

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5540376号
(P5540376)

(45) 発行日 平成26年7月2日(2014.7.2)

(24) 登録日 平成26年5月16日(2014.5.16)

(51) Int.Cl. F I
H04N 5/91 (2006.01) H04N 5/91 Z

請求項の数 6 (全 17 頁)

(21) 出願番号	特願2009-213682 (P2009-213682)	(73) 特許権者	597040902
(22) 出願日	平成21年9月15日 (2009.9.15)		学校法人東京工芸大学
(65) 公開番号	特開2011-66546 (P2011-66546A)		東京都中野区本町2丁目9番5号
(43) 公開日	平成23年3月31日 (2011.3.31)	(74) 代理人	100079049
審査請求日	平成24年8月9日 (2012.8.9)		弁理士 中島 淳
		(74) 代理人	100084995
			弁理士 加藤 和詳
		(74) 代理人	100085279
			弁理士 西元 勝一
		(74) 代理人	100099025
			弁理士 福田 浩志
		(72) 発明者	森山 剛
			神奈川県厚木市飯山1583 学校法人東京工芸大学内

最終頁に続く

(54) 【発明の名称】 コマ割り画像生成装置及びプログラム

(57) 【特許請求の範囲】

【請求項1】

入力された映像信号又は前記映像信号に対応する音声信号に基づいて、複数種類の特定の演出印象の各々について、前記特定の演出印象を有する映像部分を検出する演出検出手段と、

前記複数種類の特定の演出印象の各々について、前記映像信号の映像フレームから、前記演出検出手段によって検出された映像部分に対応する複数の映像フレームを選択するフレーム選択手段と、

前記複数種類の特定の演出印象の各々について、前記特定の演出印象に対応し、かつ、複数のコマに分割されたコマ割りの各コマに、前記フレーム選択手段によって前記特定の演出印象として選択された前記映像フレームを各々割り当てたコマ割り画像を各々生成する生成手段と、

を含むコマ割り画像生成装置であって、

前記コマ割りの各コマの境界の形状を、前記特定の演出印象の種類に応じた形状としたコマ割り画像生成装置。

【請求項2】

前記特定の演出印象は、映像中で短時間の間にカットが頻出している緊迫したシーンを含み、

前記緊迫したシーンに対応するコマ割りを、斜めの線を含む境界によって複数のコマに分割したものとし、

10

20

前記演出検出手段は、前記映像信号の連続した映像フレーム間の色情報の差分に基づいて、連続撮影の境界であるカットを検出し、カットの検出回数が閾値以上となる所定範囲を、前記緊迫したシーンとなる映像部分として検出する請求項1記載のコマ割り画像生成装置。

【請求項3】

前記特定の演出印象は、衝撃的なシーンを含み、

前記衝撃的なシーンに対応するコマ割りを、斜めの線を含む境界によって複数のコマに分割すると共に、他のコマよりも大きさが大きいコマを有するものとし、

前記演出検出手段は、前記映像信号に対応する音声信号に基づいて、分析窓内の映像信号に対応する音声信号の短時間平均パワーを算出し、算出された短時間平均パワーが閾値より多い前記分析窓の映像信号を、前記衝撃的なシーンとなる映像部分として検出する請求項1又は2記載のコマ割り画像生成装置。

10

【請求項4】

前記特定の演出印象は、会話シーン又はゆっくりしたシーンを含み、

前記会話シーン又はゆっくりしたシーンに対応するコマ割りを、規則的に配置され、かつ、各々の大きさが対応している複数のコマに分割したものとした請求項1～請求項3の何れか1項記載のコマ割り画像生成装置。

【請求項5】

前記演出検出手段は、前記映像信号に対応する音声信号に基づいて、前記会話シーン又はゆっくりしたシーンとなる映像部分を検出する請求項4記載のコマ割り画像生成装置。

20

【請求項6】

コンピュータを、

入力された映像信号又は前記映像信号に対応する音声信号に基づいて、複数種類の特定の演出印象の各々について、前記特定の演出印象を有する映像部分を検出する演出検出手段、

前記複数種類の特定の演出印象の各々について、前記映像信号の映像フレームから、前記演出検出手段によって検出された映像部分に対応する複数の映像フレームを選択するフレーム選択手段、及び

前記複数種類の特定の演出印象の各々について、前記特定の演出印象に対応し、かつ、複数のコマに分割されたコマ割りの各コマに、前記フレーム選択手段によって前記特定の演出印象として選択された前記映像フレームを各々割り当てたコマ割り画像を各々生成する生成手段

30

として機能させるためのプログラムであって、

前記コマ割りの各コマの境界の形状を、前記特定の演出印象の種類に応じた形状としたプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、コマ割り画像生成装置及びプログラムに係り、特に、映像の内容及び演出印象を表わすコマ割り画像を生成するコマ割り画像生成装置及びプログラムに関する。

40

【背景技術】

【0002】

従来、映像の要約に関しては多くの研究がある。それらの多くは、カット（カメラの切り替わる点）や大音量の箇所、あるいは画像中の特定の物体（例えば、主人公の顔）が出現する箇所を重要箇所（キーフレーム）として検出する手法である。すなわち、従来の手法は、映像中に映っている内容を基準に映像を要約しようとするものであった。

【0003】

例えば、テレビ番組等の映像信号及び音声信号から、人物領域を検出すると共に台詞を認識し、吹き出し及び効果線を付与して、自動的に漫画画像を生成する漫画生成装置が知られている（特許文献1）。

50

【 0 0 0 4 】

ところで、笑顔の人物が映った画像は、どのように映っているかによって、様々な解釈が可能である。例えば、女性が食事に誘う映像を撮影する場合を考えると、ハイアングル（人物の斜め上から撮影）では、女性が愛らしく見えるのに対して、ローアングル（人物の斜め下から撮影）では、女性が高圧的に見える。このように、映像中に映っている内容がどういう意味を持つかを解釈するためには、それがどのように映っているかという情報が不可欠である。

【 0 0 0 5 】

そして、被写体をどのように映すと、観客がどのように受け取るかは、映画の文法と呼ばれ、20世紀前半に集大成されて今日に至っている。従って、カメラワーク（レンズワーク、アングル、ショットサイズ、照明、被写体とカメラの動き）や音響（効果音、せりふの音声、音楽）といった物理的な構造から、そのシーンが観客にどのような心理的印象を与えるか（以下、演出印象）を予測することができると考えられる。

【 先行技術文献 】

【 特許文献 】

【 0 0 0 6 】

【 特許文献 1 】 特開 2 0 0 3 - 8 5 5 7 2 号 公 報

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 0 7 】

従来の映像要約法では、確かに人物の顔や物体が見て取れるとしても、どのように映っているかを表わす情報を保存していないため、映像全体の流れの中で、その人物の顔や物体がどのような意味を持つのかを解釈する上では、十分な情報を保存していない、という問題がある。

【 0 0 0 8 】

本発明は、上記の問題点を解決するためになされたもので、映像信号が表わす内容及び演出印象を表わしたコマ割り画像を生成することができるコマ割り画像生成装置及びプログラムを提供することを目的とする。

【 課題を解決するための手段 】

【 0 0 0 9 】

上記の目的を達成するために本発明に係るコマ割り画像生成装置は、入力された映像信号又は前記映像信号に対応する音声信号に基づいて、複数種類の特定の演出印象の各々について、前記特定の演出印象を有する映像部分を検出する演出検出手段と、前記複数種類の特定の演出印象の各々について、前記映像信号の映像フレームから、前記演出検出手段によって検出された映像部分に対応する複数の映像フレームを選択するフレーム選択手段と、前記複数種類の特定の演出印象の各々について、前記特定の演出印象に対応し、かつ、複数のコマに分割されたコマ割りの各コマに、前記フレーム選択手段によって前記特定の演出印象として選択された前記映像フレームを各々割り当てたコマ割り画像を各々生成する生成手段と、を含み、前記コマ割りの各コマの境界の形状を、前記特定の演出印象の種類に応じた形状としている。

【 0 0 1 0 】

本発明に係るプログラムは、コンピュータを、入力された映像信号又は前記映像信号に対応する音声信号に基づいて、複数種類の特定の演出印象の各々について、前記特定の演出印象を有する映像部分を検出する演出検出手段、前記複数種類の特定の演出印象の各々について、前記映像信号の映像フレームから、前記演出検出手段によって検出された映像部分に対応する複数の映像フレームを選択するフレーム選択手段、及び前記複数種類の特定の演出印象の各々について、前記特定の演出印象に対応し、かつ、複数のコマに分割されたコマ割りの各コマに、前記フレーム選択手段によって前記特定の演出印象として選択された前記映像フレームを各々割り当てたコマ割り画像を各々生成する生成手段として機能させるためのプログラムであって、前記コマ割りの各コマの境界の形状を、前記特定の

10

20

30

40

50

演出印象の種類に応じた形状としている。

【0011】

本発明によれば、演出検出手段によって、入力された映像信号又は映像信号に対応する音声信号に基づいて、特定の演出印象を有する映像部分を検出する。フレーム選択手段によって、映像信号の映像フレームから、演出検出手段によって検出された映像部分に対応する複数の映像フレームを選択する。

【0012】

そして、生成手段によって、特定の演出印象に対応し、かつ、複数のコマに分割されたコマ割りの各コマに、フレーム選択手段によって選択された映像フレームを各々割り当てたコマ割り画像を生成する。

10

【0013】

このように、特定の演出印象を有する映像部分に対応する複数の映像フレームを、特定の演出印象に対応するコマ割りの各コマに割り当てたコマ割り画像を生成することにより、映像信号が表わす内容と共に、それが観客にどのような心理的印象を与えるかを表わしたコマ割り画像を生成することができる。

【0014】

本発明に係る演出検出手段は、映像信号又は音声信号に基づいて、複数種類の特定の演出印象の各々について、特定の演出印象を有する映像部分を検出し、フレーム選択手段は、複数種類の特定の演出印象の各々について、映像信号の映像フレームから、特定の演出印象が検出された映像部分に対応する複数の映像フレームを選択し、生成手段は、複数種類の特定の演出印象の各々について、特定の演出印象に対応するコマ割りの各コマに、特定の演出印象として選択された映像フレームを各々割り当てたコマ割り画像を各々生成することができる。これによって、複数種類の演出印象について、映像信号が表わす内容及び演出印象を表わしたコマ割り画像を生成することができる。

20

【0015】

上記のコマ割りの各コマの境界の形状を、特定の演出印象の種類に応じた形状とすることができる。

【0016】

上記の特定の演出印象は、緊迫したシーンを含み、緊迫したシーンに対応するコマ割りを、斜めの線を含む境界によって複数のコマに分割したものとすることができる。これによって、緊迫したシーンであることを表わしたコマ割りとすることができる。

30

【0017】

また、演出検出手段は、映像信号の連続した映像フレーム間の色情報の差分に基づいて、連続撮影の境界であるカットを検出し、カットの検出回数が閾値以上となる所定範囲を、緊迫したシーンとなる映像部分として検出することができる。

【0018】

上記の特定の演出印象は、衝撃的なシーンを含み、衝撃的なシーンに対応するコマ割りを、斜めの線を含む境界によって複数のコマに分割すると共に、他のコマよりも大きさが大きいコマを有するものとすることができる。これによって、衝撃的なシーンであることを表わしたコマ割りとすることができる。

40

【0019】

また、演出検出手段は、映像信号に対応する音声信号に基づいて、衝撃的なシーンとなる映像部分を検出することができる。

【0020】

上記の特定の演出印象は、会話シーン又はゆっくりしたシーンを含み、会話シーン又はゆっくりしたシーンに対応するコマ割りを、規則的に配置され、かつ、各々の大きさが対応している複数のコマに分割したものとすることができる。これによって、会話シーン又はゆっくりしたシーンであることを表わしたコマ割りとすることができる。

【0021】

また、演出検出手段は、映像信号に対応する音声信号に基づいて、会話シーン又はゆっ

50

くりしたシーンとなる映像部分を検出することができる。

【発明の効果】

【0022】

以上説明したように、本発明のコマ割り画像生成装置及びプログラムによれば、特定の演出印象を有する映像部分に対応する複数の映像フレームを、特定の演出印象に対応するコマ割りの各コマに割り当てたコマ割り画像を生成することにより、映像信号が表わす内容と共に、それが観客にどのような心理的印象を与えるかを表わしたコマ割り画像を生成することができる、という効果が得られる。

【図面の簡単な説明】

【0023】

【図1】本発明の第1の実施の形態に係る映像機器の構成を示す概略図である。

【図2】RGBヒストグラムを示す図である。

【図3】緊張感のあるシーンに対応するコマ割りの例を示すイメージ図である。

【図4】衝撃的なシーンに対応するコマ割りの例を示すイメージ図である。

【図5】会話シーンに対応するコマ割りの例を示すイメージ図である。

【図6】(A)フレームの例を示すイメージ図、及び(B)コマの形状をフレームに重畳させた様子を示すイメージ図である。

【図7】本発明の第1の実施の形態に係る映像機器におけるコマ割り画像生成処理ルーチンの内容を示すフローチャートである。

【図8】本発明の第1の実施の形態に係る映像機器における緊迫シーンコマ画像生成処理ルーチンの内容を示すフローチャートである。

【図9】本発明の第1の実施の形態に係る映像機器におけるカット頻度算出処理ルーチンの内容を示すフローチャートである。

【図10】本発明の第1の実施の形態に係る映像機器における衝撃シーンコマ画像生成処理ルーチンの内容を示すフローチャートである。

【図11】本発明の第1の実施の形態に係る映像機器における会話シーンコマ画像生成処理ルーチンの内容を示すフローチャートである。

【図12】普通のシーンに対応するコマ割りの例を示すイメージ図である。

【発明を実施するための形態】

【0024】

以下、図面を参照して本発明の実施の形態を詳細に説明する。なお、映像信号を記録及び再生する映像機器に本発明を適用した場合を例に説明する。

【0025】

図1に示すように、第1の実施の形態に係る映像機器10は、CPUと、RAMと、後述するコマ割り画像生成処理ルーチンを実行するためのプログラムを記憶したROMとを備え、機能的には次に示すように構成されている。

【0026】

映像機器10は、DVDなどの記録媒体を介して入力された映像信号及び音声信号からなる映画データを記憶した映画記憶部12と、映画記憶部12から処理対象としての映画データを取得する映画データ取得部14と、特定の演出印象を有する映像部分を検出する演出検出部16と、特定の演出印象の種類毎に予め用意されたコマ割りデータを記憶するコマ割り記憶部18と、検出された特定の演出印象に対応するコマ割りデータをコマ割り記憶部18から取得するコマ割り取得部20と、検出された特定の演出印象に応じて、映画データの映像信号から、キーフレームを複数選択するキーフレーム選択部22と、選択されたキーフレームを、取得したコマ割りデータの各コマに割り当てて、コマ割り画像を生成し、映画データの要約として映画記憶部12に記憶させるコマ割り画像生成部24と、映画記憶部12に記憶された映画データ及びコマ割り画像を、ディスプレイなどの表示装置40に表示させる表示制御部26とを備えている。なお、フレームは、映像における伝送単位であり、1枚の画像である。また、シーンは、映像または漫画を視聴した人物が、意味的なまとまり、もしくは演出を受け取る部分である。コマは、ページを折れ線分に

10

20

30

40

50

よって分割した単位であり、ページは、紙媒体の漫画における、1枚の紙の片面である。演出は、映像または漫画を視聴した人物が、映像または漫画から受け取る印象である。

【0027】

次に、複数種類の特定の演出印象を検出する原理について説明する。まず、特定の演出印象として、緊迫感のあるシーン（あるいは、動きのあるシーン）を検出する場合について説明する。

【0028】

映画においては、緊迫感を演出するための様々な手法がある。例えば、効果音として持続的な高音を鳴らしたり、カメラもしくは被写体を大きく動かしたり、一定時間内にショットを頻繁に切り替えたりする手法である。本実施の形態では、最後に挙げた手法に着目する。この手法は、視聴者の視界を激しく変化させ、その心理状態をかく乱することを狙ったものである。これは言い換えれば、映像中における緊迫箇所とは、映像中で短時間の間にカットが頻出している箇所だと言える。従って、映像中でカットの頻度が突出している箇所を検出することによって、緊迫箇所を検出することができると考えられる。ここで、カットとは、ショットの境界であり、ショットとは、映像において、1台のカメラで連続して撮像された部分である。

【0029】

本実施の形態では、演出検出部16によって、入力された映画データの映像信号の最初から、T秒間の分析窓を1秒ずつシフトしながら、逐次、分析窓内でカットの頻度を求める。カットは、映像中のフレーム間の色分布の差分を閾値処理して求める。分析窓内におけるカット頻度が、あらかじめ定めた閾値よりも多い場合、分析窓の位置を、緊迫感のあるシーンとなる箇所として検出する。なお、カット頻度に関する閾値は、予め実験などに求めておけばよく、例えば、閾値を7（回/10秒間）とする。

【0030】

なお、ショット間のつなぎ方には、通常の切り替え以外にも、ディゾルブ（隣り合うショットの映像フレームを混ぜ合わせながらスムーズに切り替える効果）等の効果がある。本実施の形態では、緊迫感に最も寄与するものとして、通常の切り替えのみを対象とする。ショットの境界前後では、多くの場合、異なる舞台装置が撮影されており、その色分布に急激な変化が生ずる。そこで、演出検出部16では、現フレームと次フレームのそれぞれに対して、色分布として、図2に示すようなRGBヒストグラムを算出し、色分布の差分として、色チャンネルごとにヒストグラム間のバタチャリヤ（Bhattacharyya）距離を求める。ここで、バタチャリヤ距離とは、独立な二つの事象の同時確率に対する自己情報量として定義され、アフィン変換（平行移動、回転、伸縮）に不変であるため、カメラの移動やズーム、回転に関して影響を受けない。RGBチャンネル間の距離の平均値が、あらかじめ設定した閾値以上となる場合、そのフレーム間でカットが発生したと判断される。

【0031】

なお、バタチャリヤ距離に対する閾値は、実験などにより予め定めておけばよい。例えば、カット頻度の算出範囲（分析窓）を $T = 10$ 秒とし、これを $T/2 = 5$ 秒ずつシフトしながらカット頻度を算出する場合、バタチャリヤ距離に対する閾値を、0.1から0.4まで0.02刻みで変化させると、ROC（Receiver Operating Characteristic）カーブの傾きが最大となった値0.36が得られ、これを閾値として用いる。

【0032】

また、カット頻度に対する閾値は、実験などにより例えば7回と予め定めておけばよい。この場合、これを越えた分析窓位置を、統計上稀にみる緊迫シーンとなる箇所として検出し、また、目視で検出した真値としてのカット（A）、プログラムによって検出されたカット（B）として、プログラムによって検出されたカットのうちの正解カット（C）について以下の（1）式、（2）式により再現率及び適合率を算出した。その結果、再現率は53.8%、適合率は81.4%であった。

10

20

30

40

50

再現率 = $|C| / |A|$ (1)

適合率 = $|C| / |B|$ (2)

【0033】

次に、特定の演出印象として、衝撃的なシーンを検出する場合について説明する。

【0034】

演出検出部16は、入力された映画データの映像信号の最初から、T秒間（例えば1秒間）の分析窓をT/2秒ずつシフトしながら、逐次、分析窓内で、映像信号に対応する音声信号の短時間平均パワーを算出する。衝撃的なシーンの箇所は、映像中の短時間平均パワーを閾値処理して求める。分析窓内における短時間平均パワーが、予め定めた閾値よりも多い場合、分析窓の位置を衝撃的なシーンとして検出する。なお、短時間平均パワーに関する閾値は、実験などにより予め求めておけばよい。

10

【0035】

次に、特定の演出印象として、会話シーン（ゆっくりしたシーン）を検出する場合について説明する。

【0036】

演出検出部16は、入力された映画データの映像信号の最初から、T秒間（例えば62.5ms）の分析窓をT/2秒ずつシフトしながら、逐次、分析窓内で、音声検出処理を行い、音声を検出されたか否かを判断する。音声を検出された場合、当該分析窓を、音声区間としてマークする。演出検出部16は、過去M秒（例えば10秒）の音声区間の割合を算出し、音声区間の割合が、あらかじめ定めた閾値（例えば0.4）以上である場合、分析窓の位置を、会話シーンとして検出する。

20

【0037】

次に、本実施の形態の原理について説明する。

【0038】

漫画においても、映像と同様の演出技法が開発されている。例えば、非特許文献1（秋田孝宏、「「コマ」から「フィルム」へ」、マンガとマンガ映画、NTT出版、2005）や非特許文献2（竹内オサム、「マンガ表現学入門」、筑摩書房、2006）の技術が知られている。本を媒体とする漫画においては、シーンをページに分割し、ページをコマに分割する際に、ページ数やコマ数、さらにコマのサイズの大小、コマの形を変えることにより、観客がコマに描かれた絵をどのように受け取るかをコントロールする（本実施の形態ではこれらを総称して、コマ割りと呼ぶこととする）。従って、同じ被写体もしくは描画対象を、映画と漫画の双方において、同じ演出で表現することが可能であると考えられる。

30

【0039】

本実施の形態では、映画中から、特定の演出がなされている箇所を検出し、その近傍のキーフレームを、同様の演出印象を有する漫画のコマ割りに配置することにより、その箇所に映っている被写体と演出印象とを同時に要約する手法を提案する。

【0040】

次に、特定の演出印象の種類毎に予め用意されたコマ割りデータについて説明する。

【0041】

図3に、緊迫感のあるシーンに対応するコマ割りの例を示す。緊迫感のあるシーンに対応するコマ割りにおいて、コマの境界の形状に注目すると、斜めの線が多用されている。斜めの線は、映画の文法においても、観客に動きを感じさせる性質を有しており、局面が動的であることを演出する上で、映画と漫画の間で共通した規則であると考えられる。また、特定のコマに観客を注視させずコマからコマへ視線を動かすため、核となるコマが際立つようなコマの大きさの差異を持たせない。コマ割り記憶部18には、緊迫感のあるシーンに対応して、上記図3に示すようなコマ割りを表わすコマ割りデータが1種類だけ予め記憶されている。

40

【0042】

図4に、衝撃的なシーンに対応するコマ割りの例を示す。核となるコマとして特定されるほど他のコマより大きさが大きいコマがあり、当該大きなコマによって、衝撃が与えら

50

れる。当該大きなコマは、ほぼ斜めではない境界に囲まれており、一方、斜めの線である境界で囲まれた小さいいくつかのコマを有している。

【 0 0 4 3 】

例えば、衝撃的なシーンに対応するコマ割りでは、右の大きなコマに、ここで起こった事件を描き、左に縦3つ並んだコマそれぞれに、その事件を取り巻く人物3人の驚いた顔を描く。これにより、起こった事件が大変に衝撃的で、その驚きのあまり、人物全員の時間が止まったかのように息を詰まらせている衝撃的なシーンを演出することができる。コマ割り記憶部18には、衝撃的なシーンに対応して、上記図4に示すようなコマ割りを表わすコマ割りデータが1種類だけ予め記憶されている。

【 0 0 4 4 】

図5に、会話シーン(ゆっくりしたシーン)に対応するコマ割りの例を示す。中くらいの大きさで、ほぼ同じ大きさの複数のコマが規則的に配置されている。コマの境界のずれが、コマ画像の順序通りに視線を誘導する。コマ割り記憶部18には、会話シーンに対応して、上記図5に示すようなコマ割りを表わすコマ割りデータが1種類だけ予め記憶されている。

【 0 0 4 5 】

キーフレーム選択部22は、緊迫感のあるシーンが検出された分析窓内から、コマ割りのコマ数に等しい6ショット分の先頭フレームを、緊迫感のあるシーンに対応するキーフレームとして選択する。

【 0 0 4 6 】

キーフレーム選択部22は、衝撃的なシーンが検出された分析窓の位置を含むショットの先頭フレームとその後の3ショット分の先頭フレーム(合わせると、コマ割りのコマ数に等しい4ショット分の先頭フレーム)を、衝撃的なシーンに対応するキーフレームとして選択する。

【 0 0 4 7 】

キーフレーム選択部22は、会話シーンが検出された範囲の先頭のショットから、コマ割りのコマ数に等しい8ショット分の先頭フレームを、会話シーンに対応するキーフレームとして選択する。

【 0 0 4 8 】

コマ割り画像生成部24は、選択された複数のキーフレームを、当該特定の演出印象に対応するコマ割りの各コマに割り当てて、コマ割り画像を生成し、映画データの要約として映画記憶部12に記憶させる。

【 0 0 4 9 】

コマへの割り当てでは、まず、図6(B)に示すように、コマの形状を、映像のフレームへ内接する比率で拡大縮小し、重心を一致させるようにフレームに重畳する。その状態で、図6(A)に示すようなフレームにおいて、コマと重なり合ったフレーム内の領域を、そのままコマへコピーして、割り当てる。なお、コマ割りデータの複数のコマは、予め順番が定められており、選択された複数のキーフレームのフレーム番号の順と、コマの順とを合わせて、キーフレームを各コマに割り当てる。例えば、コマの位置が右上に近いほど順番が早く、右下に近いほど順番が遅くなるように、コマ割りデータのコマの順番が定められている。

【 0 0 5 0 】

表示制御部26は、映像機器10の操作部(図示省略)又は遠隔操作部(図示省略)によって、映画データの閲覧が指示されたときに、映画記憶部12から、対応する映画データ及びコマ割り画像データを読み出し、映画データの要約としてのコマ割り画像データ、又は映画データを、表示装置40に表示させる。

【 0 0 5 1 】

次に、本実施の形態に係る映像機器10の作用について説明する。まず、記録媒体を介して、映画データが映像機器10に入力されて、入力された映画データが、映画記憶部12に記憶される。

10

20

30

40

50

【 0 0 5 2 】

そして、映像機器 1 0 において、図 7 に示すコマ割り画像生成処理ルーチンが実行される。

【 0 0 5 3 】

まず、ステップ 1 0 0 において、コマ割り画像の生成対象となる映画データを、映画記憶部 1 2 から取得する。そして、ステップ 1 0 2 において、緊迫感のあるシーンに対応するコマ割り画像を生成して、映画記憶部 1 2 に記憶し、ステップ 1 0 4 において、衝撃的なシーンに対応するコマ割り画像を生成して、映画記憶部 1 2 に記憶する。

【 0 0 5 4 】

次のステップ 1 0 6 では、会話シーンに対応するコマ割り画像を生成して、映画記憶部 1 2 に記憶し、コマ割り画像生成処理ルーチンを終了する。

10

【 0 0 5 5 】

上記ステップ 1 0 2 は、図 8 に示す緊迫シーンコマ画像生成処理ルーチンによって実現される。

【 0 0 5 6 】

ステップ 1 1 0 において、時刻 t を 0 に設定し、ステップ 1 1 2 で、時刻 t から T 秒間の範囲である分析窓内で、カット頻度を算出する。

【 0 0 5 7 】

そして、ステップ 1 1 4 において、上記ステップ 1 1 2 で算出されたカット頻度が閾値以上であるか否かを判定し、カット頻度が閾値未満である場合には、ステップ 1 1 6 へ進むが、一方、カット頻度が閾値以上である場合には、緊迫シーンが検出されたと判断し、ステップ 1 2 0 へ進む。

20

【 0 0 5 8 】

ステップ 1 1 6 では、分析窓が、映像終了位置に到達したか否かを判定し、映像終了位置に到達した場合には、緊迫シーンが検出されなかったと判断し、緊迫シーンに対応するコマ割り画像を生成せずに、緊迫シーンコマ画像生成処理ルーチンを終了する。

【 0 0 5 9 】

ステップ 1 1 6 で、分析窓が、映像終了位置に到達していないと判定された場合には、ステップ 1 1 8 で、時刻 t を 1 秒だけ進めて分析窓をシフトさせて、ステップ 1 1 2 へ戻る。

30

【 0 0 6 0 】

ステップ 1 2 0 では、上記ステップ 1 1 4 でカット頻度が閾値以上であると判定された分析窓内の各ショットの先頭フレームのフレーム番号を抽出する。次のステップ 1 2 2 では、コマ割り記憶部 1 8 から、緊迫感のあるシーンに対応するコマ割りデータを取得する。

【 0 0 6 1 】

そして、ステップ 1 2 4 において、上記ステップ 1 2 0 で抽出された先頭フレームのフレーム番号を、先頭から順に、上記ステップ 1 2 2 で取得したコマ割りデータのコマ数と等しいフレーム数分だけ選択して、映像信号から、複数のキーフレームを選択する。

【 0 0 6 2 】

そして、ステップ 1 2 6 において、上記ステップ 1 2 4 で選択したキーフレームの各々を、コマ割りデータの各コマに順に割り当てて、コマ割り画像を生成し、上記ステップ 1 0 0 で取得した映画データの要約として、映画記憶部 1 2 に記憶させて、上記ステップ 1 1 6 へ戻る。

40

【 0 0 6 3 】

上記ステップ 1 1 2 は、図 9 に示すカット頻度算出処理ルーチンによって実現される。

【 0 0 6 4 】

まず、ステップ 1 3 0 で、分析窓内の先頭フレームを対象フレームに設定すると共に、カット頻度を初期値 0 に設定する。

【 0 0 6 5 】

50

そして、ステップ132において、対象フレームのRGBヒストグラムを算出し、ステップ134で、対象フレームの次のフレームのRGBヒストグラムを算出する。

【0066】

そして、ステップ136では、上記ステップ132、134で算出されたRGBヒストグラム間のバタチャリヤ距離（RGBチャンネル間のバタチャリヤ距離の平均値）を算出する。次にステップ138において、上記ステップ136で算出されたバタチャリヤ距離が閾値以上であるか否かを判定し、算出されたバタチャリヤ距離が閾値未満である場合には、対象フレームと次フレームとの間にカットは存在しないと判断し、ステップ142へ進む。一方、算出されたバタチャリヤ距離が閾値以上である場合には、対象フレームと次フレームとの間にカットが存在すると判断し、ステップ140で、カット頻度をインクリメントしてカウントし、ステップ142へ進む。

10

【0067】

ステップ142では、分析窓内の全てのフレームについて上記の処理が終了したか否かを判定し、次フレームが最終フレームに到達した場合には、全てのフレームについて上記の処理が終了したと判断し、現在のカット頻度を算出結果として、カット頻度算出処理ルーチンを終了する。一方、次フレームが最終フレームに到達していない場合には、ステップ144で、次フレームを対象フレームに設定して、上記ステップ132へ戻る。

【0068】

上記ステップ104は、図10に示す衝撃シーンコマ画像生成処理ルーチンによって実現される。

20

【0069】

ステップ150において、時刻 t を0に設定し、ステップ152で、時刻 t から T 秒間の範囲である分析窓内で、短時間平均パワーを算出する。例えば、 T 秒間の音声波形サンプル値の二乗平均を計算することにより、短時間平均パワーを算出する。

【0070】

そして、ステップ154において、上記ステップ152で算出された短時間平均パワーが閾値以上であるか否かを判定し、短時間平均パワーが閾値未満である場合には、ステップ156へ進むが、一方、短時間平均パワーが閾値以上である場合には、衝撃的なシーンが検出されたと判断し、ステップ160へ進む。

【0071】

30

ステップ156では、分析窓が、映像終了位置に到達したか否かを判定し、映像終了位置に到達した場合には、衝撃的なシーンが検出されなかったと判断し、衝撃的なシーンに対応するコマ割り画像を生成せずに、衝撃シーンコマ画像生成処理ルーチンを終了する。

【0072】

ステップ156で、分析窓が、映像終了位置に到達していないと判定された場合には、ステップ158で、時刻 t を $T/2$ 秒だけ進めて分析窓を $T/2$ だけシフトさせて、ステップ152へ戻る。

【0073】

ステップ160では、上記ステップ154で短時間平均パワーが閾値以上であると判定された分析窓の位置を含むショットの先頭フレームのフレーム番号を抽出する。次のステップ162では、コマ割り記憶部18から、衝撃的なシーンに対応するコマ割りデータを取得する。

40

【0074】

そして、ステップ164において、上記ステップ160で抽出された先頭フレームのフレーム番号のキーフレームを映像信号から選択すると共に、後に続く複数のショットの各々の先頭フレームを、キーフレームとして映像信号から選択して、上記ステップ122で取得したコマ割りデータのコマ数と等しいフレーム数分だけのキーフレームを選択する。

【0075】

そして、ステップ166において、上記ステップ164で選択したキーフレームの各々を、コマ割りデータの各コマに順に割り当てて、コマ割り画像を生成し、上記ステップ1

50

00で取得した映画データの要約として映画記憶部12に記憶させて、上記ステップ156へ戻る。

【0076】

上記ステップ106は、図11に示す会話シーンコマ画像生成処理ルーチンによって実現される。

【0077】

ステップ170において、時刻 t を0に設定し、ステップ172で、上記ステップ100で取得した映画データの音声信号に対して、時刻 t から T 秒間の範囲である分析窓内で、音声検出処理を行う。なお、音声検出処理では、従来既知の手法を用いればよいので、音声検出処理に関する説明を省略する。

10

【0078】

そして、ステップ174において、上記ステップ172での音声検出処理によって、音声が発見されたか否かを判定し、音声が発見されなかった場合には、ステップ176へ進むが、一方、音声が発見された場合には、ステップ180へ進む。

【0079】

ステップ176では、分析窓が、映像終了位置に到達したか否かを判定し、映像終了位置に到達した場合には、会話シーンが発見されなかったと判断し、会話シーンに対応するコマ割り画像を生成せずに、会話シーンコマ画像生成処理ルーチンを終了する。

【0080】

ステップ176で、分析窓が、映像終了位置に到達していないと判定された場合には、ステップ178で、時刻 t を $T/2$ 秒だけ進めて分析窓を $T/2$ だけシフトさせて、ステップ172へ戻る。

20

【0081】

ステップ180では、上記ステップ174で音声が発見されたと判定された分析窓の区間を、音声区間としてマークする。次のステップ182では、現在の分析窓の区間を含む過去 M 秒の範囲において、音声区間としてマークされた区間の割合を算出する。

【0082】

そして、ステップ184で、上記ステップ182で算出された割合が閾値以上であるか否かを判定し、割合が閾値未満である場合には、ステップ178へ移行するが、一方、割合が閾値以上である場合には、会話シーンが発見されたと判断し、ステップ184へ進む。

30

【0083】

ステップ186では、コマ割り記憶部18から、会話シーンに対応するコマ割りデータを取得する。

【0084】

そして、ステップ188において、上記ステップ184で音声区間の割合が閾値以上であると判定された過去 M 秒の範囲の映像信号の先頭のショットから、先頭フレームをキーフレームとして選択すると共に、後に続く複数のショットの各々の先頭フレームを、キーフレームとして映像信号から選択して、上記ステップ186で取得したコマ割りデータのコマ数と等しいフレーム数分だけのキーフレームを選択する。

40

【0085】

そして、ステップ190において、上記ステップ188で選択したキーフレームの各々を、コマ割りデータの各コマに順に割り当てて、コマ割り画像を生成し、上記ステップ100で取得した映画データの要約として映画記憶部12に記憶させて、上記ステップ176へ戻る。

【0086】

また、ユーザが、映像機器10の操作部(図示省略)又は遠隔操作部(図示省略)を操作して、映画データの閲覧を指示すると、映像機器10の表示制御部26によって、映画記憶部12から、指示された映画データと、当該映画データに対応し、かつ、特定の演出印象の各種類に応じたコマ割り画像を読み出し、映画データの要約として、特定の演出印

50

象の各種類に応じたコマ割り画像を表示装置40に表示させると共に、映画データを、表示装置40に表示させる。

【0087】

以上説明したように、第1の実施の形態に係る映像機器によれば、特定の演出印象として検出された映像部分に対応する複数のキーフレームを、特定の演出印象を表現する予め用意された漫画のコマ割りの各コマに割り当てたコマ割り画像を生成することにより、映像信号が表わす内容と共に、それが観客にどのような心理的印象を与えるかを表わしたコマ割り画像を、映像の要約として生成することができる。また、複数種類の演出印象について、映像信号が表わす内容と共に、それが観客にどのような心理的印象を与えるかを表わしたコマ割り画像を生成することができる。

10

【0088】

映像の要約に関しては、従来、映像に何が映ったかを基準に行われてきた。しかし、映像中の被写体が、映像全体の中でどういった意味をもつかという解釈は、その被写体がどのように映像に映ったかを示す情報があって初めて可能となる。そこで、本実施の形態では、映像中において、特定の演出印象を有する箇所を検出し、その箇所の近傍のキーフレームを、演出印象に応じた漫画のコマ割りに割り当てることで、映像の被写体と演出印象の双方を要約する手法を提案した。

【0089】

次に、第2の実施の形態について説明する。なお、第2の実施の形態に係る映像機器の構成は、第1の実施の形態と同様であるため、同一符号を付して説明を省略する。

20

【0090】

第2の実施の形態では、映像の物理的特徴に基づいて、コマ割りの各コマの形状及び大きさを決定している点と、フレームに対する物体認識結果に基づいて、フレームからコマへ切り出す領域を決定している点とが、第1の実施の形態と異なっている。

【0091】

第2の実施の形態では、コマ割り取得部20によって、検出された特定の演出印象に対応するコマ割りデータをコマ割り記憶部18から取得し、取得したコマ割りデータと、特定の演出印象が検出された映像部分の物理的特徴とに基づいて、コマ割りデータの各コマの形状及び大きさを決定し、検出された特定の演出に対応するコマ割りデータを生成する。

30

【0092】

コマ割り画像生成部24は、コマへの割り当てにおいて、映像のフレーム内の物体（例えば、人物の顔）の位置を物体認識等により検出し、映像の文法と、検出された物体の位置とに基づいて、フレームからコマへ切り出す領域を決定し、フレーム内の決定された領域を、そのままコマへコピーして、割り当てる。

【0093】

なお、第2の実施の形態に係る映像機器の他の構成及び作用については、第1の実施の形態と同様であるため、説明を省略する。

【0094】

上記の第1の実施の形態及び第2の実施の形態では、緊迫感のあるシーン、衝撃的なシーン、会話シーンに対応するコマ割り画像を生成する場合を例に説明したが、これに限定されるものではなく、更に、普通のシーンに対応するコマ割り画像を生成してもよい。例えば、緊迫感のあるシーン、衝撃的なシーン、及び会話シーンの何れでもない部分を、普通のシーンに対応する映像部分として検出して、キーフレームを選択し、普通のシーンに対応するコマ割りの各コマに割り当てて、普通のシーンに対応するコマ割り画像を生成してもよい。例えば、図12に示すような普通のシーンに対応するコマ割りデータを予め用意しておけばよい。普通のシーンに対応するコマ割りの各コマの形状及び大きさは、他の演出印象のコマ割りに見られる特徴を持たないように定めればよい。

40

【0095】

また、映画データなどの映像データから、コマ割り画像を生成する場合を例に説明した

50

が、ビデオ撮影等によって得られた映像データから、コマ割り画像を生成してもよい。ビデオ撮影によって得られた映像から特徴的な画像を抽出してインパクトのある要約ができ、また、印象深いアルバムを生成することができる。また、デジタルフォト画像などの静止画像からコマ割り画像を生成することに、本発明を応用してもよい。例えば、特定の演出印象の種類をユーザが選択し、入力されたデジタルフォト画像を、選択された特定の演出印象に対応するコマ割りの各コマに配置して、コマ割り画像を生成してもよい。コマ割りを用いて写真のアルバムを生成することができるようになり、演出効果を持つアルバムを自動生成することも可能となる。

【 0 0 9 6 】

また、記録媒体を介して映画データが入力される場合を例に説明したが、これに限定されるものではなく、放送されている映像データを入力として記憶するようにしてもよい。

10

【 0 0 9 7 】

また、映像フレーム間の色分布の差分として、RGBヒストグラム間のバタチャリヤ距離を算出する場合を例に説明したが、これに限定されるものではなく、例えば、フレーム間の他の画像特徴量の差分を算出してもよい。

【 0 0 9 8 】

また、本願明細書中において、プログラムが予めインストールされている実施形態として説明したが、当該プログラムをCDROM等の記憶媒体に格納して提供することも可能である。

【 符号の説明 】

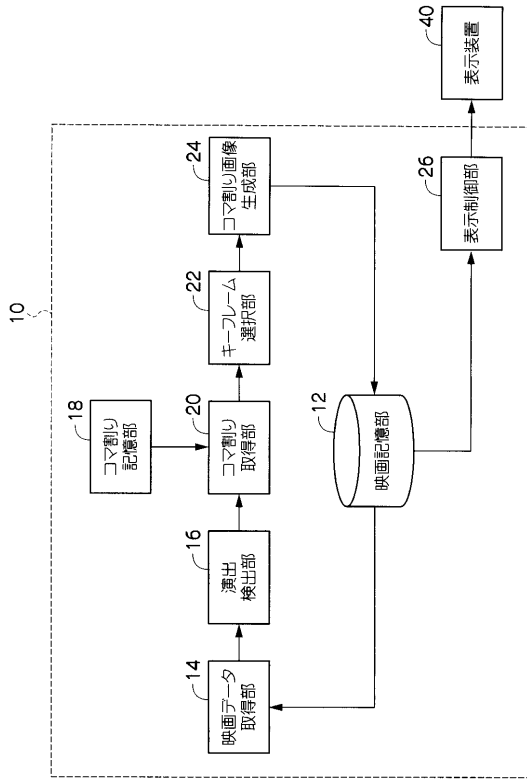
20

【 0 0 9 9 】

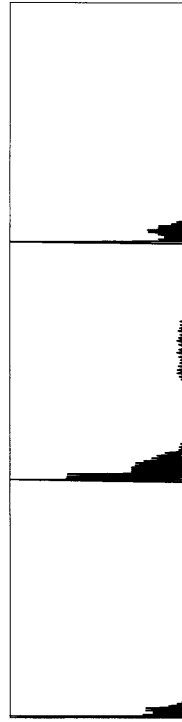
- 1 0 映像機器
- 1 2 映画記憶部
- 1 4 映画データ取得部
- 1 6 演出検出部
- 1 8 コマ割り記憶部
- 2 0 コマ割り取得部
- 2 2 キーフレーム選択部
- 2 4 コマ割り画像生成部
- 2 6 表示制御部
- 4 0 表示装置

30

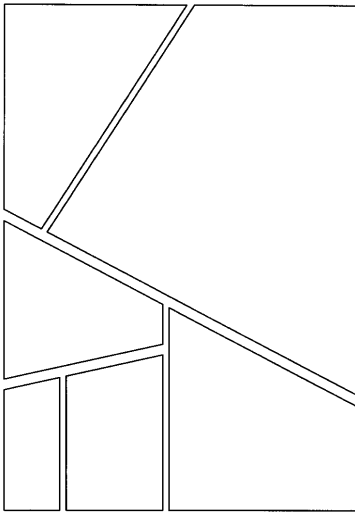
【 図 1 】



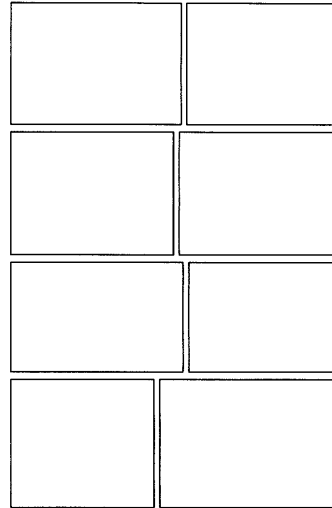
【 図 2 】



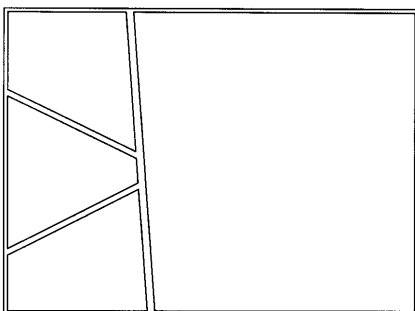
【 図 3 】



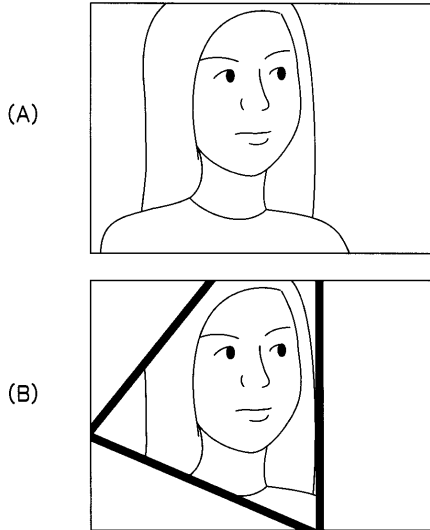
【 図 5 】



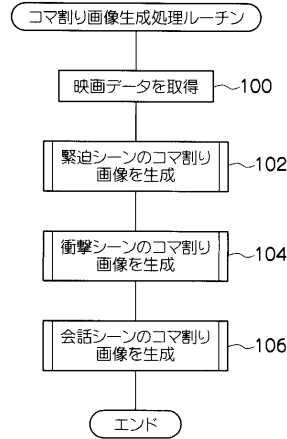
【 図 4 】



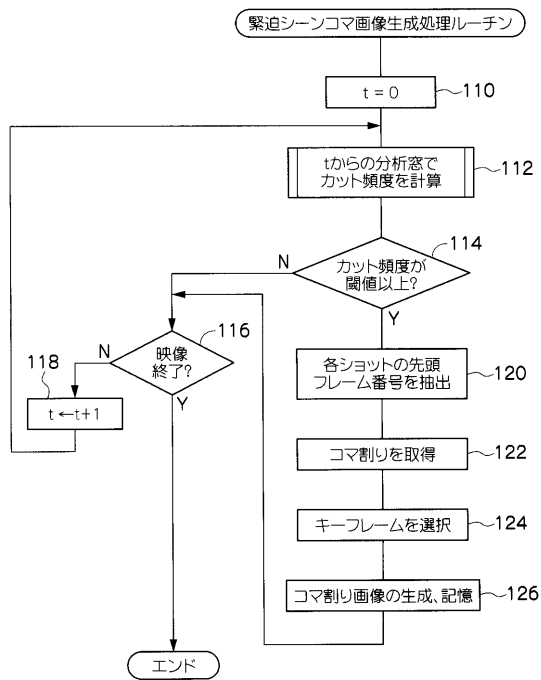
【図6】



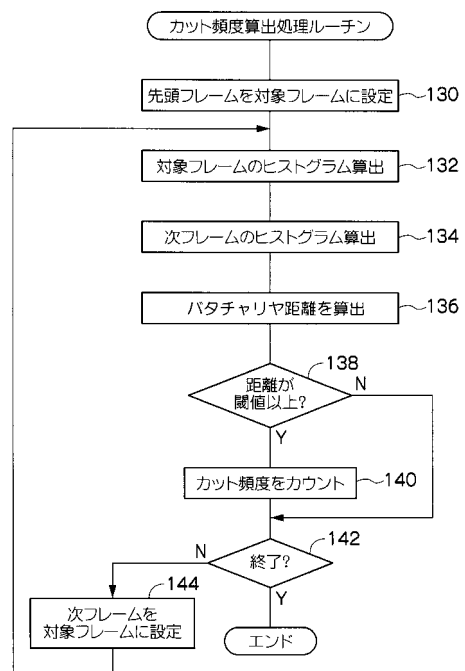
【図7】



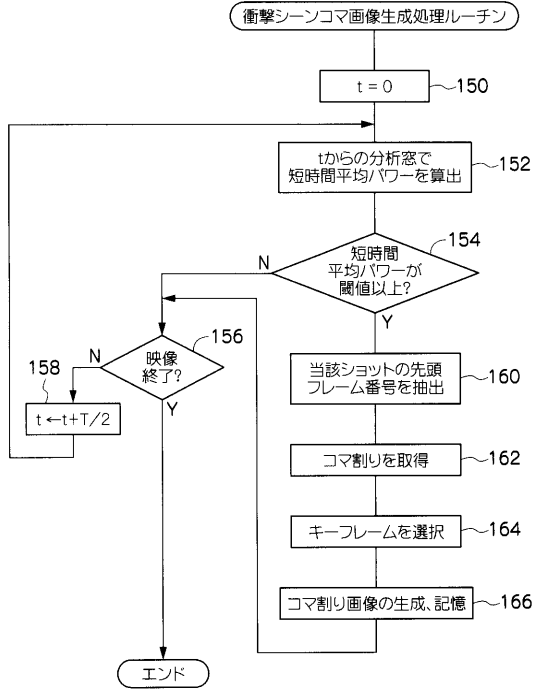
【図8】



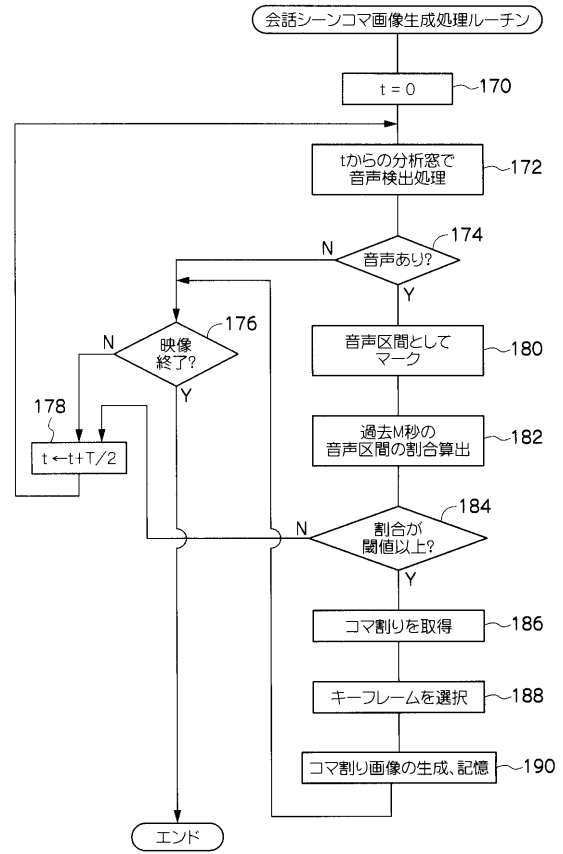
【図9】



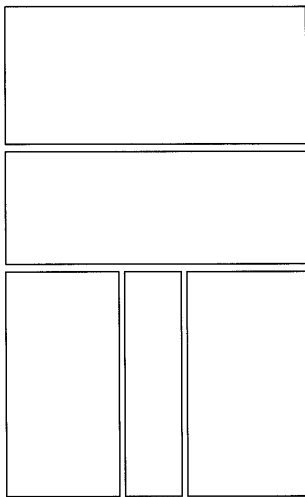
【図10】



【図11】



【図12】



フロントページの続き

審査官 後藤 嘉宏

- (56)参考文献 特開2003-085572(JP,A)
特開2009-065462(JP,A)
特開2006-014084(JP,A)
特開平03-080782(JP,A)
特開2006-050045(JP,A)
特開2005-167456(JP,A)
特開平09-265547(JP,A)
特開2008-092335(JP,A)

(58)調査した分野(Int.Cl., DB名)

H04N 5/76 - 5/956
H04N 1/38 - 1/393