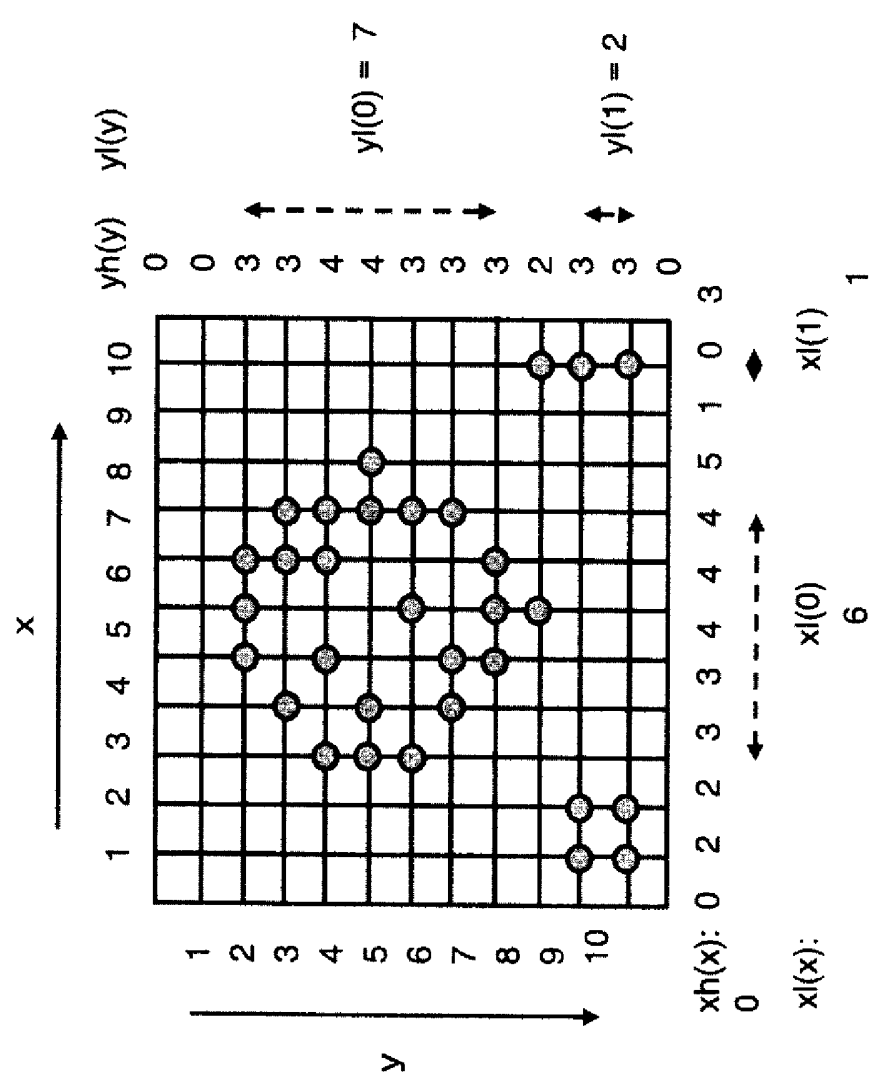


FIG.28

[図29]

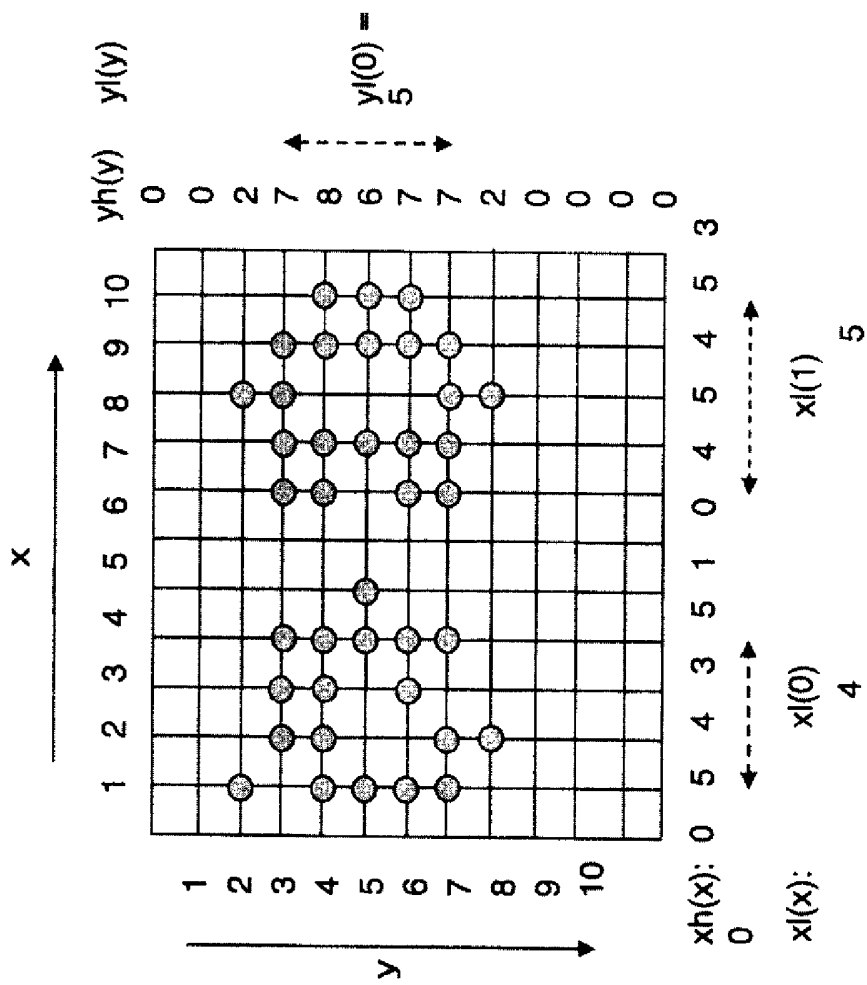


FIG.29

[図30]

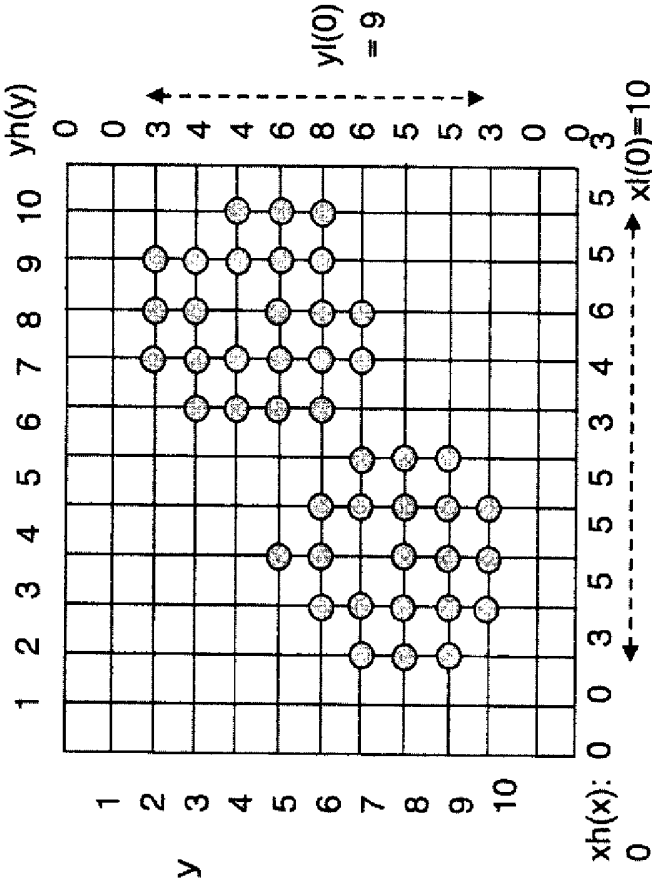


FIG.30

[図31]

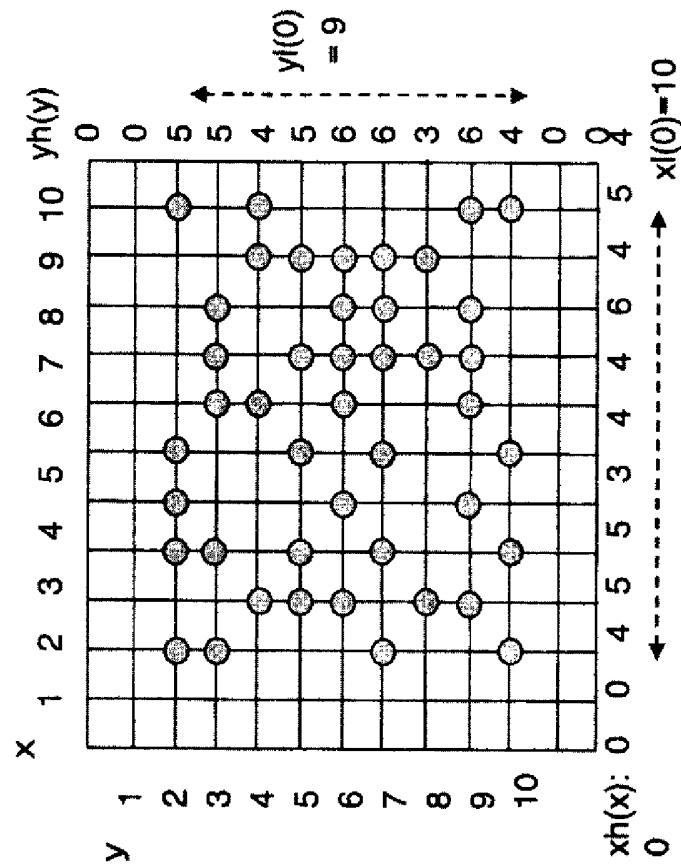


FIG.31

[図32]

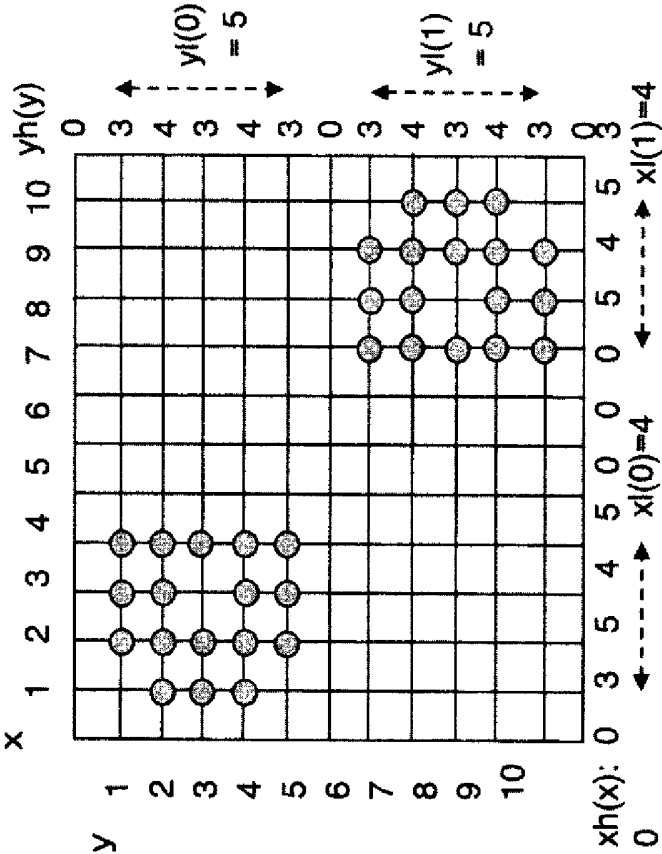
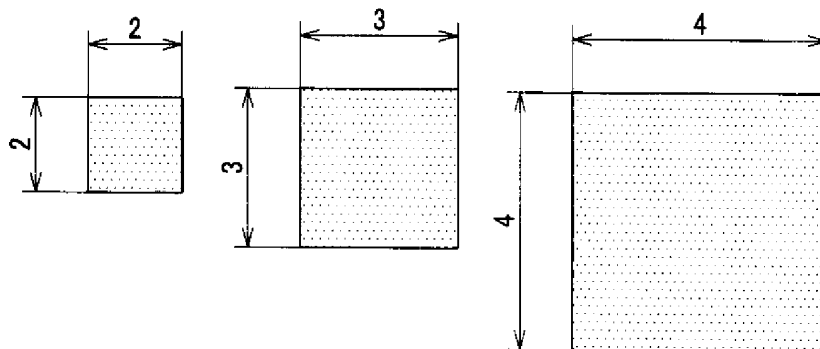


FIG.32

[図33]

**FIG.33**

[図34]

FIG.34A

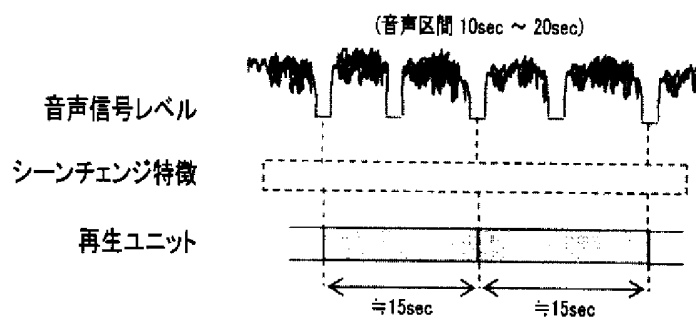


FIG.34B

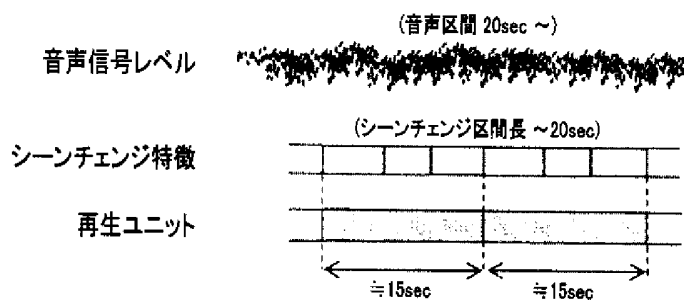


FIG.34C

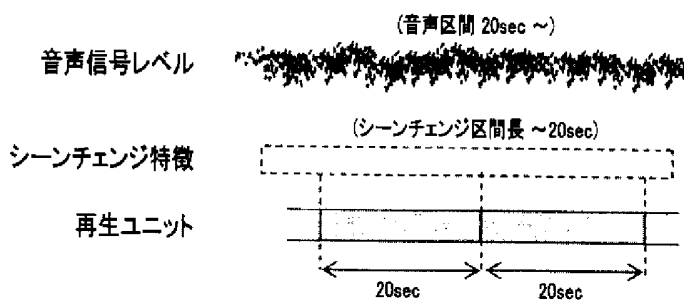


FIG.34D

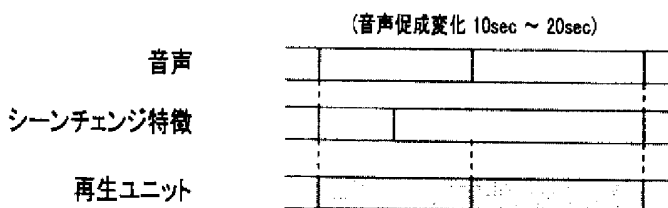
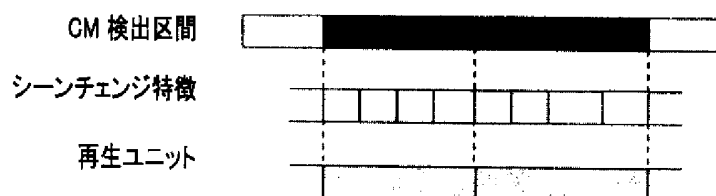


FIG.34E



[図35]

FIG.35A

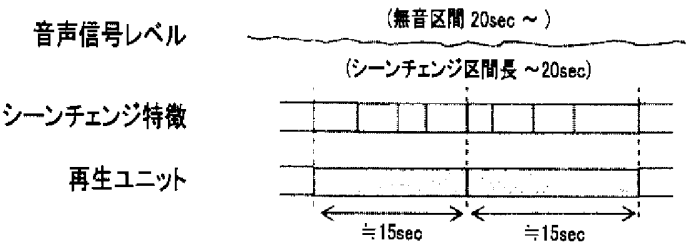
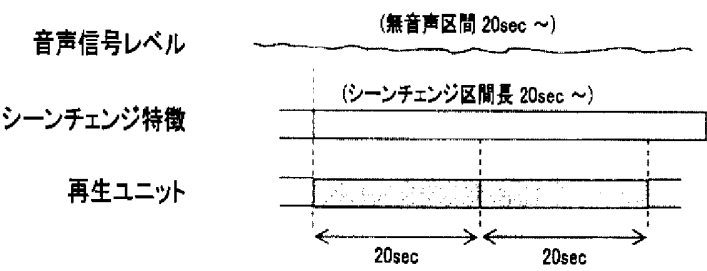


FIG.35B



[図36]

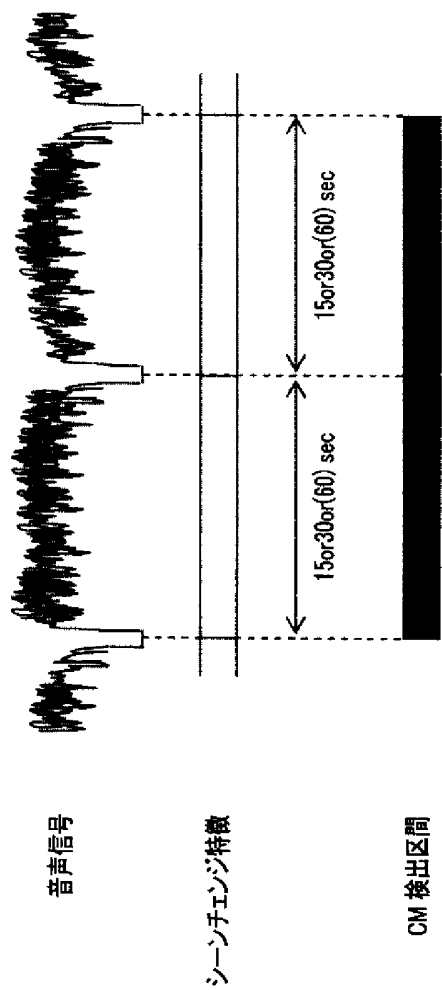


FIG.36

[図37]

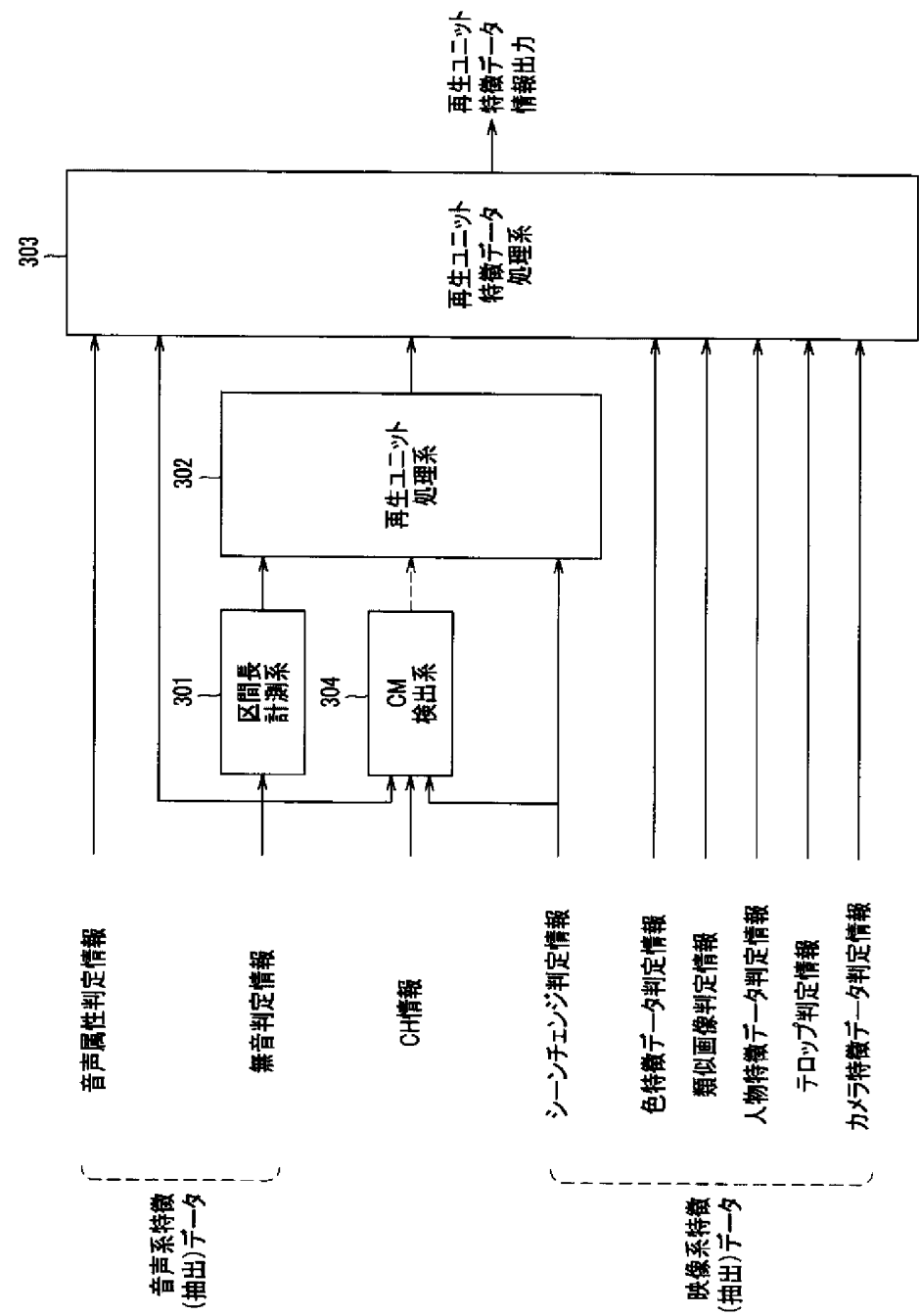


FIG.37

[図38]

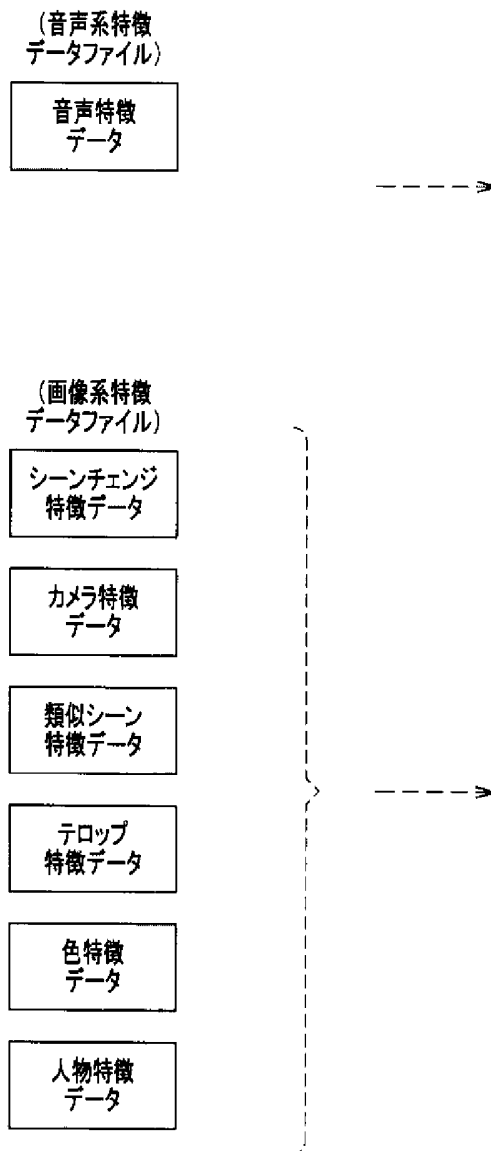


FIG.38

[図39]

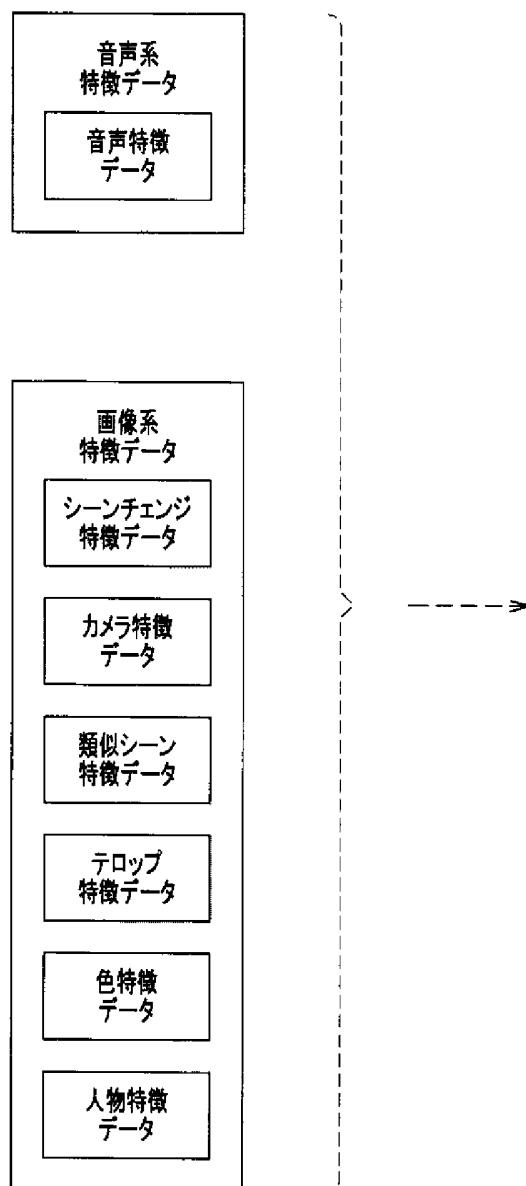
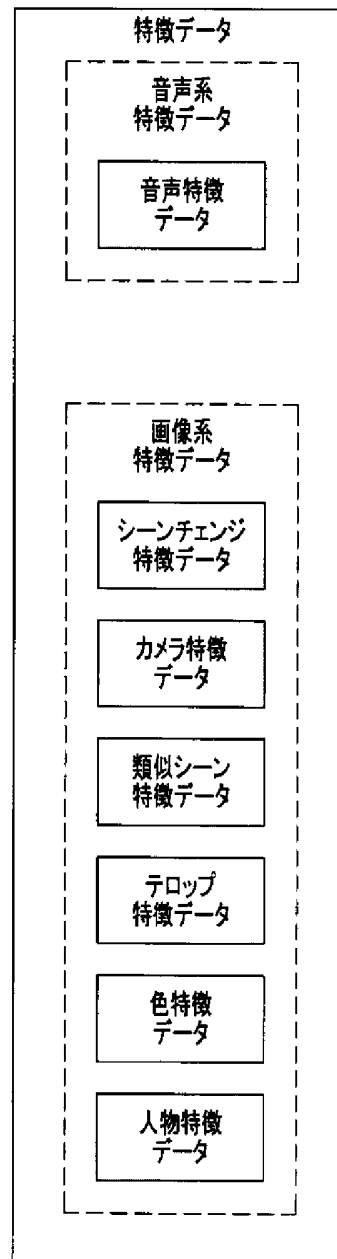


FIG.39

[図40]

**FIG.40**

[図41]

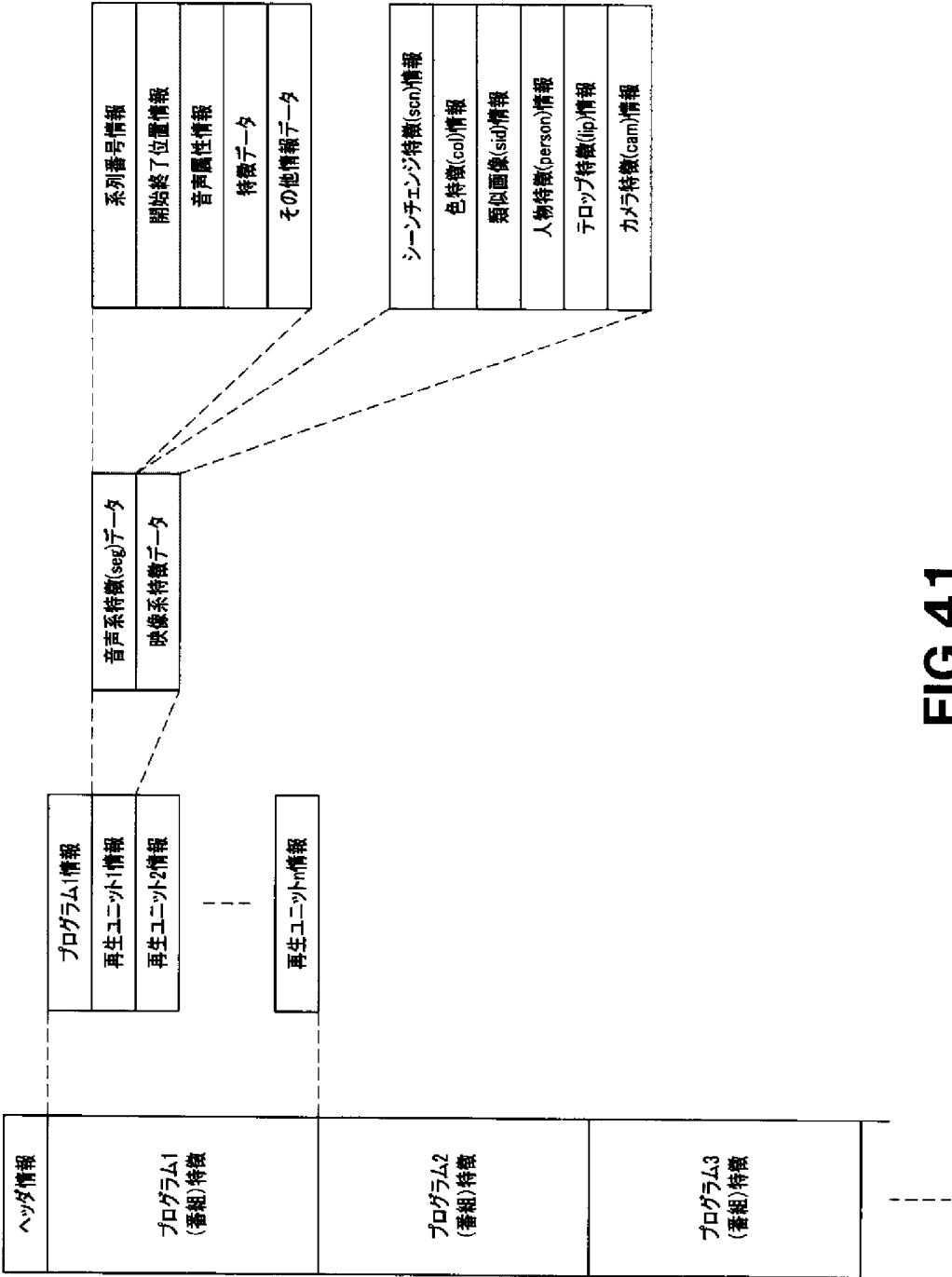


FIG.41

[図42]

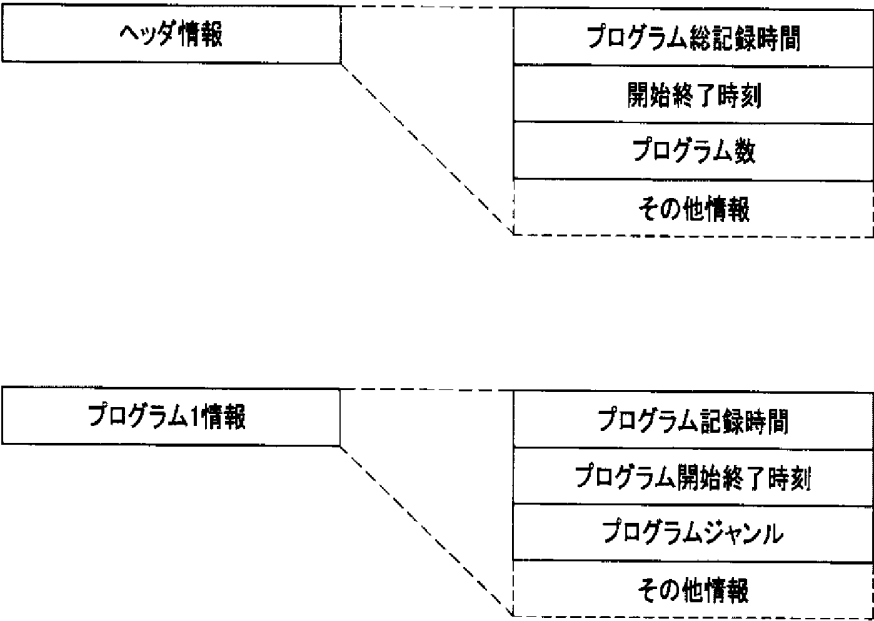


FIG.42

[図43]

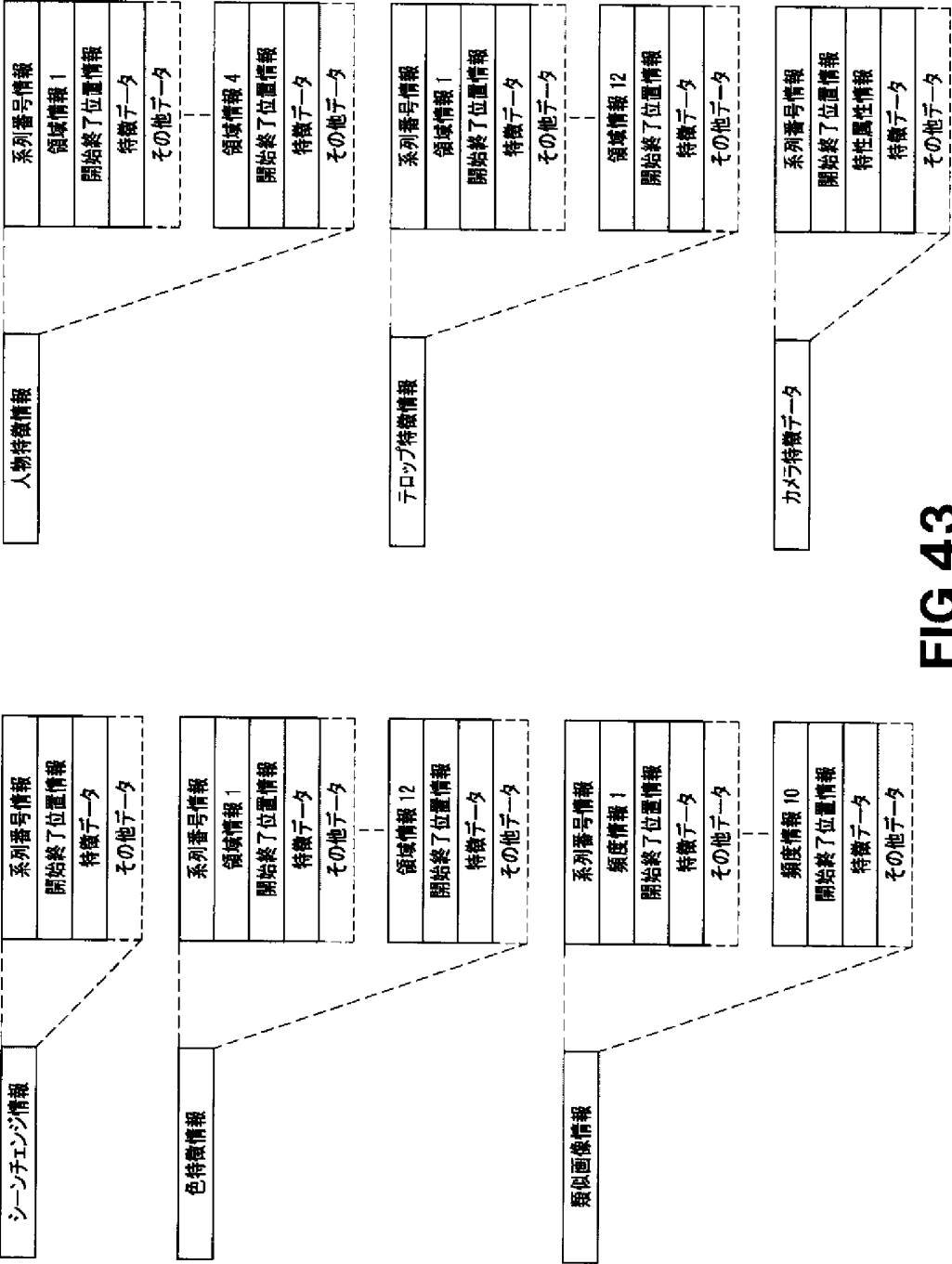


FIG.43

[図44]

(a)	(b)	(c)	(d)	(d)
0	229	0	b	0
230	1000	100	aabb	1
1001	1099	30	aa	0
1100	2600	100	aabb	1
2601	2999	30	aa	0
3000	3500	70	ca	1

(a)	(b)	(c)	(d)
230	1000	100	aabb
1100	2600	100	aabb
3000	3500	70	ca

FIG.44B

FIG.44A

[図45]

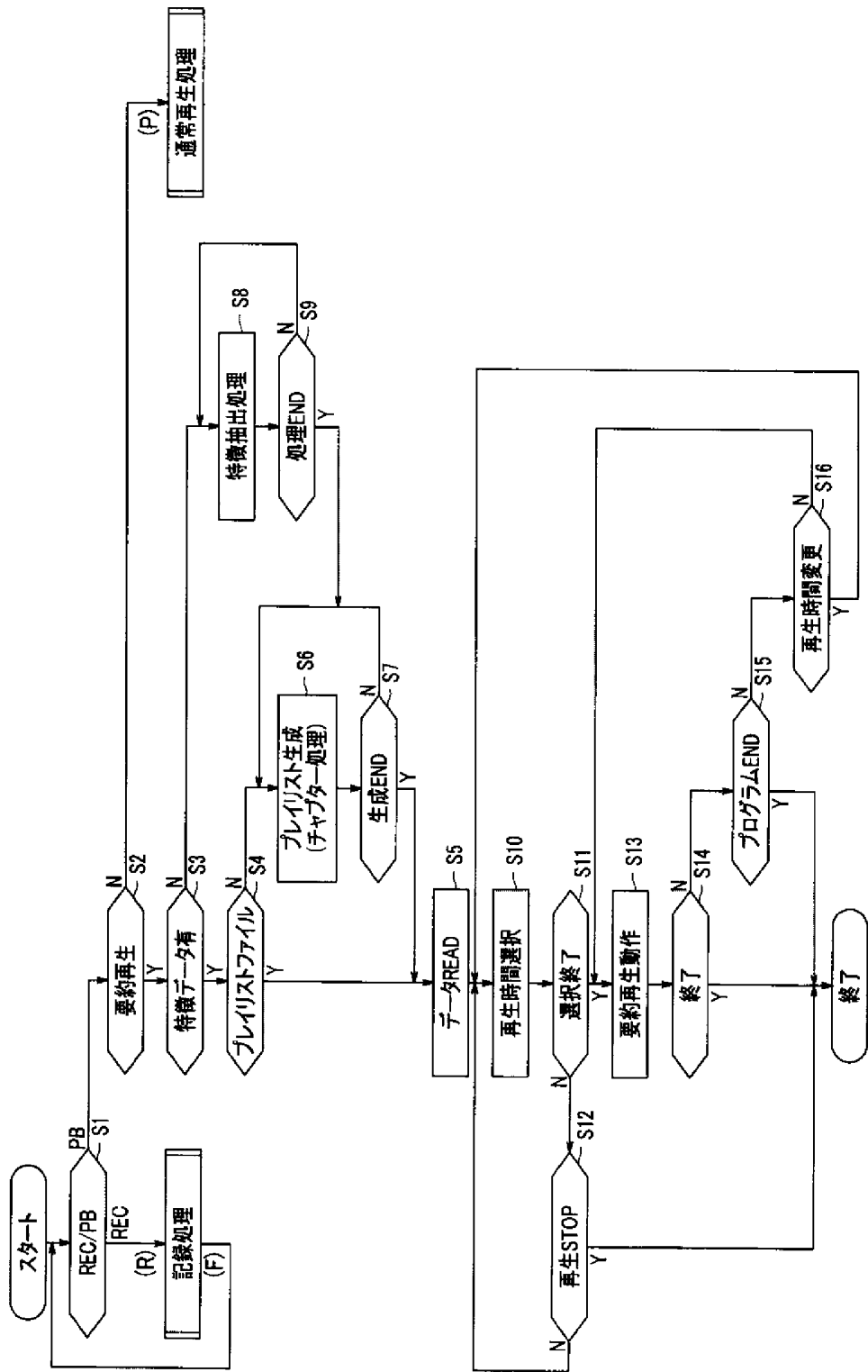


FIG.45

[図46]

	記録時間 Trec (所定プログラム記録時間)				
	Trec < 10分	10分 ≤ Trec < 30分	30分 ≤ Trec < 1時間	1時間 ≤ Trec < 2時間	Trec ≥ 2時間
選択可能 要約再生時間	Trec	Trec/5 Trec/2	Trec/6 Trec/2 Trec(2/3)	Trec/6 Trec/4 Trec/2 Trec(2/3)	Trec/12 Trec/6 Trec/4 Trec/2 Trec(2/3)
デフォルト 要約再生時間	Trec	Trec/5	Trec/2	Trec/2	Trec/4

FIG.46

[図47]

	記録時間 Trec (所定プログラム記録時間)				
	Trec < 10分	10分 ≤ Trec < 30分	30分 ≤ Trec < 1時間	1時間 ≤ Trec < 2時間	Trec ≥ 2時間
自動設定 チャプタ数	3 ~ 5	3 ~ 10	5 ~ 20	5 ~ 40	10 ~ 100

FIG.47

[図48]

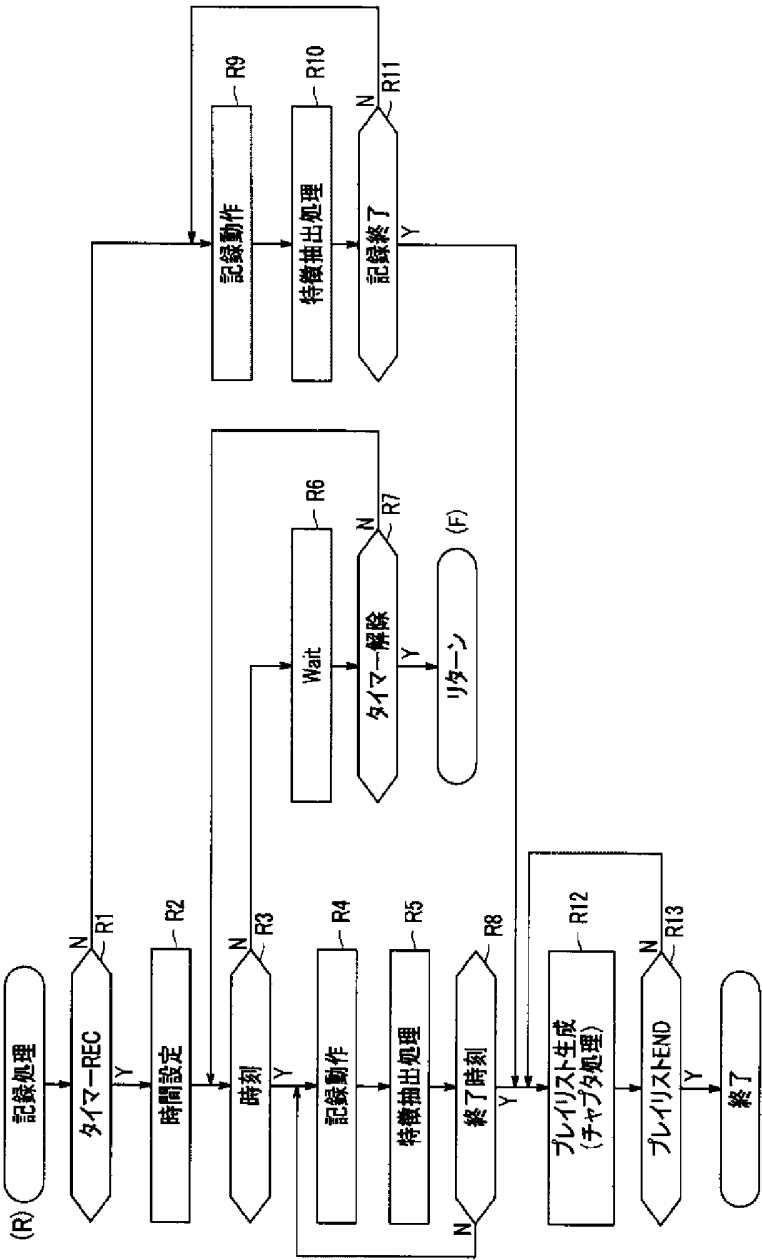


FIG.48

[図49]

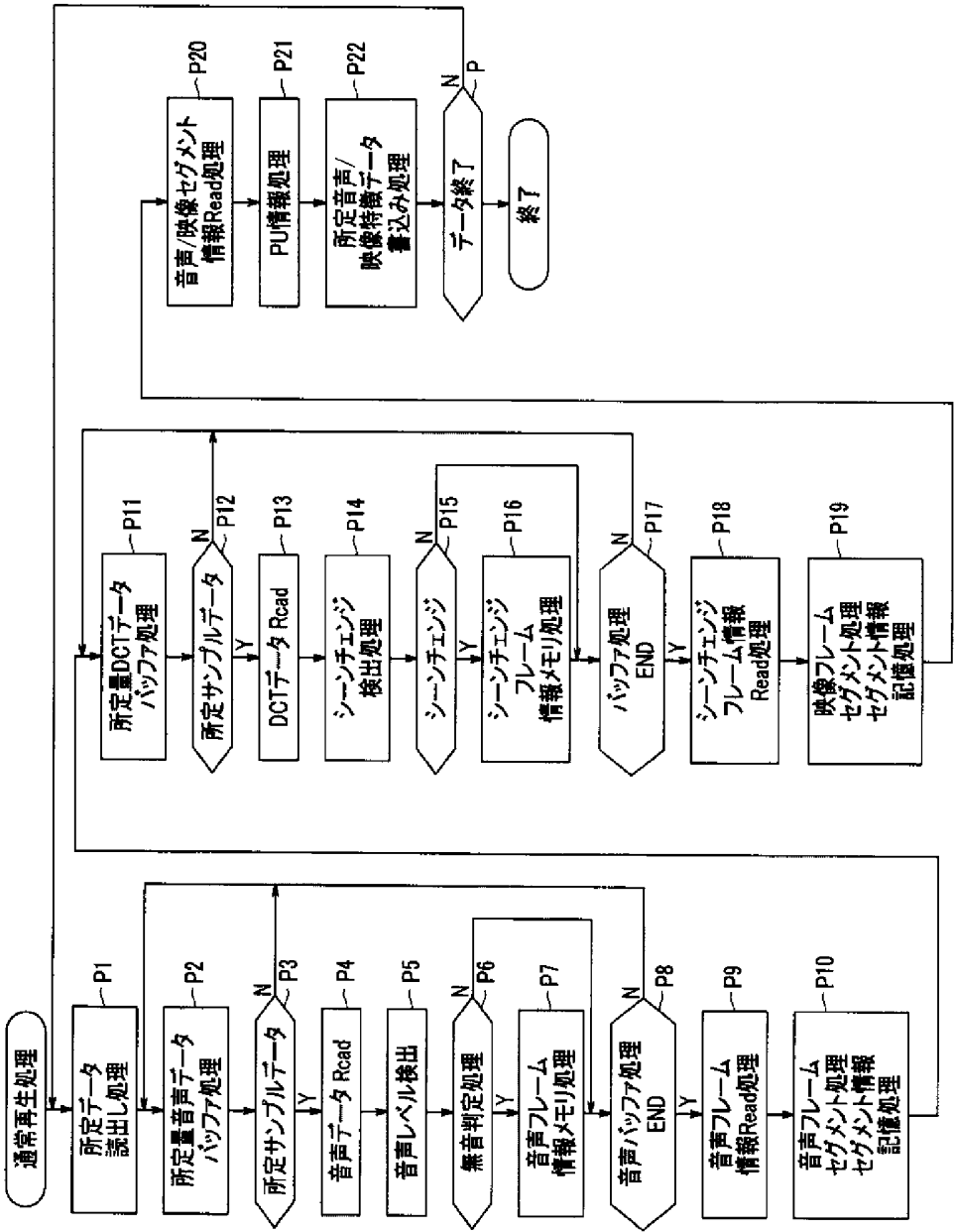


FIG.49

[図50]

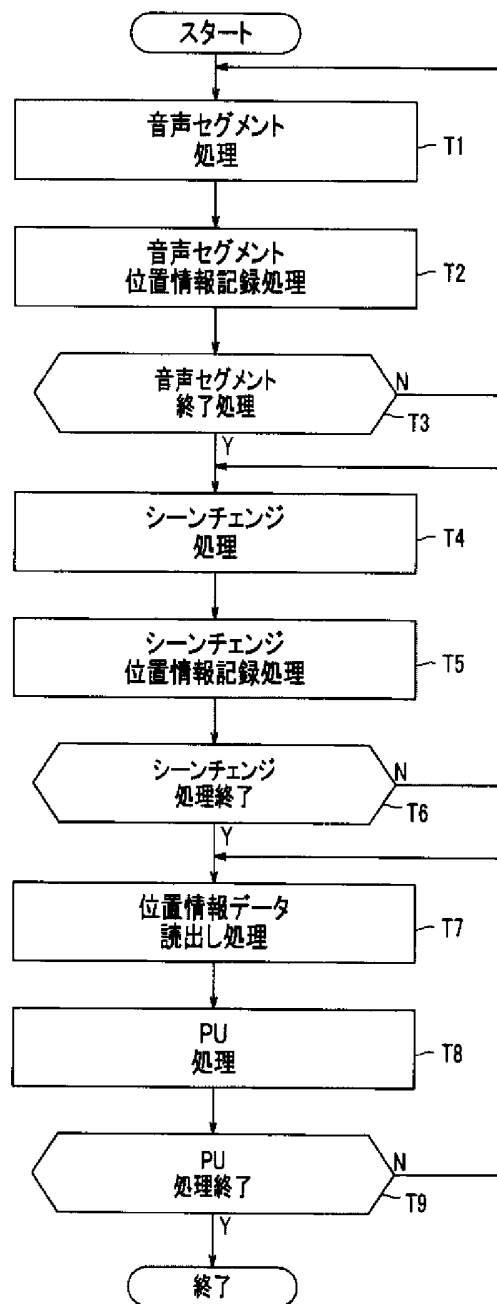


FIG.50

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2005/014598

A. CLASSIFICATION OF SUBJECT MATTER

H04N5/91 (2006.01), **G11B27/034** (2006.01), **G11B27/00** (2006.01)

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)

H04N5/91 (2006.01), **G11B27/034** (2006.01), **G11B27/00** (2006.01)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2005
Kokai Jitsuyo Shinan Koho	1971-2005	Toroku Jitsuyo Shinan Koho	1994-2005

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2002-142189 A (Canon Inc.), 17 May, 2002 (17.05.02), Full text; all drawings (Family: none)	1-35
A	JP 2003-283993 A (Sanyo Electric Co., Ltd.), 03 October, 2003 (03.10.03), Full text; all drawings (Family: none)	1-35
P, A	JP 2004-364183 A (Nippon Telegraph And Telephone Corp.), 24 December, 2004 (24.12.04), Full text; all drawings (Family: none)	1-35



Further documents are listed in the continuation of Box C.



See patent family 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 application or patent 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

07 November, 2005 (07.11.05)

Date of mailing of the international search report

15 November, 2005 (15.11.05)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. ~~H04N5/91~~ (2006.01), ~~G11B27/034~~ (2006.01), ~~G11B27/00~~ (2006.01)

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int.Cl. ~~H04N5/91~~ (2006.01), ~~G11B27/034~~ (2006.01), ~~G11B27/00~~ (2006.01)

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2005年
日本国実用新案登録公報	1996-2005年
日本国登録実用新案公報	1994-2005年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 2002-142189 A (キヤノン株式会社) 2002.05.17, 全文、全図 (ファミリーなし)	1-35
A	JP 2003-283993 A (三洋電機株式会社) 2003.10.03, 全文、全図 (ファミリーなし)	1-35
P, A	JP 2004-364183 A (日本電信電話株式会社) 2004.12.24, 全文、全図 (ファミリーなし)	1-35

☐ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

07.11.2005

国際調査報告の発送日

15.11.2005

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

梅岡 信幸

電話番号 03-3581-1101 内線 3541

5C

9075