

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5857122号
(P5857122)

(45) 発行日 平成28年2月10日 (2016. 2. 10)

(24) 登録日 平成27年12月18日 (2015. 12. 18)

(51) Int. Cl. F I
HO4N 5/93 (2006.01) HO4N 5/93 Z
G11B 27/34 (2006.01) G11B 27/34 N

請求項の数 13 (全 22 頁)

(21) 出願番号	特願2014-511511 (P2014-511511)	(73) 特許権者	513077243
(86) (22) 出願日	平成24年5月17日 (2012. 5. 17)		インテレクトチュアル ベンチャーズ ファ
(65) 公表番号	特表2014-516222 (P2014-516222A)		ンド 83 エルエルシー
(43) 公表日	平成26年7月7日 (2014. 7. 7)		アメリカ合衆国、89128 ネバダ州、
(86) 国際出願番号	PCT/US2012/038233		ラスベガス、ウエスト レイク ミード
(87) 国際公開番号	W02012/158859		ブルバード 7251、スイート 30
(87) 国際公開日	平成24年11月22日 (2012. 11. 22)		0
審査請求日	平成26年1月14日 (2014. 1. 14)	(74) 代理人	100107766
(31) 優先権主張番号	13/110, 085		弁理士 伊東 忠重
(32) 優先日	平成23年5月18日 (2011. 5. 18)	(74) 代理人	100070150
(33) 優先権主張国	米国 (US)		弁理士 伊東 忠彦
		(74) 代理人	100091214
			弁理士 大貫 進介

最終頁に続く

(54) 【発明の名称】 興味の特徴を含むビデオサマリー

(57) 【特許請求の範囲】

【請求項 1】

デジタル画像をキャプチャーするように構成された画像センサーと；

処理回路と、を含むシステムであって；

前記処理回路は、

前記画像センサーから一連の画像フレームを含んでいるビデオシーケンスを受取り、

前記ビデオシーケンスから分離した参照データを指定し、前記参照データは興味の特徴を示し、かつ、前記興味の特徴は所望の特性を含んでおり、

前記ビデオシーケンスを受取っている最中に、前記興味の特徴および前記所望の特性を含む前記画像フレームのサブセットを特定するために前記画像フレームを分析し、

前記ビデオシーケンスのビデオサマリーを形成し、前記ビデオサマリーは、前記興味の特徴および前記所望の特性を含んでいる画像フレームの前記サブセットの一部を含み、

前記ビデオサマリーは、前記ビデオシーケンスと並行して同時に生成される、

ように構成されている、

ことを特徴とするシステム。

【請求項 2】

前記システムは、さらに、

前記ビデオシーケンスを保管するように構成されたメモリーを含み、

前記処理回路は、さらに、

前記ビデオサマリーと前記ビデオシーケンスを前記メモリーの中に保管するように構

10

20

成されている、

請求項 1 に記載のシステム。

【請求項 3】

前記ビデオサマリーは、前記ビデオシーケンスから分離したファイルとして保管される

、
請求項 2 に記載のシステム。

【請求項 4】

前記処理回路は：

前記ビデオサマリーを処理するように構成された第 1 の処理パスと；

前記ビデオシーケンスを処理するように構成された第 2 の処理パスと、を含み、

前記第 1 の処理パスでは、前記第 2 の処理パスより画像フレームの解像度が低い、

請求項 1 に記載のシステム。

10

【請求項 5】

前記処理回路は、さらに、

前記ビデオシーケンスの中の追加のフレームを特定し、

前記追加のフレームを前記ビデオサマリーに追加する、

ように構成されており、

前記追加のフレームは、前記興味の特徴を含んでいる前記画像フレームのサブセットの
直前又は直後に位置するフレームを含む、

請求項 2 に記載のシステム。

20

【請求項 6】

前記追加のフレームは、圧縮されたビデオシーケンスの中で独立してエンコードされる

、
請求項 5 に記載のシステム。

【請求項 7】

前記所望の特性は、カメラ動作特性、カメラズーム特性、または、対象物動作特性のうち、少なくとも一つを含む、

請求項 1 に記載のシステム。

【請求項 8】

前記参照データは、前記興味の特徴を含んでいる参照画像、または、前記興味の特徴の
テキスト表現のうち、少なくとも一つを含む、

請求項 1 に記載のシステム。

30

【請求項 9】

コンピューターデバイスによって、画像センサーから一連の画像フレームを含んでいる
ビデオシーケンスを受取るステップと；

前記コンピューターデバイスによって、前記ビデオシーケンスから分離した参照データを
指定するステップであり、前記参照データは前記画像フレームに関する興味の特徴を示し、かつ、前記興味の特徴は所望の特性を含んでいる、ステップと；

前記コンピューターデバイスによって、前記ビデオシーケンスを受取る際に、前記興味
の特徴および前記所望の特性を含んでいる前記画像フレームのサブセットを特定するため
に前記画像フレームを分析するステップと；

前記コンピューターデバイスによって、前記ビデオシーケンスのビデオサマリーを形成
するステップであり、前記ビデオサマリーは、前記興味の特徴および前記所望の特性を含
んでいる画像フレームの前記特定されたサブセットの一部を含むステップと、を含み、

前記ビデオサマリーは、前記ビデオシーケンスと並行して同時に生成される、

ことを特徴とする方法。

40

【請求項 10】

前記方法は、さらに、

前記コンピューターデバイスによって、前記ビデオサマリー及び前記ビデオシーケンス
をメモリーの中に保管するステップ、を含む、

50

請求項 9 に記載の方法。

【請求項 1 1】

前記方法は、さらに、

前記コンピューターデバイスによって、前記ビデオシーケンスの中の追加のフレームを特定するステップであり、前記追加のフレームは、前記興味の特徴を含んでいる前記画像フレームのサブセットの直前又は直後に位置するフレームを含んでいるステップと、

前記コンピューターデバイスによって、前記追加のフレームを前記ビデオサマリーに追加するステップと、を含む、

請求項 1 0 に記載の方法。

【請求項 1 2】

インストラクションが保管されたコンピューターで読取り可能な固定の媒体であって、前記インストラクションは：

画像センサーから一連の画像フレームを含んでいるビデオシーケンスを受取るためのインストラクションと、

前記ビデオシーケンスから分離した参照データを指定するためのインストラクションであり、前記参照データは前記画像フレームに関する興味の特徴を示し、かつ、前記興味の特徴は所望の特性を含んでいる、インストラクションと、

前記ビデオシーケンスを受取る際に、前記興味の特徴および前記所望の特性を含んでいる前記画像フレームのサブセットを特定するために前記画像フレームを分析するインストラクションと、

前記ビデオシーケンスのビデオサマリーを形成するインストラクションであり、前記ビデオサマリーは、前記興味の特徴および前記所望の特性を含んでいる画像フレームの前記特定されたサブセットの一部を含むインストラクションと、を含み、

前記ビデオサマリーは、前記ビデオシーケンスと並行して同時に生成される、

ことを特徴とするコンピューターで読取り可能な固定の媒体。

【請求項 1 3】

前記媒体は、さらに、

前記ビデオサマリー及び前記ビデオシーケンスをメモリーの中に保管するインストラクション、を含む、

請求項 1 2 に記載のコンピューターで読取り可能な固定の媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、デジタルビデオ処理に関し、より特定的には、デジタルビデオサマリーを形成するためのシステム及び方法に関する。

【背景技術】

【0002】

多くのデジタルキャプチャーデバイスは、静止画像はもちろんビデオをキャプチャーすることができる。デジタルビデオコンテンツを管理することは、しかしながら、難しい仕事である。ビデオは、しばしば、ビデオの最初のフレームのサムネイル画像を用いて視覚的に表現される。これは、ビデオのコンテンツについて多くの洞察ができないことがある。与えられたビデオの中に特定のイベントが含まれているかを判断することは、しばしば、ビデオフレームの全体を視聴することを要求するものである。長時間のビデオに対して、ユーザーにとっては、そのビデオ全体を視聴することなく、ビデオのクイックサマリーを得ることができることが好ましい。

【0003】

デジタルビデオは、また、共有の観点からの現実的な問題を表している。多くのデジタルキャプチャーデバイスは、毎秒 30 から 60 フレームを、1920 × 1080 ピクセルの高解像度でビデオを記録する。圧縮した場合においてさえ、生成されたビデオの量は、比較的短時間のビデオでさえ共有することを非現実的なものとしている。

10

20

30

40

50

【0004】

より容易に共有することができるより短いバージョンへとビデオを手動で集約するために、ビデオ編集ソフトウェアを使用することができる。しかしながら、手動のビデオ編集は、時間が掛かり、骨の折れるプロセスであり、多くのユーザーは手動の編集に興味はない。自動ビデオ集約アルゴリズムも、同様に存在する。こうしたソリューションは、入力としてキャプチャーされたビデオを用いて開始され、ビデオサマリーを決定するためにビデオを分析する。発明者ラタコンダ (Ratakonda) の米国特許第5995095号明細書、タイトル“Method for hierarchical summarization and browsing of digital video”は、ビデオシーケンスに基づいて階層的なサマリーを生成する方法を開示している。発明者リら (Liet al.) の米国特許第7035435号明細書、タイトル“Scalable video summarization and navigation system and method”は、ビデオの中のそれぞれのシーン、ショット、および、フレームに対して重要度を割当て、ビデオサマリーのための主要フレームを決定するのに重要度を使用する方法が記載されている。発明者エドワードら (Edward et al.) の米国特許第7483618号明細書、タイトル“Automatic editing of a visual recording to eliminate content of unacceptably low quality and/or little or no interest”は、ビデオサマリーを決定する方法を開示しており、クオリティが低いか、興味がないコンテンツはビデオから削除される。

10

20

【0005】

しかしながら、自動ビデオ集約アルゴリズムは非常に複雑である。ビデオサマリーを決定するのに必要な分析を実行するためにビデオをデコードする必要があるからである。従って、デジタルキャプチャーデバイス上で、ジャストキャプチャービデオ (just-captured video) に対応するビデオサマリーを直ちに視聴することはできない。この欠点が、素早いレビューとキャプチャーされたビデオの共有を促進することを難しくしている。

【0006】

ビデオサマリーを作成する場合、しばしば、サマリーの中に所定の特徴を有することが望まれる。ビデオサマリーは、その中に特徴が表現されているいくつか又は全てのビデオコンテンツを含むように作成される。こうした特徴の例は、人、ペット、イベント、場所、動作、および、対象物を含んでいる。手動で、こうした注文仕立てのビデオサマリーを作成することは、退屈なプロセスであり得るものである。注文仕立てのこうしたビデオサマリーを生成するためにデスクトップソフトウェアを使用することは、素早くレビューしてビデオサマリーを共有する可能性を妨げてしまう。

30

【0007】

従って、デジタルキャプチャーデバイスにおいてビデオサマリーを算出するためのシステム及び方法を提供することが望ましい。特に、ビデオキャプチャーが完了して最小の遅れをもって、デジタルキャプチャーデバイス上でビデオサマリーを生成することができるソリューションを提供することが望ましいであろう。また、ユーザー所定の特徴を含むビデオサマリーを提供することも望ましいであろう。

40

【先行技術文献】

【特許文献】

【0008】

【特許文献1】米国特許第5995095号明細書

【特許文献2】米国特許第7035435号明細書

【特許文献3】米国特許第7483618号明細書

【特許文献4】米国特許公開第2007/0024931号公報

【特許文献5】米国特許第6292218号明細書

50

【特許文献6】米国特許第5668597号明細書

【特許文献7】米国特許出願第2011/0292288号公報

【特許文献8】米国特許出願第2011/0293018号公報

【発明の概要】

【発明が解決しようとする課題】

【0009】

本発明は、興味の特徴を含んでいるビデオフレームのサブセット(subset)を決定するためにキャプチャーする時点でビデオフレームを分析するという利点を有している。それにより、ビデオサマリーを作成する時点でビデオデータを解凍する必要を取り除いている。

10

【課題を解決するための手段】

【0010】

本発明は、ビデオシーケンスをキャプチャーして、関連するビデオサマリーを提供するためのデジタルビデオカメラシステムを表している。本発明に係るシステムは、

デジタル画像をキャプチャーするための画像センサーと；

前記画像センサー上にシーンの画像を形成するための光学系と；

データ処理システムと；

キャプチャーされたビデオシーケンスを保管するためのストレージメモリと；

前記データ処理システムに通信可能に接続されたプログラムメモリであり、前記データ処理システムにビデオサマリーを形成するための方法を実施するようにさせるように構成されたインストラクションを保管しているプログラムメモリと、

20

を含み、

前記方法は、

参照データを指定するステップであり、前記参照データは興味の特徴を示すステップと；

前記画像センサーを使用して前記シーンのビデオシーケンスをキャプチャーするステップであり、前記ビデオシーケンスは画像フレームのタイムシーケンスを含んでいるステップと；

デジタルビデオファイルを形成するためにビデオ処理パスを使用して前記キャプチャーされたビデオシーケンスを処理するステップと；

30

特徴認識アルゴリズムを使用して、前記興味の特徴を含む前記画像フレームのサブセットを特定するために、前記キャプチャーされた画像フレームを分析するステップと；

前記キャプチャーされたビデオシーケンスの中の全ての画像フレームより少ない画像フレームを含む前記ビデオサマリーを形成するステップであり、前記ビデオサマリーは、前記興味の特徴を含んでいる前記特定された画像フレームのサブセットを少なくとも部分的に含んでいるステップと；

前記デジタルビデオファイルを前記ストレージメモリの中に保管するステップと；

前記ビデオサマリーの表現を前記ストレージメモリの中に保管するステップと、

を含むことを特徴としている。

【0011】

40

本発明は、興味の特徴を含んでいるビデオフレームのサブセット(subset)を決定するためにキャプチャーする時点でビデオフレームを分析するという利点を有している。それにより、ビデオサマリーを作成する時点でビデオデータを解凍する必要を取り除いている。

【0012】

本発明は、前記保管されたデジタルビデオシーケンスを解凍する必要なく前記ビデオサマリー表現を保管するという追加の利点を有している。これにより、ビデオキャプチャーの完了の時点で最小の遅れをもって、ビデオサマリーを生成し、デジタルキャプチャーデバイス上で視聴することができる。

【0013】

50

いくつかの実施例において、ビデオサマリーは、メタデータを使用してデジタルビデオファイルにおいてエンコードされ、分離したファイルとしてビデオサマリーをエンコードする必要はない。このことは、ビデオサマリーメタデータを理解する「賢い(“smart”)」ビデオプレーヤーを用いて視聴するために、ビデオサマリーが便利に利用可能であるという利点を有しており、一方従来のプレーヤーに対してはトランスペアレント(transparent)である。

【図面の簡単な説明】

【0014】

【図1】図1は、ビデオサマリーを形成するためのシステムに係るコンポーネントを示すハイレベルの模式図である。

【図2】図2は、ビデオサマリーを形成する方法に係るフローチャートである。

【図3】図3は、デジタルビデオシーケンスの処理のための2つの異なるビデオ処理パスの使用を説明するフローチャートである。

【図4】図4は、代替的な実施例に従って、デジタルビデオシーケンスの処理のための2つの異なるビデオ処理パスの使用を説明するフローチャートである。

【図5】図5は、第2の実施例に従って、ビデオサマリーを生成する方法に係るフローチャートである。

【図6】図6は、第3の実施例に従って、ビデオサマリーを生成する方法に係るフローチャートである。

【発明を実施するための形態】

【0015】

以降の記載において、本発明の好適な実施例が通常にソフトウェアプログラムとして実行されるであろう形式で説明される。当業者であれば、そうしたソフトウェアの均等物が、また、ハードウェアにおいても構成され得ることを容易に理解するであろう。なぜなら、そうしたソフトウェアの画像操作のアルゴリズム及びシステムはよく知られたものであり、本記載は、特に、本発明が関するシステム及び方法の一部を形成する、または、より直接的に協同するアルゴリズム及びシステムに向けられたものだからである。こうしたアルゴリズム及びシステム、および、ここにおいてに包含される画像信号を生成して、そうでなければ処理を行う他の態様は、ここにおいて特に示され又は説明されないが、従来技術として知られる、そうしたシステム、アルゴリズム、コンポーネント、および、エレメントから選択され得るものである。以降の資料における本発明に従って説明されたシステムが与えられれば、本発明の実施のために役に立つソフトウェアは、ここにおいて特に示され、示唆され、または、説明されていなくても、一般的なものであり、従来技術の範囲内のものである。

【0016】

さらにまた、ここにおいて使用されるように、本発明に係る方法を実施するために使用されるコンピュータープログラムは、コンピューターで読取り可能な記憶媒体に保管され得る。例えば、磁気ディスク(ハードドライブ又はフロッピー(登録商標)ディスクといったもの)又は磁気テープといった磁気記録媒体；光ディスク、光テープ、又は機械で読取り可能なバーコードといった光記憶媒体；ランダムアクセスメモリー(RAM)又は読出し専用メモリー(ROM)といった半導体電子記憶媒体；または、本発明に従った方法を実行するように一つまたはそれ以上のコンピューターをコントロールするためのインストラクションを有するコンピュータープログラムを保管するために使用されるあらゆる他の物理的デバイス又は媒体、を含んでいる。

【0017】

本発明は、ここにおいて説明される実施例の組合せについて包括的なものである。「所定の実施例」への言及および類似のものは、本発明の少なくとも一つの実施例において存在する特徴に言及するものである。「一つの実施例」又は「所定の実施例」もしくは類似のものへ離れて言及することは、必ずしも同一の実施例を参照する必要はない。しかしながら、こうした実施例はお互いに排他的なものではない。そうであると示唆され又は当業者に

10

20

30

40

50

とって容易に明らかではない場合である。「方法」に対する言及における単数形又は複数形の使用は、限定的なものではない。そうでないことが明示的に示され、または文脈から要求されるものでなければ、用語「または」は、この発明開示において非排他的な意味で使用されていることに留意すべきである。

【 0 0 1 8 】

デジタルカメラは、画像デバイス及び信号キャプチャーと処理のために関連する回路を使用しており、ディスプレイはよく知られているので、本発明の説明は、本発明に従った方法及び装置の一部を形成する、または、より直接的に協同する所定のエレメントに対して向けられている。ここにおいて、特に示されていないか、記載されていないエレメントは、従来技術から選択されたものである。説明される実施例の所定の態様は、ソフトウェアにおいて提供される。以降の資料における本発明に従って説明されたシステムが与えられれば、本発明の実施のために役に立つソフトウェアは、ここにおいて特には示され、示唆され、または、説明されていない場合でも、一般的なものであり、従来技術の範囲内のものである。

10

【 0 0 1 9 】

以降のデジタルカメラに係る説明は、当業者はよく知っていることである。本実施例に係る多くのバリエーションが可能であり、コスト削減のため、カメラのパフォーマンスに特徴を追加または改善するために選択できることが明らかである。

【 0 0 2 0 】

図 1 は、デジタル写真システムのブロックダイアグラムを示しており、本発明に従ってビデオ画像をキャプチャーすることができるデジタルカメラ 10 を含んでいる。おそらく、デジタルカメラ 10 は、ポータブル電池で動作するデバイスであり、画像をキャプチャー又はレビューする際にユーザーによって容易に取扱うことができるよう十分に小さいものである。デジタルカメラ 10 は、デジタル画像を生成し、ストレージメモリ 30 を使用してデジタル画像ファイルとして保管する。「デジタル画像」又は「デジタル画像ファイル」という用語は、ここにおいて使用されるように、デジタル静止画像又はデジタルビデオファイルといった、あらゆるデジタル画像ファイルを参照するものである。

20

【 0 0 2 1 】

いくつかの実施例において、デジタルカメラ 10 は、モーションビデオ画像と静止画像の両方をキャプチャーする。他の実施例において、デジタルカメラ 10 は、モーションビデオ画像のみをキャプチャーするデジタルカメラである。デジタルカメラ 10 は、また、他の機能も含み得る。これらに限定されるわけではないが、デジタル音楽プレーヤー（例えば、MP3 プレーヤー）、音声録音デバイス、携帯電話、GPS 受信器、または、プログラム可能デジタルアシスタンス（PDA）の機能を含んでいる。

30

【 0 0 2 2 】

デジタルカメラ 10 は、調整可能な絞りと調整可能なシャッター 6 を有するレンズ 4 を含んでいる。好適な実施例において、レンズ 4 は、ズームレンズであり、ズームとフォーカスモータードライバー 8 によってコントロールされる。レンズ 4 は、シーン（図示なし）からの光を画像センサー 14 上にフォーカスする。センサーは、例えば、シングルチップカラー CCD 又は CMOS 画像センサーである。レンズ 4 は、画像センサー 14 上にシーンに係る画像を形成するための光学システムの一つのタイプである。他の実施例において、光学システムには、可変又は固定フォーカスのいずれかをもった固定焦点距離レンズが使用され得る。

40

【 0 0 2 3 】

画像センサー 14 の出力は、アナログ信号プロセッサ（ASP）及びアナログ - デジタル変換（A/D）コンバーター 16 によってデジタル形式へ変換され、一時的にバッファメモリー 18 の中に保管される。バッファメモリー 18 の中に保管された画像データは、続いて、ファームウェアメモリー 28 の中に保管されているエンベットされたソフトウェアプロセッサ（例えば、ファームウェア）を使用して、プロセッサ 20 によって操作される。いくつかの実施例において、ソフトウェアプログラムは、読出し専用メモリー（

50

ROM)を使用してファームウェアメモリーの中に永久に保管されている。他の実施例において、ファームウェアメモリー28は、例えば、フラッシュEPROMメモリーを使用して、変更され得る。そうした実施例において、外部デバイスは、有線インターフェイス38又は有線モデム50を使用して、ファームウェアメモリー28の中に保管されたソフトウェアプログラムを更新することができる。そうした実施例において、ファームウェアメモリー28は、また、画像センサー較正データ、ユーザー設定選択、および、カメラが電源オフされる際に保持されなければならない他のデータを保管するためにも使用され得る。いくつかの実施例において、プロセッサ20は、プログラムメモリー(図示なし)を含み、ファームウェアメモリー28の中に保管されたソフトウェアプログラムは、プロセッサ20によって実行される前にプログラムメモリーの中へとコピーされる。

10

【0024】

プロセッサ20の機能は、シングルプログラム可能プロセッサを使用して、または、マルチプログラム可能プロセッサの使用によって提供されることが理解されよう。一つまたはそれ以上のデジタル信号プロセッサ(DSP)デバイスを含むものである。代替的に、プロセッサ20は、カスタム回路(例えば、デジタルカメラでの使用のために特別にデザインされた一つまたはそれ以上のカスタム集積回路(IC))、または、プログラム可能プロセッサとカスタム回路との組合せによって提供され得る。プロセッサ20と図1に示される種々のコンポーネントのいくつか又は全てのものとの間のコネクタは、共通のデータバスを使用して作成され得る。例えば、いくつかの実施例において、プロセッサ20と、バッファメモリー18、ストレージメモリー30、および、ファームウェアメモリー28との間の接続が、共通のデータバスを使用して作成され得る。

20

【0025】

処理された画像は、次に、ストレージメモリー30を使用して保管される。ストレージメモリー30は、当業者に知られているあらゆる形式のメモリーであってよいことが理解されるべきである。これらに限定されるわけではないが、取出し可能フラッシュメモリーカード、内部フラッシュメモリーチップ、磁気メモリー、または、光メモリーを含んでいる。いくつかの実施例において、ストレージメモリー30は、内部フラッシュメモリーチップと、セキュリティデジタル(SD)カードといった、取出し可能フラッシュメモリーカードへの標準インターフェイスの両方を含むことができる。代替的に、異なるメモリーカードフォーマットが使用され得る。マイクロSDカード、コンパクトフラッシュ(登録商標)(CF)カード、マルチメディアカード(MMC)、xDカード、または、メモリースティックといったものである。

30

【0026】

画像センサー14は、タイミングジェネレーター12によってコントロールされており、列(row)とピクセルを選択し、ASPとA/Dコンバーター16のオペレーションを同期させるために種々のクロック信号を生成する。画像センサー14は、例えば、およそ4000x3000ピクセルの静止画像ファイルを生成するために12.4メガピクセル(4088x3040ピクセル)を有することができる。カラー画像を生成するために、画像センサーは一般的にカラーフィルターアレイ(array)を用いてオーバーレイされ、異なるカラーピクセルを含むピクセルのアレイを有する画像センサーを提供する。異なるカラーピクセルは、多くの異なるパターンで配置され得る。一つの実施例として、周知のバイヤー(Bayer)カラーフィルターアレイを使用して配置され得る。発明者バイヤー(Bayer)の米国特許第3971065号明細書、タイトル"Color imaging array"に記載されているようにである。第2の実施例として、異なるカラーピクセルが配置され得る。発明者コンプトン(Compton)とハミルトン(Hamilton)の米国特許公開第2007/0024931号公報、タイトル"Image sensor with improved light sensitivity"に記載されているようにである。これらの実施例は限定的なものではなく、多くの他のカラーパターンが使用されてよい。

40

【0027】

50

画像センサー14、タイミングジェネレーター12、および、ASPとA/Dコンバーター16は、分離して集積回路を作り上げることができ、または、COMS画像センサーを用いて一般的に行われているように単一集積回路として作り上げられてもよい。いくつかの実施例において、この単一集積回路は、図1に示される他の機能のいくつかを実施例することができる。プロセッサ20によって提供される機能のいくつかを含むものである。

【0028】

画像センサー14は、タイミングジェネレーター12によって第1のモードにおいて作動された場合に、より低解像度のセンサー画像データのモーションシーケンスを提供するのに有効である。画像を構成するために、ビデオ画像をキャプチャーする際に使用され、キャプチャーされるべき静止画像をプレビューする際にも使用される。このプレビューモードセンサー画像データは、HD解像度の画像データ、例えば、1280×720ピクセルを用いて提供され得る。または、VGA解像度の画像データ、例えば、640×480ピクセルを用いて、もしくは、画像センサーの解像度に比べて非常に少ないデータの行と列を有する他の解像度を使用して提供され得る。

10

【0029】

プレビューモードセンサー画像データは、同一のカラーを有する隣接したピクセルの値を結合することにより、または、ピクセル値のいくつかを除去することにより、もしくは、他のカラーのピクセルの値を除去する一方でいくつかのカラーピクセルの値を結合することによって提供され得る。プレビューモード画像データは、発明者パルルスキーら(P Prulski et al.)の米国特許第6292218号明細書、タイトル"Electric camera for initiating capture of still images while previewing motion images"に記載されているようにである。

20

【0030】

画像センサー14は、また、タイミングジェネレーター12によって第2のモードにおいて作動された場合に、より高解像度のセンサー画像データのモーションシーケンスを提供するのに有効である。この最終モードセンサー画像データは、高解像度出力画像データとして提供され、高い輝度レベルを有するシーンのために画像センサーの全てのピクセルを含んでおり、例えば、4000×3000ピクセルを有する12メガピクセルの最終画像データであり得る。より低い輝度レベルにおいて、最終センサー画像データは、画像センサー上のいくつかの数量の類似カラーピクセルを「ビニング("binning")」することによって提供され得る。信号レベルを増加するためであり、従って、センサーの「ISO速度("ISO speed")」が増す。

30

【0031】

ズームとフォーカスモータードライバー8は、プロセッサ20によって供給されるコントロール信号によってコントロールされ、好適な焦点距離を提供して、画像センサー上にシーンをフォーカスする。画像センサー14の露出レベルは、F値(f/number)と調整可能な絞り及び調整可能なシャッターの露出時間、タイミングジェネレーター12を介した画像センサーの露出期間、および、ASPとA/Dコンバーター16のゲイン設定(例えば、ISO速度)をコントロールすることによってコントロールされる。プロセッサ120は、また、シーンを照明することができるフラッシュ2もコントロールする。

40

【0032】

デジタルカメラ10のレンズ4は、「スルーザレンズ("through-the-lens")」オートフォーカスを使用することで第1モードにおいてフォーカスされ得る。発明者パルルスキーら(P Prulski et al.)の米国特許第5668597号明細書、タイトル"Electronic Camera with Rapid Automatic Focus of an Image upon a Progressive Scan Image Sensor"に記載されているようにである。このことは、ズームとフォーカスドライバー8を使用して、レンズ4の焦点位置を近接焦点位

50

置から無限大焦点位置の間の範囲の数多くの位置に対して調整することによって達成される。一方、プロセッサ20は、画像センサー14によってキャプチャーされた画像の中心部分に対して最大のシャープネス値 (sharpness value) を提供する最も近い焦点位置を判断する。最も近い焦点位置に対応する焦点距離は、次に、好適なシーンモードの自動設定といった、いくつかの目的で使用され得るし、他のレンズとカメラの設定を伴って、画像ファイルにおけるメタデータとして保管され得る。

【0033】

プロセッサ20は、メニューと低解像度カラー画像を生成し、それらはディスプレイメモリ36の中に一時的に保管され、画像ディスプレイ32上に表示される。画像ディスプレイ32は、典型的には、アクティブマトリクスカラー液晶ディスプレイ (LCD) 10 であるが、有機発光ダイオードディスプレイといった、他のタイプのディスプレイが使用され得る。ビデオインターフェイス44は、デジタルカメラ10からのビデオ出力信号を、フラットパネルHDTVディスプレイといった、ビデオディスプレイ46を提供する。プレビューモード、または、ビデオモードにおいて、バッファメモリー18からのデジタル画像データは、プロセッサ20によって操作され、典型的にはカラー画像として、画像ディスプレイ32上に表示されるモーションプレビュー画像のシリーズを形成する。レビューモードにおいて、画像ディスプレイ32上に表示される画像は、ストレージメモリー30の中に保管されたデジタル画像ファイルから画像データを使用して生成される。

【0034】

画像ディスプレイ32上に表示されるグラフィックユーザーインターフェイスは、ユーザーコントロール34によって提供されるユーザー入力に応じてコントロールされる。ユーザーコントロール34は、ビデオキャプチャーモード、静止画キャプチャーモード、および、レビューモードといった種々のカメラモードを選択するために、そして、静止画像のキャプチャーとモーション画像の記録を開始するために使用される。いくつかの実施例において、上記の第1モード (つまり、静止プレビューモード) は、ユーザーがシャッターボタンを部分的に押す際に開始される。これは、ユーザーコントロール34のうちの一つであり、ユーザーが完全にシャッターボタンを押す際に第2モード (つまり、静止画像キャプチャーモード) が開始される。ユーザーコントロール34は、また、カメラを電源オンし、レンズ4をコントロールし、そして、写真撮影プロセスを開始するためにも使用される。ユーザーコントロール34は、典型的には、ボタン、ロッカー (rocker) 30 スイッチ、ジョイスティック、または、回転ダイアルのいくつかの組合せを含んでいる。いくつかの実施例において、ユーザーコントロール34のいくつかは、画像ディスプレイ32上にオーバーレイされたタッチスクリーンを使用して提供される。他の実施例において、追加のステータスディスプレイ又は画像ディスプレイが使用され得る。

【0035】

ユーザーコントロール34を使用して選択することができるカメラモードは、「タイマー」モードを含んでいる。「タイマー」モードが選択された場合、ユーザーが完全にシャッターボタンを押してから、プロセッサ20が静止画像キャプチャーを開始するまでに短い遅延 (例えば、10秒) が生じる。

【0036】

プロセッサ20に接続された音声コーデック (codec) 22は、マイクロフォン24から音声信号を受取り、音声信号をスピーカー26に対して提供する。これらのコンポーネントは、ビデオシーケンス又は静止画像と一緒に、音声トラックを記録してプレイバック (playback) することができる。デジタルカメラ10が、カメラと携帯電話の組合せといったマルチ機能デバイスである場合、電話での会話のためにマイクロフォン24とスピーカー26が使用され得る。

【0037】

いくつかの実施例において、スピーカー26は、ユーザーインターフェイスの一部として使用され得る。例えば、ユーザーコントロールが押されたこと、または、所定のモードが選択されたことを示す種々の音声信号を提供する。いくつかの実施例において、マイク 50

ロフォン24、音声コーデック22、および、プロセッサ20は、音声認識を提供するために使用され、ユーザーは、ユーザーコントロール34より、むしろ、音声コマンドを使用してプロセッサ20に対してユーザー入力を提供することができる。スピーカ26は、また、電話の着信をユーザーに知らせるためにも使用され得る。このことは、ファームウェアメモリ28の中に保管されている標準のリングトーン(ring tone)を使用して、または、無線ネットワーク58からダウンロードされ、ストレージメモリ30の中に保管されたカスタムリングトーンを使用することによって行われ得る。加えて、着信電話の無音(例えば、非音声)の通知のためにバイブレーションデバイス(図示なし)が使用され得る。

【0038】

10

いくつかの実施例において、デジタルカメラ10は、また、カメラのあらゆる動作に関するデータを提供する加速度計をも含んでいる。望ましくは、加速度計27は、3つの直交する方向のそれぞれ(合計6方向の入力について)に対して並進および回転加速度の両方を検知する。

【0039】

プロセッサ20は、また、画像センサー14からの画像データの追加的な処理を提供する。表現されたsRGB画像データを生成するためであり、ストレージメモリ30における「完了」画像ファイル、よく知られたExif-JPEG画像ファイルといったもの、の中に圧縮され保管される。

【0040】

20

デジタルカメラ10は、有線インターフェイス38を介してコンピューター40に接続されたインターフェイス/リチャージャー48に接続されている。コンピューターは、家又はオフィスにあるデスクトップコンピューター又はポータブルコンピューターであり得る。有線インターフェイス38は、例えば、よく知られたUSB2.0インターフェイス仕様に従うものでよい。インターフェイス/リチャージャー48は、有線インターフェイス38を介して、デジタルカメラ10の一式の再充電可能バッテリー(図示なし)に対して電力を供給することができる。

【0041】

デジタルカメラ10は、無線モデム50を含み、ラジオ周波数帯52上で無線ネットワーク58とインターフェイスする。無線モデム50は、種々の無線インターフェイスプロトコルを使用することができる。よく知られたブルートゥース無線インターフェイス又はよく知られた802.11無線インターフェイスといったものである。コンピューター40は、インターネット70を介して、Kodak EasyShare Galleryといった、写真サービスプロバイダー72に画像をアップロードすることができる。他のデバイス(図示なし)は、写真サービスプロバイダー72によって保管された画像にアクセスすることができる。

30

【0042】

代替の実施例において、無線モデム50は、ラジオ周波数(例えば、無線)リンク上で、3GSM(登録商標)ネットワークといった、携帯電話ネットワーク(図示なし)とコミュニケーションする。デジタルカメラ10からデジタル画像ファイルをアップロードするためにインターネット70と接続するものである。こうしたデジタル画像ファイルは、コンピューター40又は写真サービスプロバイダー72に提供され得る。

40

【0043】

これから、図2に関して本発明が説明される。最初に、ビデオシーケンスキャプチャステップ200において、画像センサー14を使用してデジタル画像シーケンス205をキャプチャするために、デジタルカメラ10(図1)といったデジタルビデオキャプチャデバイスが使用される。デジタルビデオシーケンス205は画像フレームのタイムシーケンスを含んでいる。

【0044】

ビデオシーケンスキャプチャステップ200の最中、画像センサー14を使用して参

50

照画像 215 をキャプチャーするために参照画像キャプチャーステップ 210 が実行される。参照画像 215 は、所定の人間を含んでいる。参照画像 215 は、キャプチャーされたデジタルビデオシーケンス 205 からの画像フレームであってよく、参照画像 215 は、ユーザーコントロール 34 の使用を通じて選択され得る。例えば、ビデオキャプチャーの最中に、ユーザーは好適なユーザーコントロールボタンを押すことによって参照画像 215 のキャプチャーを要求することができる。ボタンは、バッファメモリー 18 における現在のビデオフレームを参照画像 215 として指定するように、プロセッサ 20 に対して信号を送付する。参照画像 215 は、所定の興味のある人間を含んでいる。望ましい実施例において、参照画像 215 は、十分な空間解像度で所定の人間の顔の前面を含んでおり、参照画像 215 から顔面特徴を決定することができる。いくつかの実施例において、参照画像 215 は、一人の人間を唯一含んでおり、所定の人間が、参照画像 215 における唯一の人間として特定される。他の実施例において、参照画像 215 は、複数の人間を含んでおり、所定の人間は多くの手段のうちの一つによって選択され得る。最も大きな人間を選択、参照画像 215 の中心に最も近い人間を選択、視認できる最も大きな顔面をもった人間を選択、または、好適なユーザーインターフェイスを使用して手動で人間を選択（例えば、ポインティングデバイスを使用してユーザーに顔面を選択させることによる）を含んでいる。代替的に、参照画像 215 におけるそれぞれの人間が、所定の人間であることを指定され得る。いくつかの実施例において、所定の人間は、参照画像 215 に在る顔面を顔面認識データベースにおける既知の顔面と比較し、既知の顔面を選ぶことによって、選択され得る。他の実施例においては、ユーザーが手動で参照画像 215 の中に含まれる所定の人間を指定できるようにユーザーインターフェイスが提供される。

【0045】

キャプチャーされたビデオシーケンス処理ステップ 220 は、ビデオ処理パス (path) を使用してキャプチャーされたデジタルビデオシーケンス 205 を処理して、デジタルビデオファイル 225 を形成する。ビデオ処理パスは、例えば、ビデオ圧縮ステップを含み得る。MPEG 及び H.263 規格において規定されているようにビデオ圧縮アルゴリズムは、当業者によく知られている。

【0046】

デジタルビデオシーケンス 205 のキャプチャーの最中に、キャプチャーされた画像フレーム分析ステップ 240 は、所定の人間を含む画像フレーム 245 のサブセットを特定するために人間認識アルゴリズムを使用して、キャプチャーされたデジタルビデオシーケンス 205 における画像フレームを分析する。望ましい実施例において、人間認識アルゴリズムは、顔面認識アルゴリズムであってよい。キャプチャーされた画像フレーム分析ステップ 240 は、参照画像における所定の人間の顔面と同一の顔面を含んでいる画像フレームを特定する。顔面認識アルゴリズムは、よく知られた従来技術である。例えば、著者タークらの (Turk et al.) 記事 “Eigenfaces for Recognition” (Journal of Cognitive Neuroscience, Vol. 3, pp. 71 - 86, 1991) は、本発明と一緒に使用することができる顔面認識アルゴリズムを記載している。代替的に、人間認識アルゴリズムは、画像フレームの中の人間を参照画像 215 における所定の人間と一致されるあらゆるアルゴリズムであってよい。こうしたアルゴリズムは、性別分類、身長見積り、および、衣服分析を含み、当業者にとって知られたそうしたアルゴリズムの中から選択され得る。

【0047】

ビデオサマリー形成ステップ 250 は、キャプチャーされたデジタルビデオシーケンス 205 における全ての画像フレームより少ない画像フレームを含むビデオサマリー 255 を形成する。ビデオサマリー 255 は、所定の人間を含む画像フレーム 245 の特定されたサブセットの少なくとも一部を含んでいる。本発明の一つの実施例において、ビデオサマリー 255 は、所定の人間を含んでいるそれらの画像フレームだけを使用して形成される。

【0048】

10

20

30

40

50

いくつかの実施例において、ビデオサマリー 255 は、所定の人間を含んでいる画像フレーム 245 のサブセットの一部のみを含んでいる。例えば、ビデオサマリー 255 は、規定の数量より少ない画像フレームを含むように制限され得る。代替的な実施例において、ビデオサマリー 255 は、画像フレーム 245 のサブセットにおける連続した画像フレームのそれぞれのグループから、一つの画像フレームを含むことができる。このようにして、ビデオサマリー 255 は、所定の人間を含んでいる画像フレーム 245 の特定されたサブセットから選択された一式の静止画像から成る「スライドショー (“ s l i d e s h o w ”) 」であり得る。

【 0 0 4 9 】

別の実施例において、ビデオ 255 は、また、所定の人間を含んでいる画像フレーム 245 の特定されたサブセットの他に、追加の画像フレームを含んでいる。本発明の一つの実施例において、追加の画像フレームは、所定の人間を含んでいる画像フレーム 245 の特定されたサブセットの中の画像フレームの直前又は直後の画像フレームを含んでいる。そうしたフレームは、例えば、遷移フレームとして、ビデオサマリー 255 の一つのセクションから別のセクションまでの遷移期間を許容するように選定される。そうしたフレームは、また、ビデオサマリー 255 がデジタルビデオファイル 225 から容易に抽出できる画像のグループを含むように、選定され得る。MPEG といったビデオ圧縮規格は、いくつかのフレームが独立して（他のフレームを参照することなく）エンコードされるように、ビデオシーケンスをエンコードする。そして、一時的なシーケンシャルフレームのいくつかのグループ又は画像のグループが、画像のグループの外のあらゆるフレームを参照することなくエンコードされるようにである。従って、こうした画像のグループを表している圧縮されたビデオデータは、圧縮されたビデオデータをデコードする必要なく、圧縮されたデジタルビデオファイル 225 から抽出され得るものである。

【 0 0 5 0 】

本発明の別の実施例において、追加の画像フレームは、キャプチャーされたデジタルビデオシーケンス 205 の重要な部分であると判断された他の部分を含んでいる。デジタルビデオシーケンス 205 のこうした他の重要な部分は、主要フレーム抽出、または、ビデオサマリアルゴリズムを実行することにより特定され得る。こうしたアルゴリズムは、発明者ディーバー (Deever) の米国特許出願第 2011/0292288 号公報、タイトル “ Method for determining key video frames ”、2011 年 12 月 1 日出版、および、発明者ディーバー (Deever) の米国特許出願第 2011/0293018 号公報、タイトル “ Video summary method and system ” において説明されている。

【 0 0 5 1 】

米国特許出願第 2011/0293018 号公報には、ビデオサマリーを形成する方法が開示されており、画像フレームが、特徴値を判断するためにキャプチャーの時点において分析される。こうした特徴値は、圧縮されたビデオシーケンスを解凍することなく分析される。ビデオサマリーを含む主要なビデオスニペット (snippet) を特定するためである。

【 0 0 5 2 】

米国特許出願第 2011/0292288 号公報には、主要なビデオスニペットを特定する方法が開示されており、キャプチャーの時点で判断された特徴値を使用して、デジタルビデオシーケンスが分析される。時間の関数として重要値を決定するためである。重要値は、デジタルビデオシーケンスに対するワーブタイム (warped-time) 表現を形成するために使用される。ワーブタイム表現は、デジタルビデオシーケンスの中のいくつかのビデオシーケンスに対してより大きな時間的加重付けを与え、他のものに対してより小さな時間的加重付けを与える。ワーブタイムデジタルビデオシーケンス表現は、一式の等しい時間インターバルへと分割される。主要フレームが、対応する時間的インターバルにおけるビデオフレームに関する特徴値を分析することによって、それぞれの時間的インターバルから選択される。そうしたクライテリアは、ズームイン工程の完了後に間も

10

20

30

40

50

なく生じる主要ビデオフレームを選択すること、および、ビデオフレームの中心領域において並みのレベルのローカルな動作を有する主要なビデオフレームを選択することを含み得る。

【0053】

いくつかの実施例において、主要なビデオフレームスニペットは、主要なビデオフレームのそれぞれを取り囲んでいる一式のビデオフレームを選択することによって形成され得る。例えば、主要なビデオスニペットは、主要なビデオフレームの前後2秒間のビデオフレームを選択することによって形成され得る。こうし、4秒間の長い主要なビデオスニペットを形成している。

【0054】

代替的には、主要なビデオフレームフレームがランク付けされて、最も高くランク付けされた主要なビデオフレームに対応する主要なビデオフレームのサブセットだけに対して主要なビデオフレームスニペットが形成され得る。主要なビデオフレームをランク付けすることは、カメラ固定パターンを決定するためにデジタルビデオシーケンスを分析すること、および、カメラ固定パターンに応じて主要ビデオフレームをランク付けすることを含み得る。キャプチャー時点で判断されたグローバルな動作に対応する特徴値は、デジタルビデオシーケンスの全体を通してカメラの固定領域を示すグローバルな動作トレースを提供することができる。高い固定領域に対応しているビデオフレーム、つまり、ビデオキャプチャー全体の高い割合でカメラが固定された状態で居る領域は、高くランク付けされる。ランク付けプロセスは、それぞれのステップで最高位の次にランク付けされる主要なビデオフレームを選択することで、反復的に実行され得る。そして、それぞれのステップで、プロセスは、既にランク付けプロセスに含まれている固定領域を表している主要なビデオフレーム上の表現されていない固定領域を表している主要なビデオフレームを促進することができる。主要なビデオフレームがランク付けされると、主要なビデオスニペットの中に含まれるように、トップランクの主要なビデオフレームが選定され得る。こうしたビデオスニペットによって表された追加の画像フレームは、ビデオサマリーを形成するために、所定の人間を含んでいる画像フレームの特定されたサブセットと結合され得る。

【0055】

デジタルビデオファイル保管ステップ230は、デジタルビデオファイル225をストレージメモリ30に保管し、保管されたデジタルビデオファイル235を生成する。ビデオサマリー表現保管ステップ260は、ストレージメモリ30の中にビデオサマリー表現265を保管する。一つの実施例において、ストレージメモリ30の中にビデオサマリー表現265を保管することは、ビデオサマリー255に対応するデジタルビデオシーケンスにおける画像フレームの指示を提供するフレーム識別メタデータを保管することを含んでいる。フレーム識別メタデータは、保管されたデジタルビデオフレーム235に関連して保管され得る。例えば、ビデオサマリー表現265は、保管されたデジタルビデオファイル235の中にビデオサマリーメタデータとして保管することができ、ビデオサマリーの中に含まれているスニペットに対応する一連の開始フレームおよび終了フレームを指示している。これにより、ビデオサマリー表現265を、フレーム識別メタデータを保管するために必要な少量のメモリの他に、あらゆる追加の物理的メモリを使用することなく保管することができる。ビデオサマリー255は、次に、従来のプレーヤーに対してトランスペアレントである間にビデオサマリーメタデータを理解する「スマート(“smart”)」ビデオプレーヤーを使用して再生され得る。

【0056】

別の実施例において、ビデオサマリー表現265は、デジタルビデオファイルサマリーである。この実施例において、ビデオサマリー255は、保管されたビデオファイル235とは分離したデジタルビデオファイルとして保管される。この場合において、ビデオサマリー表現265は、保管されたデジタルビデオファイル235と独立して視聴又は共有することができるデジタルビデオファイルサマリーである。望ましい実施例では、保管されたビデオサマリー表現265は、標準のビデオプレーヤーを使用して再生されるように

10

20

30

40

50

適合されたフォーマットでのデジタルビデオファイルサマリーである。

【 0 0 5 7 】

いくつかの実施例において、デジタルビデオファイルサマリーは、ビデオシーケンスのキャプチャーが完了した後に生成され得る。ビデオサマリーの所望のフレームは、保管されたビデオファイル 2 3 5 から抽出され得る。ビデオサマリーの所望のフレームが独立してエンコードされた画像のグループに対応して選択されている場合、デジタルビデオファイルサマリーは、圧縮されたビデオデータをデコードすることなく所望のフレームに対応する圧縮データを解凍することによって生成することができる。

【 0 0 5 8 】

いくつかの実施例において、デジタルビデオファイルサマリーは、デジタルビデオファイル 2 2 5 を形成するために使用されるものとは異なるビデオ処理パスを使用して、キャプチャーされたデジタルビデオシーケンス 2 0 5 の画像フレームから形成される。例えば、多くのビデオキャプチャーデバイスは、2つのビデオストリームを同時にキャプチャー及びエンコードすることができる。これらのデバイスにおいては、それぞれの画像フレームをキャプチャーするために、一つの画像センサー 1 4 が使用されている。そして、それぞれの画像フレームは、2つの異なる処理パスを使用して処理される。一つの処理パスは、デジタルビデオファイル 2 2 5 を生成することができる。第 2 の処理パスは、ビデオサマリー 2 5 5 を保管するために使用されるデジタルビデオファイルサマリーを生成し得る。第 2 の処理パスは、それぞれの画像フレームに係る空間解像度が低いバージョンを生成することができ、より低い解像度のビデオをエンコードすることができることにおいて、第 1 の処理パスとは異なっている。多くのビデオキャプチャーデバイスは、1 0 8 0 p 解像度ビデオと Q V G A 解像度ビデオとを同時にキャプチャーすることができる。

【 0 0 5 9 】

図 3 は、2つの処理パスを有するビデオキャプチャー処理を示している。キャプチャーされた画像フレーム 3 1 0 は、第 1 のビデオ処理パス 3 2 0 と第 2 のビデオ処理パス 3 3 0 の両方に向けられる。第 1 のビデオ処理パス 3 2 0 は、第 1 の空間解像度での第 1 のエンコードされた画像フレーム 3 5 0 を提供する第 1 のビデオエンコーダーを含んでいる。第 2 のビデオ処理パス 3 3 0 は、第 2 の空間解像度を有する変更された画像フレーム 3 7 0 を生成する任意の画像リサイザー (r e s i z e r) 3 6 0 を含んでいる。第 2 のビデオ処理パス 3 3 0 は、また、変更された画像フレーム 3 7 0 をエンコードする第 2 のビデオエンコーダー 3 8 0 を含み、第 2 のエンコードされた画像フレーム 3 9 0 を提供する。当業者であれば、第 1 のビデオ処理パス 3 2 0 及び第 2 のビデオ処理パス 3 3 0 は、また、任意で他のビデオ処理プロセス 3 0 0 を含み得ることが離解されよう。色とトーン (t o n e) のスケール処理、ノイズ削減、シャープ化 (s h a r p e n i n g)、および、画像安定化、といったものである。これらの他のビデオ処理ステップ 3 0 0 が最初に適用されるものとして示されているが、それらは、第 1 のビデオ処理パス 3 2 0 及び第 2 のビデオ処理パス 3 3 0 の中のいかなるポイントにおいても適用することができ、もしくは、さらに他のステップとの間に点在してよいことが認識されるべきである。当業者であれば、また、図 3 は、2つのビデオ処理パスが異なっている一つの可能なやり方を説明するものであり、本発明の範囲において他の相違も可能であることを理解するであろう。

【 0 0 6 0 】

デジタルビデオファイル 2 5 5 (図 2) を形成するために使用されたものとは異なるビデオ処理パスを使用して、デジタルビデオファイル 2 2 5 と同時に、デジタルビデオファイルサマリーを生成することができる。図 4 は、図 3 で示されたやり方のバリエーションを示しており、2つのビデオファイルを同時に生成するために使用され得る。図 3 のように、それぞれのキャプチャーされた画像フレーム 3 1 0 は、第 1 のエンコードされた画像フレーム 3 5 0 を生成するための第 1 のビデオエンコーダー 3 4 0 を含んでいる第 1 のビデオ処理パス 3 2 0 に向けられる。第 1 のエンコードされた画像フレーム 3 5 0 は、デジタルビデオファイル 2 2 5 の中に含まれる。

【 0 0 6 1 】

10

20

30

40

50

キャプチャーされた画像フレーム310は、また、第2のビデオ処理パス430にも向けられる。第2のビデオ処理パス430は、所定の人間を含む画像検査460の追加を除いて、図3における第2のビデオ処理パス330と類似している。、所定の人間を含む画像検査460は、画像が所定の人間を含んでいるかどうかを参照画像215から判断するためにキャプチャーされた画像フレーム310を分析する人間認識アルゴリズムを使用する。含んでいない場合、キャプチャーされた画像フレーム310は、画像廃棄ステップ470を使用して廃棄される。所定の人間がキャプチャーされた画像フレーム310の中に存在する場合、処理は、変更された画像フレーム370を決定するために画像リサイザー360に進み、そして、第2のエンコードされた画像フレーム390を提供するために第2のビデオエンコーダー380に進む。第2のエンコードされた画像フレーム390は、デジタルビデオファイルサマリーの中に含まれている。代替的な実施例では、画像リサイザー360は、第2のビデオ処理パスにおいてより早く適用され得る。所定の人間を含む画像検査460の以前にである。他の実施例において、画像リサイザー360は、第2の処理パス430の中には全く存在しなくてもよい。

10

【0062】

図4に示された第1のビデオ処理パス320及び第2のビデオ処理パス430は、2つの分離したビデオファイルを生成するために使用される。第1のビデオファイル(デジタルビデオファイル225)は、全てのキャプチャーされた画像フレーム310を含んでおり、第1の空間解像度でエンコードされる。第2のビデオファイル(デジタルビデオファイルサマリー)は、所定の興味のある人間を含んでいるキャプチャーされた画像フレーム310だけを含んでおり、第2の空間解像度でエンコードされる。第2のビデオファイルは、第1のビデオファイルを用いて同時に形成される。

20

【0063】

代替的に、デジタルビデオファイルサマリーは、デジタルビデオファイルを用いて部分的にだけ同時に形成され得る。図4において説明されるように、最初のデジタルビデオファイルサマリーが生成され得る。キャプチャーが完了した後で、デジタルビデオファイルサマリーは、保管されたデジタルビデオから抽出された追加のデータを用いて増加され得る。この場合、デジタルビデオファイルサマリーは、第2のビデオ処理パスによって元々エンコードされていない画像フレームを含むように拡張することができる。これにより、キャプチャー後のデジタルビデオファイルサマリー全体を形成することによるよりも早くデジタルビデオファイルサマリーを生成することができる。

30

【0064】

図5は、本発明の代替的な実施例を示している。この場合、所定の人間を含んでいる参照画像215は、ビデオシーケンスキャプチャーステップ200と共に同時にキャプチャーされないで、むしろ、独立した参照画像キャプチャーステップ510においてキャプチャーされる。参照画像キャプチャーステップは、ビデオシーケンスキャプチャーステップ200の前後のいずれで実行されてもよい。参照画像215は、静止キャプチャーモードにおいて獲得された一つの画像であってよく、または、存在しているキャプチャーされたデジタルビデオシーケンスからの画像フレームであってよい。いくつかの実施例において、デジタルビデオカメラは、デジタルビデオカメラの所有者にとっての一式の興味のある人間に対応する一式の人間画像を保管することができる(例えば、それぞれの家族メンバーに対する人間画像が保管され得る)。そして、好適なユーザーインターフェイスを使用して、一式の人間画像から参照画像215が選択される(例えば、既定の人間の名前のリスト)。参照画像215は、デジタルビデオカメラ10上にキャプチャーされ得るし、または、代替的に、分離した画像キャプチャーデバイス上にキャプチャーされてデジタルビデオカメラ10の中にインポート(import)され得る。

40

【0065】

いくつかの実施例においては、所定の人間の異なるビューを含んでいる複数の参照画像215が指定され得る。キャプチャーされた画像フレーム分析ステップ240は、画像フレームが所定の人間を含んでいるかどうかをより信頼性を持って判断できるようにするた

50

めに、複数の参照画像 215 を使用することができる。

【0066】

いくつかの実施例において、キャプチャーされた画像フレーム分析ステップ 240 は、ビデオシーケンスキャプチャーステップ 200 と同時に行われる。デジタルビデオファイル 225 の生成と同時にビデオサマリー 255 を形成するようにである。他の実施例において、ビデオサマリー 255 は、デジタルビデオファイル 225 がキャプチャーされ保管された後で実行されるポスト処理オペレーションを使用して形成され得る。いくつかの場合において、キャプチャーされた画像フレーム分析ステップ 240 は、デジタルカメラ上で実行され得る。代替的には、キャプチャーされたデジタルビデオファイル 235 が、ホストコンピューターといった、ある別のデバイス上にロードされた後で実行されてもよい。図 5 における残りのエレメントは、図 2 に係る実施例において示されたものと同等のものである。

10

【0067】

図 6 に関して本発明の代替的な実施例がこれから説明される。この場合においては、参照画像 215 に基づいてキャプチャーされた画像フレームを分析するよりむしろ、キャプチャーされた画像フレーム分析ステップ 640 は、興味の特徴を示している一式の参照データ 615 に基づいて画像フレーム 245 のサブセットを選択する。興味の特徴は、参照データ指定ステップ 610 において指定される。図 6 における残りのエレメントは、図 2 に係る実施例において示されたものと同等のものである。

【0068】

この参照データ 615 は、多くの形式であり得る。いくつかの実施例において、参照データ 615 は、興味の特徴のテキスト表現であってよい。例えば、参照データ 615 は、個人、対象物、場所、または、イベントの名前であってよい。いくつかの実施例において、参照データ 615 は、興味の特徴の数学的表現であってよい。例えば、参照データ 615 は、色ヒストグラム、固有値、または、あらゆる他の特徴ベクトルであってよい。当業者であれば、本発明の範囲内で数多くの画像情報が表現され得るやり方があることを認識するであろう。いくつかの実施例において、参照データ 615 は、非画像情報と関連してよい。例えば、参照データ 615 は、音声信号、グローバルポジショニングシステム (GPS) データ、自動焦点データ、自動露出データ、自動ホワイトバランスデータ、加速度データ、ジャイロスコープデータ、または、赤外線センサーデータであってよい。当業者であれば、情報のタイプには幅広い種類があり、画像フレーム 245 のサブセットを特定するためのキャプチャーされた画像フレーム分析ステップ 640 による使用のための参照データ 615 として提供され得ることを認識するであろう。

20

30

【0069】

デジタルビデオシーケンス 205 のキャプチャーの最中に、キャプチャーされた画像フレーム分析ステップ 640 は、特徴認識アルゴリズムを使用してデジタルビデオシーケンス 205 中の画像フレームを分析し、参照データ 615 によって指定されたように興味の特徴を含む画像フレームのサブセットを特定する。例えば、参照データ 615 によって指定された興味の特徴が、以前に指定された参照顔面画像に対する所定の名前の人間である場合、参照顔面画像に関する画像フレームを分析することによって、画像フレームが所定の名前の人間を含むかを判断するために、顔面認識アルゴリズムが使用され得る。別の実施例においては、参照データ 615 によって指定された興味の特徴がイベントラベル (例えば、「ゴルフスイング」) である場合、画像フレームが指定されたイベントに対応するかを判断するために特徴認識アルゴリズムが使用され得る。いくつかの場合、特徴認識アルゴリズムは、好適なイベントラベルを判断するために一連の画像フレームを分析する必要がある (例えば、どの画像フレームがゴルフスイングの動作特性を含んでいるかを検知するためである)。別の実施例では、参照データ 615 によって指定された興味の特徴が対象物ラベル (例えば、「くじら」) である場合、画像フレームが指定された対象物を含むかを判断するために特徴認識アルゴリズムが使用され得る。これらの全ての実施例は、キャプチャーの時に興味の特徴を特定することができる状況であり、画像フレームのサ

40

50

ブセット 2 4 5 をビデオサマリーに含めるかを判断する。

【 0 0 7 0 】

いくつかの実施例において、参照データ 6 1 5 によって指定された興味の特徴は、画像フレームに関する一般的な特性に対応してよい。例えば、参照データ 6 1 5 によって指定された興味の特徴は、関連するカメラ動作が少ない画像フレームであり得る。この場合、特徴認識アルゴリズムは、ジャイロスコープデータ、加速度データ、または、指定されたカメラ動作条件を満たしている画像フレームを特定するための画像ベースの動作見積りデータが使用され得る。同様に、参照データ 6 1 5 によって指定された興味の特徴は、カメラのズーム工程を追いかける画像フレームであってよい。この場合、ズームレンズデータ又は画像ベースのズーム見積りデータを分析するために特徴認識アルゴリズムが使用され、カメラズームがちょうど完了したところの画像フレームを特定し得る。別の実施例において、参照データ 6 1 5 によって指定された興味の特徴は、シーンにおいて大規模な対象物動作を伴う画像フレームであってよい。この場合、シーンにおける対象物動作の量を定量化するために特徴認識アルゴリズムが使用され、対象物動作条件を満たしている画像フレームを特定することができる。当業者であれば、これらは、本発明の範囲内にある多くの興味の特徴および特徴認識アルゴリズムの実施例に過ぎないことを理解するであろう。

10

【 0 0 7 1 】

コンピュータープログラム製品は、一つまたはそれ以上のストレージ媒体を含み得る。例えば、磁気ディスク（フロッピー（登録商標）ディスクといったもの）又は磁気テープといった磁気ストレージ媒体；光ディスク、光テープ、または、機械で読取り可能なバーコードといった光ストレージ媒体；ランダムアクセスメモリー（RAM）、または、読み出し専用メモリー（ROM）といった半導体電子ストレージデバイス；もしくは、本発明に従った方法を実行するように一つまたはそれ以上のコンピューターをコントロールするためのインストラクションを有するコンピュータープログラムを保管するために使用される、あらゆる他の物理的デバイス又は媒体、である。

20

【符号の説明】

【 0 0 7 2 】

- 2 フラッシュ
- 4 レンズ
- 6 調整可能な絞りと調整可能なシャッター
- 8 ズームとフォーカスマータードライバー
- 10 デジタルカメラ
- 12 タイミングジェネレーター
- 14 画像センサー
- 16 ASPおよびA/Dコンバーター
- 18 バッファメモリー
- 20 プロセッサ
- 22 音声コーデック
- 24 マイクロフォン
- 25 圧力センサー
- 26 スピーカー
- 27 加速度計
- 28 ファームウェアメモリー
- 30 ストレージメモリー
- 32 画像ディスプレイ
- 34 ユーザーコントロール
- 36 ディスプレイメモリー
- 38 有線インターフェイス
- 40 コンピューター
- 42 傾きセンサー

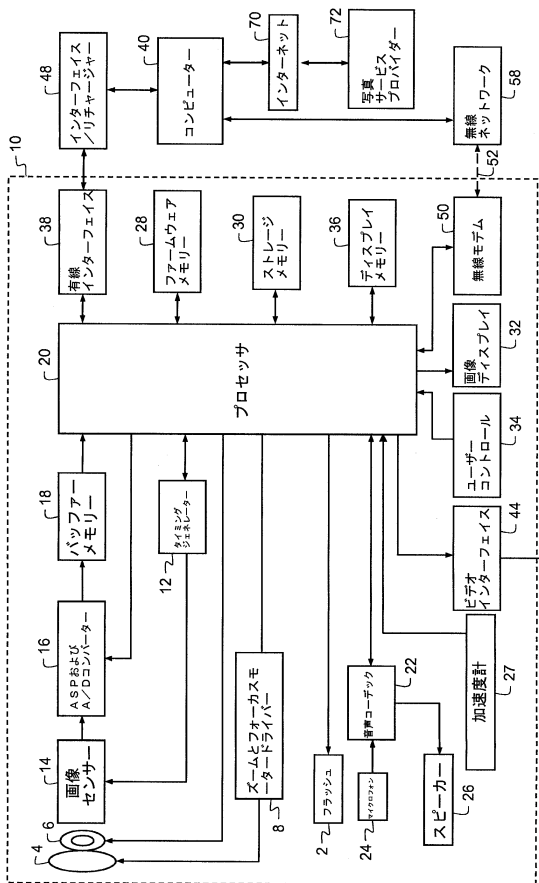
30

40

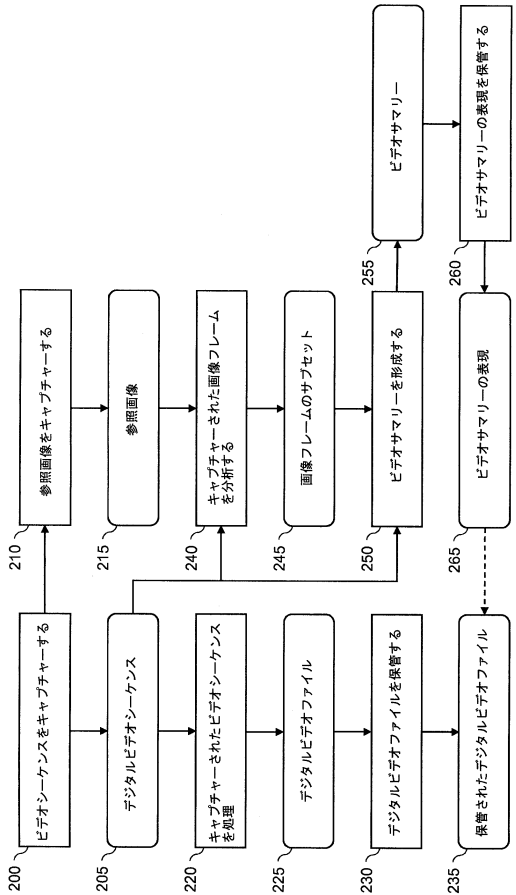
50

4 4	ビデオインターフェイス	
4 6	ビデオディスプレイ	
4 8	インターフェイス/リチャージャー	
5 0	無線モデム	
5 2	ラジオ周波数帯	
5 8	無線ネットワーク	
7 0	インターネット	
7 2	写真サービスプロバイダー	
2 0 0	ビデオシーケンスをキャプチャーする	
2 0 5	デジタルビデオシーケンス	10
2 1 0	参照画像をキャプチャーする	
2 1 5	参照画像	
2 2 0	キャプチャーされたビデオシーケンスを処理	
2 2 5	デジタルビデオファイル	
2 3 0	デジタルビデオファイルを保管する	
2 3 5	保管されたデジタルビデオファイル	
2 4 0	キャプチャーされた画像フレームを分析する	
2 4 5	画像フレームのサブセット	
2 5 0	ビデオサマリーを形成する	
2 5 5	ビデオサマリー	20
2 6 0	ビデオサマリーの表現を保管する	
2 6 5	ビデオサマリーの表現	
3 0 0	他のビデオ処理ステップ	
3 1 0	キャプチャーされた画像フレーム	
3 2 0	第1のビデオ処理パス	
3 3 0	第2のビデオ処理パス	
3 4 0	第1のビデオエンコーダー	
3 5 0	第1のエンコードされた画像フレーム	
3 6 0	画像リサイザー	
3 7 0	変更された画像フレーム	30
3 8 0	第2のビデオエンコーダー	
3 9 0	第2のエンコードされた画像フレーム	
4 3 0	第2のビデオ処理パス	
4 6 0	所定の人間を含む画像検査	
4 7 0	画像廃棄ステップ	
5 1 0	参照画像キャプチャーステップ	
6 1 0	参照データ指定ステップ	
6 1 5	参照データ	
6 4 0	キャプチャーされた画像フレーム分析ステップ	

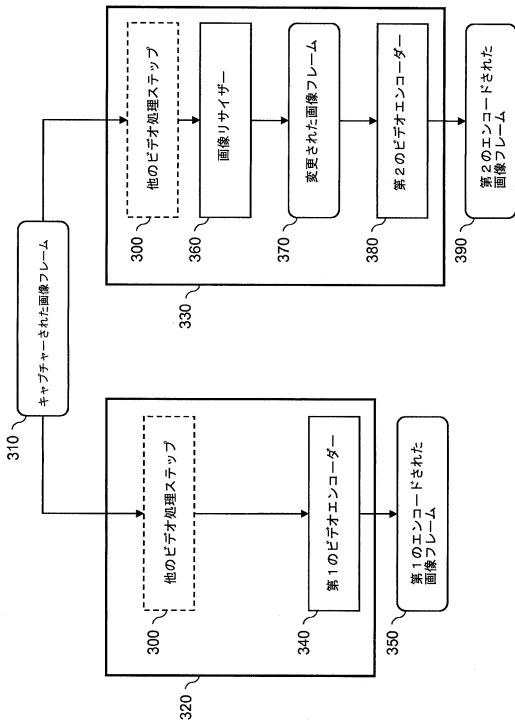
【図 1】



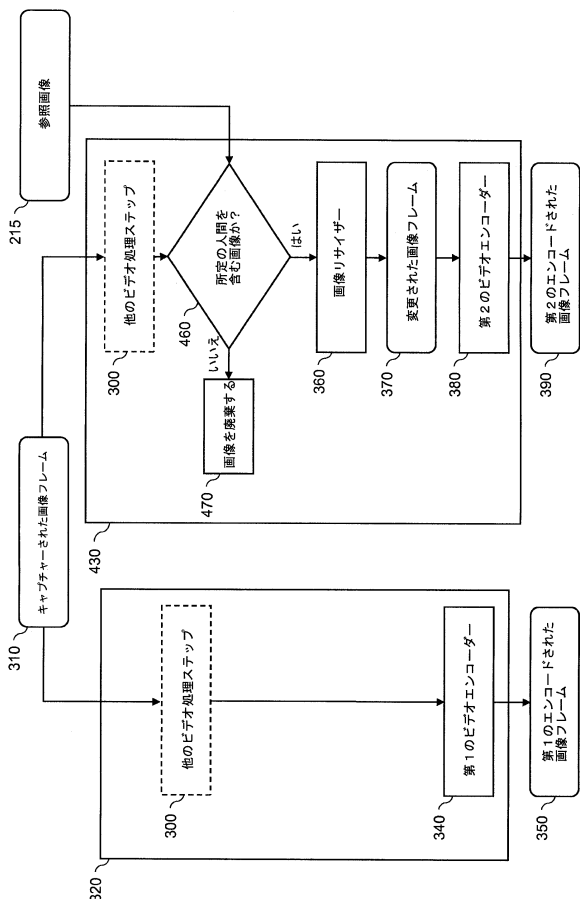
【図 2】



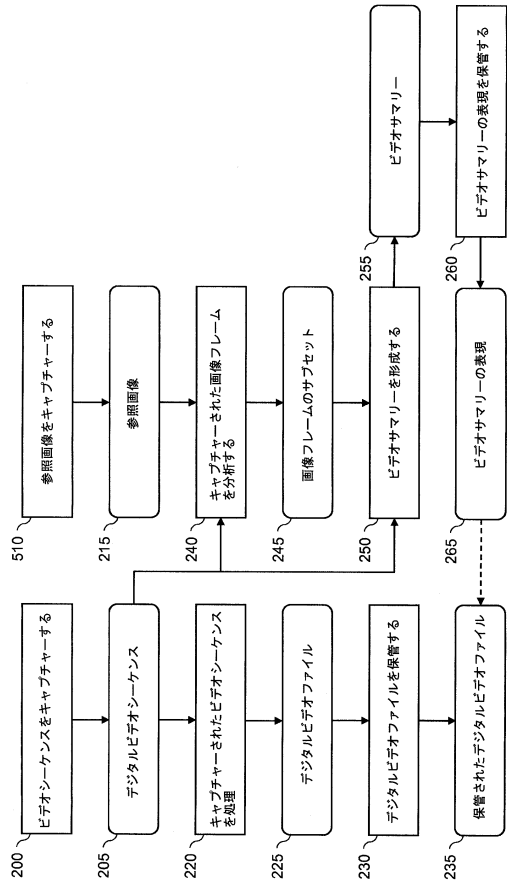
【図 3】



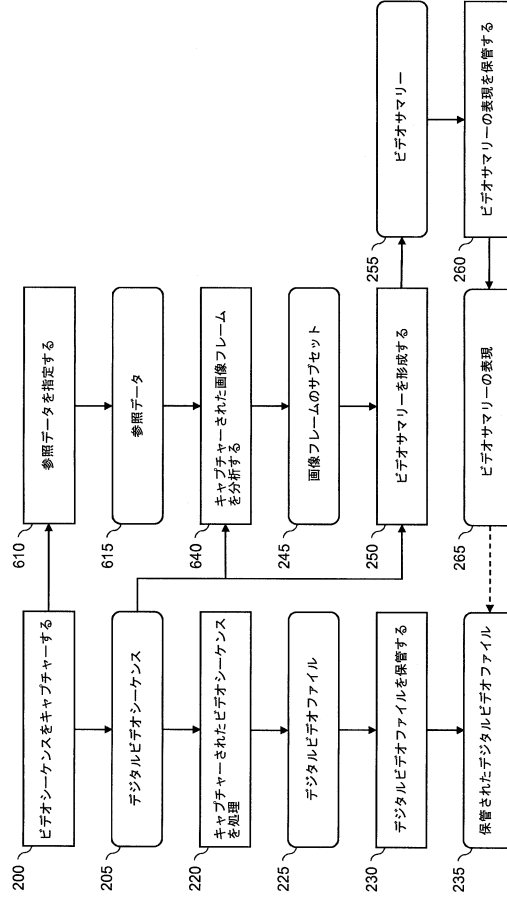
【図 4】



【 図 5 】



【 図 6 】



フロントページの続き

- (72)発明者 カーン, キース, ストール
 アメリカ合衆国 1 4 6 5 0 - 2 2 0 1 ニューヨーク州 ロチェスター ステイト・ストリート
 ・ 3 4 3
- (72)発明者 ピルマン, ブルース, ハロルド
 アメリカ合衆国 1 4 6 5 0 - 2 2 0 1 ニューヨーク州 ロチェスター ステイト・ストリート
 ・ 3 4 3
- (72)発明者 ディーヴァー, アーロン, トーマス
 アメリカ合衆国 1 4 6 5 0 - 2 2 0 1 ニューヨーク州 ロチェスター ステイト・ストリート
 ・ 3 4 3
- (72)発明者 マッコイ, ジョーン, アール.
 アメリカ合衆国 1 4 6 5 0 - 2 2 0 1 ニューヨーク州 ロチェスター ステイト・ストリート
 ・ 3 4 3
- (72)発明者 ラザヴィー, フランク
 アメリカ合衆国 1 4 6 5 0 - 2 2 0 1 ニューヨーク州 ロチェスター ステイト・ストリート
 ・ 3 4 3
- (72)発明者 グレツィンガー, ロバート
 アメリカ合衆国 1 4 6 5 0 - 2 2 0 1 ニューヨーク州 ロチェスター ステイト・ストリート
 ・ 3 4 3

審査官 松元 伸次

- (56)参考文献 特開2010-219607(JP, A)
 特開2009-077026(JP, A)
 特開2010-158046(JP, A)
 特開2006-163746(JP, A)
 特開2009-010839(JP, A)

(58)調査した分野(Int.Cl., DB名)

G 1 1 B 2 7 / 0 0 - 2 7 / 0 6
 H 0 4 N 1 / 4 0 - 1 / 4 0 9
 1 / 4 6
 1 / 6 0
 5 / 7 6
 5 / 7 6 5
 5 / 8 0 - 5 / 9 1
 5 / 9 1 5
 5 / 9 2
 5 / 9 2 2
 5 / 9 2 8 - 5 / 9 3
 5 / 9 3 7 - 5 / 9 4
 5 / 9 5 - 5 / 9 5 6