

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第4660861号  
(P4660861)

(45) 発行日 平成23年3月30日(2011.3.30)

(24) 登録日 平成23年1月14日(2011.1.14)

(51) Int.Cl.

F 1

HO4N	5/91	(2006.01)	HO4N	5/91	N
HO4N	5/222	(2006.01)	HO4N	5/222	Z
G10L	11/00	(2006.01)	G10L	11/00	402K
HO4N	5/765	(2006.01)	HO4N	5/781	510F
HO4N	5/781	(2006.01)	GO6T	1/00	340A

請求項の数 6 (全 12 頁) 最終頁に続く

(21) 出願番号

特願2006-241465 (P2006-241465)

(22) 出願日

平成18年9月6日(2006.9.6)

(65) 公開番号

特開2008-66956 (P2008-66956A)

(43) 公開日

平成20年3月21日(2008.3.21)

審査請求日

平成21年3月5日(2009.3.5)

早期審査対象出願

(73) 特許権者 306037311

富士フィルム株式会社

東京都港区西麻布2丁目26番30号

(74) 代理人 100083116

弁理士 松浦 憲三

(72) 発明者 宮坂 恭正

神奈川県川崎市麻生区万福寺1丁目2番2号 新百合トウェンティワン8F 富士フィルムソフトウェア株式会社内

(72) 発明者 寺横 素

埼玉県朝霞市泉水3丁目11番46号 富士写真フィルム株式会社内

審査官 竹中 辰利

最終頁に続く

(54) 【発明の名称】楽曲画像シンクロ動画シナリオ生成方法、プログラムおよび装置

## (57) 【特許請求の範囲】

## 【請求項1】

楽曲の構成に関する特徴量及び前記楽曲の拍または裏拍に関する特徴量からなる楽曲の特徴量を抽出するステップと、

抽出された楽曲の構成に関する特徴量に基づいて楽曲の構成を抽出し、該抽出結果に基づいて楽曲を複数の構成部分に分割するステップと、

画像の特徴量を解析するステップと、

前記画像の特徴量が共通または類似する画像を同一の画像群に分類するステップと、

前記楽曲の構成に関する特徴量と前記分類された画像群の特徴量とに応じて、前記楽曲の構成部分と前記分類された画像群とを関連づけるステップと、

前記楽曲の各構成部分における拍または裏拍に関する特徴量と前記楽曲の各構成部分に関連づけられた画像群に含まれる各画像の特徴量とに応じて、前記楽曲の各構成部分の拍または裏拍と前記各画像とを関連づけるステップと、

関連づけられた楽曲と画像とを同期して再生可能な動画シナリオを生成するステップと、

を含む楽曲画像シンクロ動画シナリオ生成方法。

## 【請求項2】

前記画像の特徴量は、画像のボケまたはブレに関する評価値を含み、

前記評価値に基づいて、前記楽曲の構成部分に関連づけるべき画像を排除するステップと、

10

20

をさらに含む請求項<sub>1</sub>に記載の楽曲画像シンクロ動画シナリオ生成方法。

【請求項3】

前記動画シナリオに基づいて、楽曲に同期して再生可能な動画を生成するステップをさらに含む請求項<sub>1</sub>または<sub>2</sub>に記載の楽曲画像シンクロ動画シナリオ生成方法。

【請求項4】

前記生成された動画を再生するステップをさらに含む請求項<sub>3</sub>に記載の楽曲画像シンクロ動画シナリオ生成方法。

【請求項5】

請求項<sub>1</sub>～<sub>4</sub>のいずれかに記載の楽曲画像シンクロ動画シナリオ生成方法を演算装置に実行させるための楽曲画像シンクロ動画シナリオ生成プログラム。

10

【請求項6】

請求項<sub>5</sub>に記載の楽曲画像シンクロ動画シナリオ生成プログラムを記憶する記憶部と、前記記憶部の楽曲画像シンクロ動画シナリオ生成プログラムを実行する演算装置と、を備える楽曲画像シンクロ動画シナリオ生成装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は楽曲にシンクロした動画を生成する技術に関する。

【背景技術】

【0002】

引用文献1は、音楽データからテンポ情報を取り出し、そのテンポ情報から画像データに対してカット編集を施すことで、ビデオ編集を行う。

20

【0003】

引用文献2は、入力素材から記述子を分析し、それをもとにスタイル情報を定義し、入力素材・入力素材から得た記述子・スタイル情報を基に作業セットを作成し、それを実行することで出力データを生成するシステムである。入力・出力には少なくとも1つの映像・オーディオ各データが含まれており、記述子の取得は、入力素材を分析することで得られる。また、記述子はユーザからも受け取ることができる。

【特許文献1】特開2004-96617号公報

30

【特許文献2】特表2005-506643号公報

【発明の開示】

【発明が解決しようとする課題】

【0004】

動画全体の流れにストーリーを持たせる、タイミングよく楽曲にマッチした画像を出力させるなど、画像の表示タイミングと楽曲の相関は、出力動画の質を大きく左右する。

【0005】

ところが、引用文献1では、楽曲からテンポ情報のみを使用してビデオを編集する方法を示しており、テンポ情報以外の楽曲の特徴量を使用する方法は明示されていない。また、引用文献2では、楽曲にシンクロしたムービーを作成できるが、どのタイミングでどの画像を出力させるかは明示されていない。

40

【0006】

そこで、本発明は、単に楽曲の特徴点にシンクロするだけでなく、画像の属性にも同期性を持たせ、出力動画のクオリティを大幅に向上させることを目的とする。

【課題を解決するための手段】

【0007】

上記課題を解決するため、本発明に係る楽曲画像シンクロ動画シナリオ生成方法は、楽曲のフレーズに関する特徴量及び楽曲の構成部分の拍または裏拍に関する特徴量からなる楽曲の特徴量を抽出するステップと、抽出された楽曲の特徴量に基づいて楽曲の構成を抽出し、該抽出結果に基づいて楽曲を複数の構成部分に分割するステップと、画像の特徴量を解析するステップと、楽曲の構成部分に対応する特徴量と画像の特徴量とに応じて、樂

50

曲と画像とを関連づけるステップと、関連づけられた楽曲と画像とを同期して再生可能な動画シナリオを生成するステップと、を含む。

【0008】

この発明によると、画像の特徴量に応じて楽曲の構成部分と画像とが関連づけられるから、従来に比して、楽曲の再生内容とマッチした画像の同期再生が可能となる。

【0009】

特徴量が共通または類似する画像を同一の画像群に分類するステップと、楽曲のフレーズに関する特徴量と画像群の特徴量とに応じて、楽曲のフレーズと画像群とを関連づけるステップと、楽曲の構成部分の拍または裏拍に関する特徴量と楽曲の構成部分に関連づけられた画像群に含まれる各画像の特徴量とに応じて、楽曲の構成部分の拍または裏拍と画像とを関連づけるステップと、をさらに含んでもよい。10

【0010】

画像の特徴量は、画像のボケまたはブレに関する評価値を含み、評価値に基づいて、楽曲の構成部分に関連づけるべき画像を排除するステップと、をさらに含んでもよい。

【0011】

動画シナリオに基づいて、楽曲に同期して再生可能な動画を生成するステップをさらに含んでもよい。

【0012】

生成された動画を再生するステップをさらに含んでもよい。

【0013】

上記の楽曲画像シンクロ動画シナリオ生成方法を演算装置に実行させるための楽曲画像シンクロ動画シナリオ生成プログラムも本発明に含まれる。20

【0014】

上記の楽曲画像シンクロ動画シナリオ生成プログラムを記憶する記憶部と、記憶部の楽曲画像シンクロ動画シナリオ生成プログラムを実行する演算装置と、を備える楽曲画像シンクロ動画シナリオ生成装置も本発明に含まれる。

【発明の効果】

【0015】

この発明によると、画像の特徴量に応じて楽曲の構成部分と画像とが関連づけられるから、従来に比して、楽曲の再生内容とマッチした画像の同期再生が可能となる。30

【発明を実施するための最良の形態】

【0016】

<第1実施形態>

図1は本発明の好ましい実施形態に係るシナリオ作成装置のブロック図である。

【0017】

画像入力装置11は、シナリオを生成するために必要な画像群を入力するものであり、例えば、メモリカードリーダーなどが考えられる。

【0018】

楽曲入力装置12は、シナリオを生成するために必要な楽曲を入力するものであり、例えば、メモリカードリーダーなどが考えられる。40

【0019】

入力装置13は、ユーザがシステムの各種操作を行うためのものであり、例えば、キーボードやマウスなどが考えられる。

【0020】

シナリオ生成装置14は、入力された楽曲を解析したり、シナリオを生成したり、画像を選択したりするものである。

【0021】

図2は、シナリオ生成装置14に含まれる画像処理装置20のブロック図である。画像処理装置20は主として、複数の処理モジュールを含む個別処理部21、全体処理部22、および制御部23を備える。50

## 【0022】

個別処理部21は、画像または画像に付帯する各種情報（撮影日時等）の少なくともいずれか一方を用いて特徴量もしくは個別評価値を算出する機能を備える。ここで、特徴量とは、画像もしくは画像群の入力に対して、一意的・絶対的な結果が得られる情報（例えば、ある画像のエッジの強さを表すボケ情報）を指す。また、個別評価値とは、画像もしくは画像群の入力に対して、ユーザの好みにより結論が変わりうる多義的・相対的な情報（例えば、ボケの観点から当該画像の評価度合を表すボケ評価値）を指す。以下の表1に、個別処理部の例を記述する。

## 【0023】

【表1】

10

個別処理部の例	機能の説明
イベントカテゴライズ部	画像付帯情報に含まれる画像撮影時刻を用いて画像群をグループ分けし、画像群のイベント情報（特徴量の一例）を算出する
イベント重要度算出部	イベント情報や類似グループ情報（特徴量）などを用いてイベント重要度（個別評価値の一例）を算出する
類似判定部	複数の画像から当該画像同士の類似度（特徴量）を算出する
類似画像グルーピング	画像群の類似度を用いて類似画像グループ情報（特徴量）を算出する
顔検出部	画像から当該画像中に写っている人物の顔を検出し、顔情報（特徴量）を算出する
顔評価部	顔情報から当該画像の顔の観点からの評価値を表す顔評価値（個別評価値）を算出する
明るさ判定部	画像から当該画像の明るさ情報（特徴量）を算出する
ボケブレ判定部	画像から当該画像のボケブレ情報（特徴量）を算出する
ボケブレ評価値	明るさ情報から当該画像のボケないしほれの観点からの評価値を表すボケブレ評価値（個別評価値）を算出する

20

30

全体処理部22は、上記特徴量のうちの一部または全部の合計、もしくは上記個別評価値のうちの一部または全部の合計、または上記特徴量および上記個別評価値のうちの一部または全部の合計をもとに、画像全体評価値を算出する機能を備える。

40

## 【0024】

例えば、イベント情報（特徴量）、顔評価値（個別評価値）、明るさ評価値（個別評価値）の3個の数値を元に特定の画像の画像全体評価値（プリントなど特定の用途に向か否かを表す数値）を算出する。

## 【0025】

ただし、本実施形態では、画像のプリントは前提としておらず、画像の画像全体の評価を行う必要がない場合は、全体処理部22は設けなくてもよい。

## 【0026】

50

制御部23は、画像処理装置20と外部機器（他の画像処理装置やオペレータの操作対象となる操作装置）とのインターフェースを持つほか、個別処理部21や全体処理部22への処理の制御を受け持つ。

#### 【0027】

以下、図3のフローチャートを参照し、動画シナリオ作成装置の実行する動画シナリオ生成処理の流れを説明する。

#### 【0028】

S1では、楽曲から特徴量を抽出する。楽曲の特徴量とは、楽曲の拍や裏拍、アクセント、テンポの変化点などを指す。楽曲の特徴量を抽出する方法は、従来の手法で可能であり、例えば、特開2003-263162号公報、特開2005-027751号公報などが挙げられる。10

#### 【0029】

S2では、楽曲の構成（これも楽曲の特徴量に含めてよい）を抽出し、抽出された構成に従って、楽曲を複数の部分に分割する。例えば、以下のような分割方法が考えられる。

#### 【0030】

（1）テンポの変わり目で楽曲の構成を区切る。

#### 【0031】

（2）拍の頭から8拍を1構成として区切る。

#### 【0032】

（3）Aメロ、Bメロ、サビなど一般的に「フレーズ」と呼ばれているものに分ける。20

#### 【0033】

（4）ある閾値以上のテンポを持つ曲に対しては、4拍を1構成とし、ある閾値以下のテンポを持つ曲に対しては、8拍を1構成として区切る。

#### 【0034】

なお、フレーズを取得する方法は、従来の手法で可能であり、例えば、特開平9-90978号公報、特開2004-233965号公報などが挙げられる。

#### 【0035】

S3では、画像を解析する。画像の解析は、画像処理装置20によって実施される。

#### 【0036】

S4では、S2/S3の解析結果に基づき、楽曲の構成部分と画像群とを関連づける。30  
すなわち、S2で求めた楽曲構成に関する特徴量（楽曲における構成部分の出現順序や、Aメロ、Bメロ、サビなど）と、S3で求めた画像群の特徴量（特定の画像群に属する画像に共通するイベント情報や撮影日など）とを照合し、その照合結果に基づいて、両者を関連づける。関連づけの具体例は次のようなものが考えられる。

#### 【0037】

例1）同日の撮影日時の付帯情報を有する画像は同一の画像群に分類するなど、撮影されたイベントの異同に応じたカテゴライズによって時系列にグルーピングされた画像群を順番に、楽曲の構成要素の先頭からフレーズ1つ1つに当てはめていくことで、両者を関連づける（図4参照）。フレーズに対して、同一または類似の属性（撮影日時など）を持った画像群（類似画像群）を割り当てるので、楽曲の区切りに自然に画像が切り替わる動画を再生するシナリオを作成することができる。40

#### 【0038】

例2）類似画像群の各々を楽曲構成部分の各々に関連づける（図5参照）。類似画像が楽曲構成部分ごとに切り替わるので、メリハリのある動画を再生するシナリオを生成することができる。

#### 【0039】

例3）同じフレーズには同じ画像群を関連づける（図6参照）。同じ画像群を反復利用することで、楽曲になじんだ動画を再生するシナリオを生成することができる。

#### 【0040】

例4）楽曲のサビに相応しい所定の特徴量を有する画像群をサビに関連づける（図7参50

照)。

#### 【0041】

例えば、サビに相応しい特徴を有する画像群としては、以下のようなものが考えられる。

#### 【0042】

(1) 輝度の最も高い画像群

(2) 直近のイベント(あるいは最新の撮影日時)で撮影された画像群

(3) 予め登録されているユーザが被写体として撮影されている画像群

上記のような例のいずれかあるいは一部または全部の組み合わせで、サビに関連づけるのに相応しい特徴を有する画像群を特定し、これをサビと関連づける。

10

#### 【0043】

S5では、S1で求めた拍または裏拍と、S3で求めた個々の画像の特徴量を照合し、その結果に従って、各構成部分における拍または裏拍と、その構成部分に関連づけられている画像群に属する個々の画像とを関連づける。具体例は以下のようなものが考えられる。

#### 【0044】

例1) 拍や裏拍に合わせて、対応する画像群の画像の各々を順次表示させるように関連づける。

#### 【0045】

例2) 例1の方法に加えて、フレーズ内の拍数より画像群の画像枚数の方が少ない場合は、特定の画像あるいはランダムに選択された画像を繰り返し使用する。

20

#### 【0046】

例3) 顔が抽出されている画像の属する画像群(顔画像群)に対しては、拍や裏拍に合わせて、顔に対するズームアウトやズームインといった表示効果(エフェクト)を取り入れる。

#### 【0047】

例4) アクセントとなる拍に対しては、その拍が含まれる楽曲構成に対応する画像群のうち最も明るい画像を特定して関連づける。

#### 【0048】

例5) 楽曲から抽出された全ての拍を使うのではなく、パターンを決め、それに合わせて、画像を割り当てる。例えば図8に示すように、楽曲構成内を8拍で切り、1拍目、3拍目、6拍目の裏拍、8拍目に画像を表示させる。このようなパターンを増やし、繰り返すことによって、単調な画像切り替えを避けられる。上記のような例を組み合わせて楽曲の構成部分全てに対し、画像を割り当てる。

30

#### 【0049】

S6では、動画再生シナリオを作成する。シナリオの形式は特に問われず、例えば図9に示したようなXML(Extensible Markup Language)でもよいし、あるいはSMIL(Synchronized Multimedia Integration Language)でもよいし、バイナリ形式であってもよい。シナリオには画像の選択方法や選択された画像の表示時間等が記述されており、これに基づいて画像を表示させることにより、楽曲と同期した動画を再生することができる。

40

#### 【0050】

<第2実施形態>

図10は第2実施形態に係る動画シナリオ生成処理の流れを示すフローチャートである。

#### 【0051】

S11～S13、S15～S17は、それぞれ第1実施形態のS1～S3、S4～S6と同様である。

#### 【0052】

ただし、S14では、S13の画像解析から得たボケブレ評価値が、所定の値より高いものを、楽曲の構成部分に関連づけるべき画像群から排除する。S16の処理では、S1

50

3で排除された画像の関連づけを行わない。

【0053】

このように、ボケやブレのあるクオリティの低い画像を動画から排除できるので、必然的にクオリティの高い動画を再生するシナリオを生成することができる。

【0054】

なお、画像を排除した結果、画像が不足した場合は、ボケブレ評価値の低い画像を繰り返し使うなどして補うとよい。

【0055】

<第3実施形態>

上述のようにして作成されたシナリオに基づいて、動画を作成する装置も本発明に含まれる。 10

【0056】

図11は本発明の好ましい実施形態に係る動画作成装置のブロック図である。

【0057】

この装置は、シナリオ生成装置(図1)と同様、画像入力装置11、楽曲入力装置12、入力装置13、シナリオ生成装置14を備えており、さらに、動画生成装置15を備えている。

【0058】

図12は動画生成処理の流れを示すフローチャートである。

【0059】

S21では、シナリオ生成装置14によってシナリオを生成する(例えば図9のようなXMLファイル)。 20

【0060】

S22では、動画生成装置15は、画像入力装置11・楽曲入力装置12にそれぞれ入力された画像および楽曲のうち、シナリオに指定されている、動画生成に必要な画像および楽曲を取得する。シナリオに指定されているデータは、例えばPC上に保存されているデータであったり、Web上に公開されているデータである。

【0061】

そしてS23では、動画生成装置15は、シナリオの記述に基づき、取得した画像および楽曲から、動画を生成する。具体的には、取得した全ての画像データに対して、ある瞬間でのエフェクトや変更を考慮した画像処理を施し、オーダー順に重ね合わせて、フレーム画像を作成する。このフレーム画像を時間順につなげ合わせることで動画を作成する。作成された動画の圧縮記録方式は、アニメーションGIF、MPEGなど何でもよい。 30

【0062】

<第4実施形態>

上記のようにして作成された動画を再生出力する動画出力装置も本発明に含まれる。

【0063】

図13は本発明の好ましい実施形態に係る動画出力装置のブロック図である。動画出力装置は、動画作成装置(図11)と同様、画像入力装置11、楽曲入力装置12、入力装置13、シナリオ生成装置14、動画生成装置15を備えており、さらに、動画再生装置16および出力装置17を備えている。 40

【0064】

動画再生装置16は、動画生成装置15の生成した動画を、出力装置17用の再生信号に変換する装置であり、例えば、パソコン上で動作するネットワークメディアプレーヤーなどの動画プレーヤーや、MPEGデコーダなどが挙げられる。

【0065】

出力装置17は、動画再生装置16からの再生信号に基づいて動画を出力するとともに、楽曲音声を出力する装置であり、例えば、ディスプレイヤやスピーカーが挙げられる。

【図面の簡単な説明】

【0066】

【図1】シナリオ作成装置のブロック図

【図2】画像処理装置のブロック図

【図3】第1実施形態に係るシナリオ生成処理の流れを示すフローチャート

【図4】楽曲の構成要素と画像群との関連づけ（楽曲のフレーズと画像のイベントの関連づけ）の一例を示す図

【図5】楽曲の構成要素と画像群との関連づけ（楽曲のサビと類似像群の関連づけ）の一例を示す図

【図6】楽曲の構成要素と画像群との関連づけ（楽曲の各フレーズと各類似像群の関連づけ）の一例を示す図

【図7】楽曲の構成要素と画像群との関連づけ（楽曲のサビとサビ用画像群の関連づけ）の一例を示す図

【図8】拍または裏拍と画像との関連づけの一例を示す図

【図9】シナリオの一例を示す図

【図10】第2実施形態に係るシナリオ生成処理の流れを示すフローチャート

【図11】動画作成装置のブロック図

【図12】動画生成処理の流れを示すフローチャート

【図13】動画出力装置のブロック図

【符号の説明】

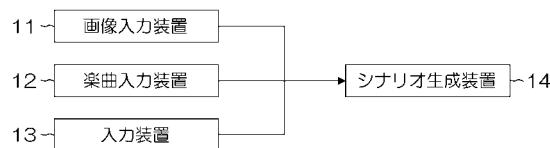
【0067】

11：画像入力装置、12：楽曲入力装置、13：入力装置、14：シナリオ生成装置  
、15：動画生成装置、16：動画再生装置、17：出力装置

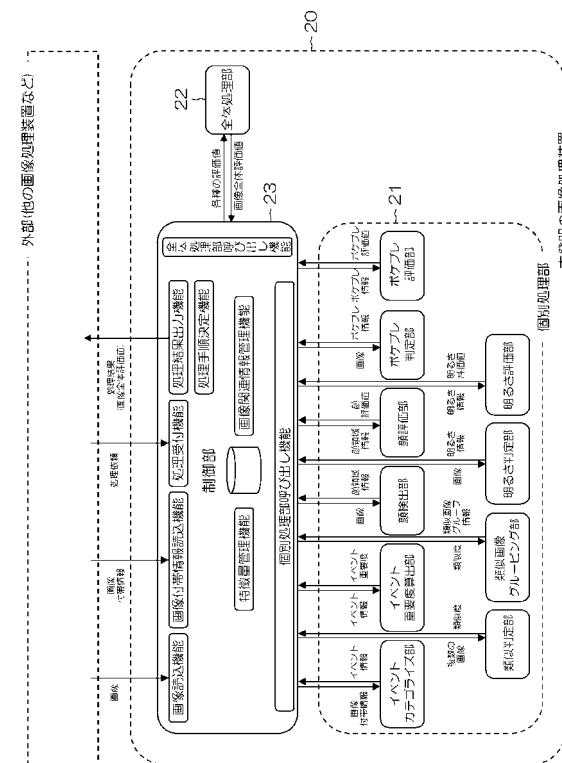
10

20

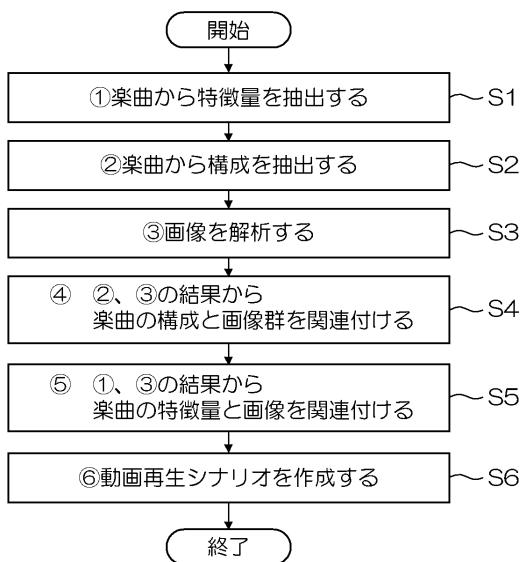
【図1】



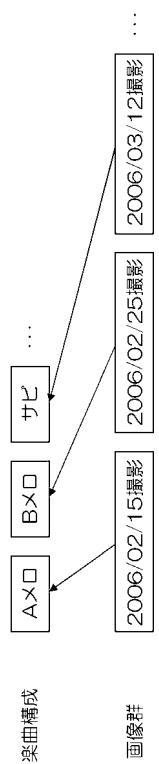
【図2】



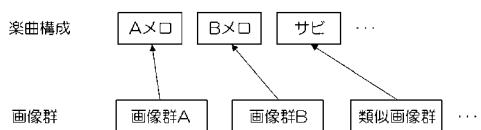
【図3】



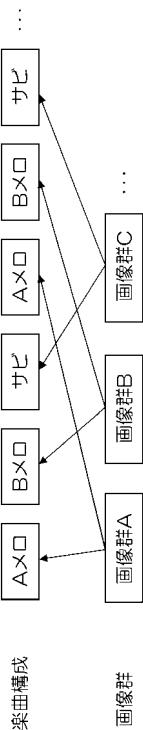
【図4】



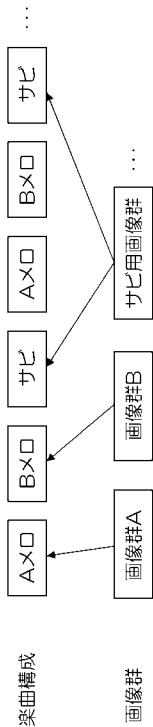
【図5】



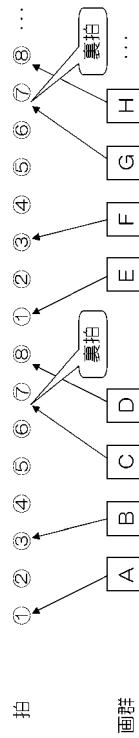
【図6】



【図7】



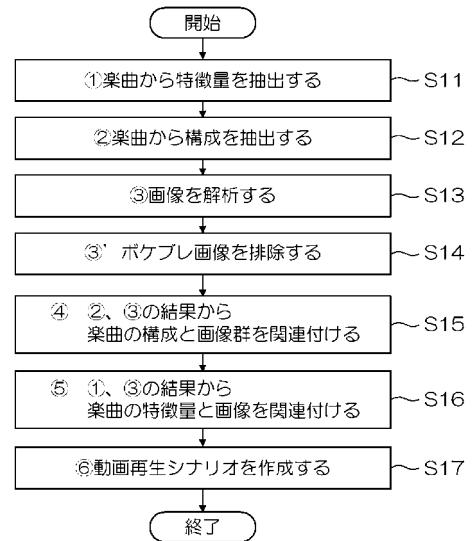
【図8】



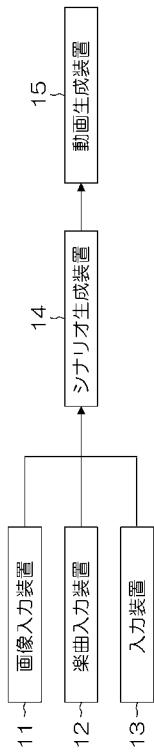
【図9】

```
<?xml version="1.0" encoding="UTF-B"?>
<scenario xmlns="http://xmlins.fujifilm.com/scenario">
  <image-select method="filename-sort" order="ascending" />
  <audio dur="3.00" arc="music.wav" />
  <image id="1" >
    <file src="1.jpg" />
    <time begin="0:00" end="1:00" />
  </image>
  <image id="2" >
    <file src="2.jpg" />
    <time begin="1:00" end="2:00" />
  </image>
  <image id="3" >
    <file src="3.jpg" />
    <time begin="2:00" end="3:00" />
  </image>
</scenario>
```

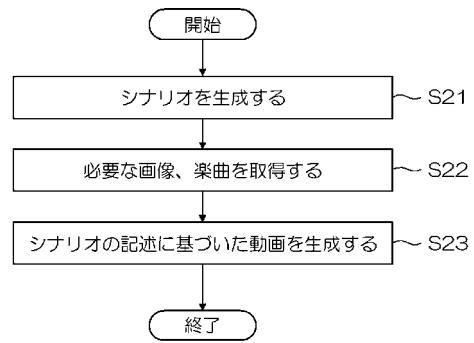
【図10】



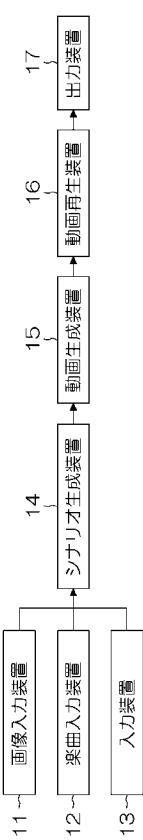
【図 1 1】



【図 1 2】



【図 1 3】



---

フロントページの続き

(51)Int.Cl.	F I
G 0 6 T 1/00 (2006.01)	G 0 6 T 7/00 3 0 0 F
G 0 6 T 7/00 (2006.01)	H 0 4 N 5/262
H 0 4 N 5/262 (2006.01)	H 0 4 N 5/93 Z
H 0 4 N 5/93 (2006.01)	

(56)参考文献 特開2005-033554 (JP, A)  
特開2005-318295 (JP, A)  
特開2005-242833 (JP, A)

## (58)調査した分野(Int.Cl., DB名)

H 0 4 N	5 / 7 6 - 5 / 9 5 6
G 0 6 T	1 / 0 0
G 0 6 T	7 / 0 0
G 1 0 L	1 1 / 0 0
H 0 4 N	5 / 2 2 2
H 0 4 N	5 / 2 6 2