

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2019-192986

(P2019-192986A)

(43) 公開日 令和1年10月31日(2019.10.31)

(51) Int.Cl.		F I		テーマコード (参考)
HO4N 5/93 (2006.01)		HO4N 5/93		5C053
GO6F 16/00 (2019.01)		GO6F 17/30	320A	
GO6F 16/70 (2019.01)		GO6F 17/30	170D	
HO4N 5/783 (2006.01)		GO6F 17/30	360Z	
		HO4N 5/783		

審査請求 未請求 請求項の数 9 O L (全 17 頁)

(21) 出願番号 特願2018-80849 (P2018-80849)
 (22) 出願日 平成30年4月19日 (2018.4.19)

(71) 出願人 899000068
 学校法人早稲田大学
 東京都新宿区戸塚町1丁目104番地
 (71) 出願人 504137912
 国立大学法人 東京大学
 東京都文京区本郷七丁目3番1号
 (74) 代理人 110002675
 特許業務法人ドライト国際特許事務所
 (72) 発明者 森島 繁生
 東京都新宿区戸塚町1丁目104番地 学
 校法人早稲田大学内
 (72) 発明者 粥川 青汰
 東京都新宿区戸塚町1丁目104番地 学
 校法人早稲田大学内

最終頁に続く

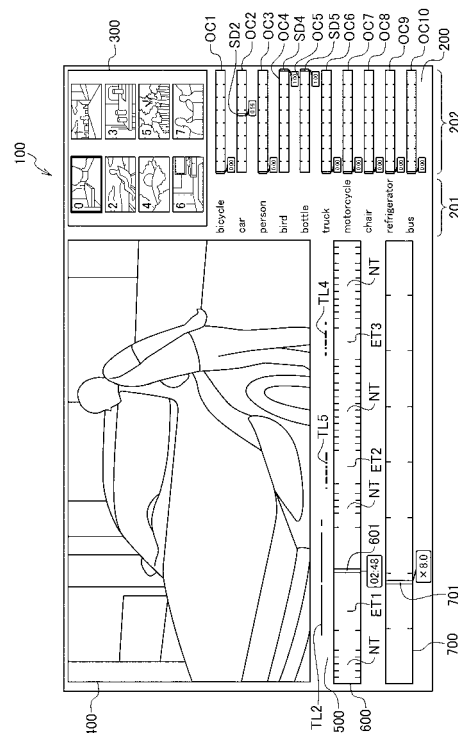
(54) 【発明の名称】 表示制御装置、表示制御方法及び表示制御プログラム

(57) 【要約】

【課題】 動画の内容を反映したインタフェースにより高速閲覧を可能とする表示制御装置、表示制御方法及び表示制御プログラムを提案する。

【解決手段】 表示制御装置10は、複数フレームの映像から構成される動画からフレーム毎に映像に含まれるオブジェクトを検出し、検出されたオブジェクトのうち、動画内での出現態様（各オブジェクトが出現する頻度及び/又は出現するタイミング）に基づく条件を満たすオブジェクトを選択し、選択されたオブジェクトのそれぞれについてユーザによる重要度の指定を受け付け可能なキューOC1~OC10をインタフェース100に提示する。

【選択図】 図12



【特許請求の範囲】**【請求項 1】**

複数フレームの映像から構成される動画から、前記動画に含まれるオブジェクトを検出する検出部と、

前記検出部により検出されたオブジェクトの中から、前記動画内での出現態様に基づく条件を満たすオブジェクトを選択する選択部と、

前記選択部により選択されたオブジェクトのそれぞれについてユーザの指定を受け付け可能な指定部をインタフェースに提示する表示制御部と、
を備える表示制御装置。

【請求項 2】

前記検出部は、フレーム毎に前記動画に含まれるオブジェクトを検出し、

前記選択部は、前記検出部により検出されたオブジェクトの中から、前記動画内での出現頻度及び出現タイミングの一方又は双方に基づいてオブジェクトを選択する、請求項 1 に記載の表示制御装置。

【請求項 3】

前記選択部は、前記検出部により検出されたオブジェクトのうち、前記動画内での出現頻度が相対的に低いオブジェクトを選択対象から除外する、請求項 2 に記載の表示制御装置。

【請求項 4】

前記選択部は、前記検出部により検出されたオブジェクトのうち、前記動画内での出現頻度が相対的に高いオブジェクトを選択対象から除外する、請求項 2 に記載の表示制御装置。

【請求項 5】

前記選択部は、前記検出部により検出されたオブジェクトの中で同時に出現する 2 以上のオブジェクトがある場合、当該 2 以上のオブジェクトのうちの一つを選択する、請求項 2 に記載の表示制御装置。

【請求項 6】

前記インタフェースに提示された前記指定部のうち少なくとも一つの指定部が指定された場合、前記表示制御部は、前記複数フレームのうち、当該少なくとも一つの指定部に対応するオブジェクトが出現する第 1 フレームの再生速度が、前記第 1 フレーム以外の第 2 フレームの再生速度よりも遅くなるように前記動画を表示させる、請求項 1 から 5 のいずれか 1 項に記載の表示制御装置。

【請求項 7】

前記表示制御部は、

前記選択部により選択されたオブジェクト毎に前記インタフェースに提示する前記指定部の色が異なるように表示し、

前記インタフェースに提示された前記指定部のうち少なくとも一つの指定部が指定された場合、当該少なくとも一つの指定部に対応するオブジェクトが出現する時間軸上の区間を当該少なくとも一つの指定部の色で前記インタフェースに提示する、請求項 1 から 6 のいずれか 1 項に記載の表示制御装置。

【請求項 8】

複数フレームの映像から構成される動画から、前記動画に含まれるオブジェクトを検出する検出ステップと、

前記検出ステップで検出されたオブジェクトの中から、前記動画内での出現態様に基づく条件を満たすオブジェクトを選択する選択ステップと、

前記選択ステップで選択されたオブジェクトのそれぞれについてユーザの指定を受け付け可能な指定部をインタフェースに提示する提示ステップと、
を含む表示制御方法。

【請求項 9】

コンピュータを、

10

20

30

40

50

複数フレームの映像から構成される動画から、前記動画に含まれるオブジェクトを検出する検出部と、

前記検出部により検出されたオブジェクトの中から、前記動画内での出現態様に基づく条件を満たすオブジェクトを選択する選択部と、

前記選択部により選択されたオブジェクトのそれぞれについてユーザの指定を受け付け可能な指定部をインタフェースに提示する表示制御部と、

として機能させる表示制御プログラム。

【発明の詳細な説明】

【技術分野】

10

【0001】

本発明は、表示制御装置、表示制御方法及び表示制御プログラムに関する。

【背景技術】

【0002】

近年、ウェアラブルカメラの小型化及び普及に伴い、頭部に装着されたカメラによって撮影した動画（以下、「一人称視点映像」と呼ぶ。）が撮影される機会が増加している。一人称視点映像を閲覧することで、撮影者がどこへ行き、何をしていたかなどの詳細な記録を、撮影者の視線を通して共有可能となる。両手が空いた状態で撮影可能な一人称視点映像は撮影時の負担が非常に小さいため、日常生活、レジャー、スポーツ、個人技能の解析など、様々な対象の記録を残す手段として今後普及して行くことが予想される。さらに、カメラ付きのスマートフォンの普及に伴い、誰でも気軽に写真が撮影可能となったことで、ソーシャル・ネットワーキング・サービス（SNS）において大量の写真が投稿、共有される時代が到来したように、今後、次なる撮影手段としてウェアラブルカメラが普及することで、一人称視点映像を共有し、閲覧する機会が増えていくことが予想される。

20

【0003】

しかしながら、一人称視点映像では常時撮影が基本であるため、撮影者の視覚体験をもれなく記録可能である反面、長時間かつ冗長なシーンを含む場合が多くなってしまい、映像を閲覧するのに時間を要するという問題点がある。この問題点を解決する手法として、シーンごとに再生速度を変化させる高速閲覧手法が研究されている。

【0004】

30

例えば、特許文献1及び非特許文献1では、一人称視点映像を閲覧する手がかり（キュー）として、Movement（移動）、Stop（静止）、Hand（手の動作）、Person（人物との対話）という撮影者の基本的な行動に対応した4つのキューを搭載したインタフェースが提案されている。ユーザがこれらのキューの重要度を指定することで、一人称視点映像の中で注目部分と非注目部分が設定され、注目部分は通常速度で再生され、非注目部分は高速で再生される。これにより、一人称視点映像からユーザの関心の高いシーンを効率的且つ高速に発見し、閲覧することが可能となる。

【先行技術文献】

【特許文献】

【0005】

40

【特許文献1】特願2016-241446

【非特許文献】

【0006】

【非特許文献1】K. Higuchi, R. Yonetani, and Y. Sato, "EgoScanning: Quickly Scanning First-Person Videos with Egocentric Elastic Timelines" In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 2017, p. 6536-6546

【発明の概要】

【発明が解決しようとする課題】

【0007】

50

しかしながら、特許文献 1 及び非特許文献 1 で提案されたインタフェースでは、ユーザが指定可能なキューが 4 つのキューに固定されており、動画の内容が考慮されていないため、システムが有効に働く動画が限定されるという問題があった。

【0008】

本発明は以上の点を考慮してなされたもので、動画の内容を反映したインタフェースにより高速閲覧を可能とする表示制御装置、表示制御方法及び表示制御プログラムを提案することを目的とする。

【課題を解決するための手段】

【0009】

本発明の表示制御装置は、複数フレームの映像から構成される動画から、前記動画に含まれるオブジェクトを検出する検出部と、前記検出部により検出されたオブジェクトの中から、前記動画内での出現態様に基づく条件を満たすオブジェクトを選択する選択部と、前記選択部により選択されたオブジェクトのそれぞれについてユーザの指定を受け付け可能な指定部をインタフェースに提示する表示制御部と、を備えるものである。

10

【0010】

本発明の表示制御方法は、複数フレームの映像から構成される動画から、前記動画に含まれるオブジェクトを検出する検出ステップと、前記検出ステップで検出されたオブジェクトの中から、前記動画内での出現態様に基づく条件を満たすオブジェクトを選択する選択ステップと、前記選択ステップで選択されたオブジェクトのそれぞれについてユーザの指定を受け付け可能な指定部をインタフェースに提示する提示ステップと、を含むものである。

20

【0011】

本発明の表示制御プログラムは、コンピュータを、複数フレームの映像から構成される動画から、前記動画に含まれるオブジェクトを検出する検出部と、前記検出部により検出されたオブジェクトの中から、前記動画内での出現態様に基づく条件を満たすオブジェクトを選択する選択部と、前記選択部により選択されたオブジェクトのそれぞれについてユーザの指定を受け付け可能な指定部をインタフェースに提示する表示制御部と、として機能させるものである。

【発明の効果】

【0012】

本発明によれば、動画に含まれる各オブジェクトについてユーザの指定を受け付け可能な指定部を動的に生成してインタフェースに提示することにより、より効率的に動画から特定のシーンを発見することができる。

30

【図面の簡単な説明】

【0013】

【図 1】本発明の実施形態の表示制御装置のハードウェア構成を示すブロック図である。

【図 2】本実施形態の表示制御装置により実行されるキュー生成処理のフローチャートである。

【図 3】動画からのオブジェクト検出を説明する図である。

【図 4】動画から検出されたオブジェクトのキューの候補を示す図である。

40

【図 5】有効なキューの組み合わせを選択するキュー選択アルゴリズムの例を説明する図である。

【図 6 A】キュー選択アルゴリズムにおいて、動画の一部のフレームのみに出現するオブジェクトがある場合を説明する図である。

【図 6 B】キュー選択アルゴリズムにおいて、動画の大部分のフレームで出現するオブジェクトがある場合を説明する図である。

【図 6 C】キュー選択アルゴリズムにおいて、同時に異なるオブジェクトが出現する場合を説明する図である。

【図 6 D】キュー選択アルゴリズムにおいて、閲覧に有効なオブジェクトのキューが選択される場合を説明する図である。

50

【図 7】キュー選択アルゴリズムにおいて、キューの候補の中から有効なキューが選択されることを説明する図である。

【図 8】貪欲法を用いたキュー選択アルゴリズムの手順を説明する図である。

【図 9】キュー選択アルゴリズムの結果に基づいて生成されるオブジェクトキューを示す図である。

【図 10】本実施形態の表示制御装置により表示されるインタフェースであって、オブジェクトキューが操作されていない状態の表示例を示す図である。

【図 11】本実施形態の表示制御装置により実行される閲覧設定処理のフローチャートである。

【図 12】本実施形態の表示制御装置により表示されるインタフェースであって、オブジェクトキューが操作されたときの表示例を示す図である。

【図 13 A】本実施形態のインタフェースの評価実験のために用いる複数の映像の内容を示す表である。

【図 13 B】評価実験の設計を示す表である。

【図 14】本実施形態のインタフェースを用いた場合において、映像毎に検出回数の多いオブジェクトとキュー選択アルゴリズムで選択されたキュー（オブジェクト）とを示す表である。

【図 15】インタフェースの高速閲覧性能を図る尺度としての平均閲覧速度を説明する図である。

【図 16】評価実験で用いた映像毎に、本実施形態のインタフェース及び従来のインタフェースの平均閲覧速度及び 95% 信頼区間を示す表である。

【図 17】本実施形態のインタフェース及び従来のインタフェースに対する主観評価結果を示す図である。

【発明を実施するための形態】

【0014】

以下、図面を参照して本発明の実施形態について説明する。各図において、同一又は同様の要素には同一の符号を付し、重複する説明は省略する。

【0015】

図 1 は、本発明の実施形態に係る表示制御装置 10 のハードウェア構成を示すブロック図である。表示制御装置 10 は、CPU (Central Processing Unit) 11 と、RAM (Random Access Memory) 12 と、ROM (Read Only Memory) 13 と、通信部 14 と、入力部 15 と、表示部 16 とを有し、各部はバスを介して相互に接続されている。

【0016】

CPU 11 は、ROM 13 に格納された表示制御プログラム等の各種のアプリケーションプログラムを実行する。また、CPU 11 は、通信部 14 又は入力部 15 からの入力データを演算し、演算結果を RAM 12 又は ROM 13 に格納させたり、表示部 16 に表示させたりする。具体的に、CPU 11 は、ROM 13 に格納された表示制御プログラムにしたがって、動画からフレーム毎にオブジェクトを検出する検出部と、検出されたオブジェクトの中から、動画中の各オブジェクトの出現態様（各オブジェクトが出現する頻度及び出現するタイミング）に基づく条件を満たすオブジェクトを選択する選択部と、選択されたオブジェクトのそれぞれについてユーザが重要度を指定可能なキュー（指定部）を表示部 16 上のユーザインタフェース（UI）に提示する表示制御部としての機能を有する。

【0017】

RAM 12 は、半導体素子で構成される揮発性メモリであり、CPU 11 によって実行される各種処理で生じたデータ等を記憶する。

【0018】

ROM 13 は、半導体記憶素子で構成される不揮発性メモリであり、CPU 11 によって実行される各種のアプリケーションプログラムやデータを格納している。なお、表示制御装置 10 で実行可能な各種のアプリケーションプログラムは、表示制御装置 10 が読み

10

20

30

40

50

取り可能なメモリカード等の外部記録媒体に格納されていてもよいし、通信ネットワークを介して提供されてもよい。

【0019】

通信部14は、表示制御装置10をLAN(Local Area Network)又はWAN(Wide Area Network)等の通信ネットワークに接続する。

【0020】

入力部15は、キーボード、マウス、タッチパネル等の入力デバイスからなり、ユーザの操作によってデータの入力を受け付ける。

【0021】

表示部16は、LCD(Liquid Crystal Display)等の表示デバイスからなり、CPU11による演算結果を表示する。具体的に、表示部16は、CPU11により実行された表示制御プログラムに基づいて、動画表示のためのUI(図10及び図12参照。)を表示する。

10

【0022】

なお、表示制御装置10は、一般のパーソナルコンピュータ又はスマートフォンのCPUによって表示制御プログラムを実行することで実現されるようにしてもよい。

【0023】

次に、表示制御装置10により実行される処理について説明する。

図2は、表示制御装置10により実行されるキュー生成処理を示すフローチャートである。

20

【0024】

まず、表示制御装置10は、図3に示すように、入力された動画(一人称視点映像)からフレーム毎に動画に含まれるオブジェクト(物体、人物等)を検出する(ステップS1)。動画からオブジェクトを検出するのは、検出されたオブジェクトの一部を重要度指定のためのキューとしてUIに提示することにより、ユーザは提示されたキューから映像の内容(撮影された場所や撮影者の行動等)を推定し、且つ動画中の特定のシーンに容易にアクセスできるようにするためである。

【0025】

ステップS1では、例えば、公知のオブジェクト検出方法であるYOLOv2(J. Redmon and A. Farhadi. YOLO9000: better, faster, stronger. In In Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR '17), pp. 7263-7271, 2017.)を用いる。ここで、公知のCOCO dataset(T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. Microsoft COCO: Common Objects in Context. In Proceedings of the European Conference on Computer Vision (ECCV), pp. 740-755, 2014.)を用いて合計80種類のオブジェクトが検出可能なネットワークを利用する。なお、オブジェクトの検出方法及び検出可能なオブジェクトの数については、これらの公知技術に限定されるものではない。

30

【0026】

ステップS1において動画からオブジェクトが検出されると、表示制御装置10は、図4に示すように、検出された全てのオブジェクトについて、重要度を指定するためのキューを暫定的に設定する。ここで、ステップS1で検出された全てのオブジェクトに対応するキューをUIの限られたスペースに表示することは困難である。また、動画中に数フレームしか現れないオブジェクトや、逆に常時現れ続けるオブジェクトなど、適応的な高速閲覧に必ずしも適さないオブジェクトがキューとして利用される可能性がある。そこで、表示制御装置10は、キュー選択アルゴリズムを用いて、ステップS1で検出された全てのオブジェクトに対応するキューの候補の中から、動画の高速閲覧に有効なオブジェクトのキューを選択する(ステップS2)。

40

【0027】

以下、図5~図8を参照して、動画の高速閲覧に有効なオブジェクトのキューを選択するキュー選択アルゴリズムについて詳細に説明する。

50

【 0 0 2 8 】

ここでは、図 5 に示すように、すでに 2 つのオブジェクトのキュー (c_1 、 c_2) が選択済みであるとき、新たな 4 つのオブジェクトのキューの候補 (c_3 、 c_4 、 c_5 、 c_6) の中から有効なキューを 1 つ選択する例を説明する。図 5 に示す各キューの斜線部分は、時間軸上でオブジェクトが出現する箇所 (キューが強調する箇所) を示している。これら 4 つのキューの候補 (c_3 、 c_4 、 c_5 、 c_6) は、図 5 に示すように、動画全体における出現態様 (オブジェクトが出現する頻度及び出現するタイミング) が異なっている。最初に選択される 2 つのキュー c_1 及び c_2 の選択方法については後述する。

【 0 0 2 9 】

(i) 動画のごく一部のみ出現するオブジェクトのキュー (Irrelevant cue)

10

図 6 A に示すように、キュー c_3 は、動画内のごくわずかなシーンのみ強調するキューであり、動画の要旨に無関係なノイズである場合が多い。また、選択済みの 2 つのキュー (c_1 、 c_2) にキュー c_3 を追加した場合、これらのキューを任意に組み合わせても (c_1+c_3 、 c_2+c_3 、 $c_1+c_2+c_3$ など) 強調できないシーンが動画中に多く存在してしまう。よって、キュー c_3 は動画の高速閲覧に有効なキューとして選択されない。

【 0 0 3 0 】

(ii) 動画の大部分に出現するオブジェクトのキュー (Dominant cue)

反対に、図 6 B に示すキュー c_4 のように、動画の大部分に出現するオブジェクトのキューは、動画の内容を反映したキューである反面、ユーザがキュー c_4 の重要度を高く指定すると、動画の大部分が強調された冗長な動画が再生されてしまい、動画からシーンを絞り込む効果がない。よって、キュー c_4 も動画の高速閲覧に有効なキューとして選択されない。

20

【 0 0 3 1 】

(iii) 他のオブジェクトと同時に出現するオブジェクトのキュー (Overlapping cue)

動画内で同時に出現するオブジェクトとしては、例えば、お皿とフォーク、ディスプレイとキーボードなどがある。図 6 C に示すキュー c_5 のように、動画の一部区間を適度に強調するがその強調箇所が既に選択済みのキュー (図 6 C ではキュー c_2) の強調箇所と被っている場合 (例えば、キュー c_2 がお皿、キュー c_5 がフォークに対応する場合)、ユーザがキュー c_2 の重要度を指定した場合でもキュー c_5 の重要度を指定した場合でも、同じ箇所が強調された動画が再生されることとなる。よって、キュー c_5 も動画の高速閲覧に有効なキューとして選択されない。

30

【 0 0 3 2 】

(iv) 最適なキュー (Acceptable cue)

図 6 D に示すキュー c_6 は、動画の一部区間を強調でき、且つ上記 (i) ~ (iii) のいずれにも該当しないキューである。キュー c_6 が選択済みの 2 つのキュー c_1 、 c_2 に追加され、結果として得られた 3 つのキューを任意に組合せると、動画の一部を強調することができる様々なパターン (c_1+c_2 、 c_1+c_6 、 c_2+c_6 、 $c_1+c_2+c_6$) を得ることができる。

【 0 0 3 3 】

そのため、図 7 に示すように、4 つのオブジェクトのキューの候補 (c_3 、 c_4 、 c_5 、 c_6) の中から有効なキューとして c_6 が選択され、結果として 3 つの有効なキュー c_1 、 c_2 、 c_6 が選択されることとなる。

40

【 0 0 3 4 】

このように、動画から検出された全てのオブジェクトについて暫定的に設定されたキューの中から上記のキュー選択アルゴリズムに従って最適なキューを一つずつ選択して追加する作業を、予め設定されたキューの個数まで繰り返すことで、最終的にユーザに提示する最適なキューの組み合わせが決定される。

【 0 0 3 5 】

キュー選択アルゴリズムに従って最適なキューを選択するために、キューの組み合わせ $C_N = \{c_1, \dots, c_N\}$ に対し式 (1) に示す有効度評価関数 $F(C_N)$ を用いる。

【数 1】

$$F(C_N) = A(C_N) - \max_{i \in N} (B(c_i)) \cdot \cdot \cdot (1)$$

【0036】

式(1)において $A(C_N)$ は N 個のキューの組み合わせで強調するフレームの総数、 $B(c_i)$ は1つのキュー c_i が強調するフレームの数を示す。キュー選択アルゴリズムでは有効度評価関数 $F(C_N)$ が最大となるようなキューの組み合わせ C_N を選択する。

【0037】

例えば、図5に示すように、既に2つのキューが選択済みで、新たに3番目のキューを選択する場合($N=3$ の場合)、 $c_1+c_2+c_3$ (図6A)や $c_1+c_2+c_5$ (図6C)は強調するフレームの総数 $A(C_3)$ が相対的に少ない一方で、 $c_1+c_2+c_4$ (図6B)や $c_1+c_2+c_6$ (図6D)は強調するフレームの総数 $A(C_3)$ が相対的に多い。また、キュー c_4 が強調するフレームの数 $B(c_4)$ は相対的に多い一方で、キュー c_6 が強調するフレームの数 $B(c_6)$ は相対的に少ない。すなわち、 $A(C_N)$ を導入することで、動画内でわずかしき出現しないキュー c_3 や選択済みのキューと強調箇所が被るようなキュー c_5 は排除される。また、 $\max(B(c_i))$ を導入することで、キュー c_4 のように動画全体を通して極端に出現頻度の高いオブジェクトのキュー(動画の大部分を強調してしまうキュー)を排除することができる。よって、既に2つのキュー(c_1, c_2)が選択済みの場合、4つのキューの候補(c_1, c_2, c_3, c_4)のうち有効度評価関数 $F(C_3)$ が最大となるキューは c_6 となる。

【0038】

有効なキューの選択のアルゴリズムでは、全探索を用いて有効なキューの組み合わせを求めるのは計算量が膨大となりNP困難となることから、本実施形態では貪欲法を用いて有効なキューの組み合わせを求める。ただし、最初に選択される2つのキューについては、図8に示すように、有効度評価関数 $F(C_2)$ が最大となるようなキューの組み合わせ $C_2 = \{c_1, c_2\}$ を全探索によって求める。 $C_3 = \{c_1, c_2, c_3\}$ を求める場合は、選択済みの $C_2 = \{c_1, c_2\}$ を保持し、 $F(C_3)$ が最大となるようなキュー c_3 を求める。このように貪欲法を用いて最適なキューの組み合わせを求める処理を予め設定されたキューの個数 m まで繰り返すことで、最終的にユーザに提示する最適なキューの組み合わせ $C_m = \{c_1, \dots, c_m\}$ が決定される。

【0039】

なお、本実施形態では、有効度評価関数として式(1)を採用したが、上記(iv)の最適なキューが選択可能であれば、他の関数を採用してもよい。

【0040】

キュー選択アルゴリズムによって最適なキューが選択されると、表示制御装置10は、選択された最適なキューを、表示部16に表示されたインタフェース100上に提示させる(ステップS3)。図9に、キュー選択アルゴリズムによって選択された最適なキュー(オブジェクトキュー)の一例を示す。図9のオブジェクトキューは、図10に示すように、インタフェース100のオブジェクトキュー領域200に提示される。図9及び図10ではキューが10個($m=10$)提示された例を示している。なお、本実施形態では、最適なキューの個数 m が予め設定されるものとしているが、ユーザが最適なキューの個数を任意に指定可能としてもよい。

【0041】

オブジェクトキュー領域200は、図9に示すように、オブジェクト名表示領域201とキュー表示領域202とを有し、オブジェクト名表示領域201上の各オブジェクト名に対応して、重要度を指定するためのキューOC1~OC10が表示されている。キューOC1~OC10のそれぞれにはスライダSD1~SD10が設けられており、スライダSD1~SD10を操作することにより、重要度が0.00~1.00の間で指定可能となっている。また、キューOC1~OC10は互いに異なる色で表示されており(例えば、キューOC1は赤色、キューOC2は緑色など)、各オブジェクトが色によって区別可能となっている。

10

20

30

40

50

【 0 0 4 2 】

図 1 0 は、表示部 1 6 に表示されるインタフェース 1 0 0 であって、オブジェクトキュー領域 2 0 0 のキュー OC 1 ~ OC 1 0 のスライダ SD 1 ~ SD 1 0 が全て 0.00 の位置にある場合（すなわち、全てのキューの重要度が 0.00 である場合）を示す。図 1 0 に示すように、インタフェース 1 0 0 は、オブジェクトキュー領域 2 0 0 と、動画選択領域 3 0 0 と、再生画面領域 4 0 0 と、オブジェクトタイムライン 5 0 0 と、伸縮タイムライン 6 0 0 と、再生速度設定領域 7 0 0 とから構成される。

【 0 0 4 3 】

動画選択領域 3 0 0 は、再生画面領域 4 0 0 に表示中の動画以外の動画へのリンクとなる領域である。図 1 0 では 8 個の動画のサムネイルが表示されているが、表示される動画のサムネイルの個数は限定されない。また、動画のサムネイルの他に、動画の内容を示すテキスト情報が表示されていてもよい。

10

【 0 0 4 4 】

再生画面領域 4 0 0 は、動画選択領域 3 0 0 から選択された動画を再生する領域である。

オブジェクトタイムライン 5 0 0 は、再生画面領域 4 0 0 と伸縮タイムライン 6 0 0 との間に設けられ、キュー OC 1 ~ OC 1 0 のうち 1 以上のキューのスライダが操作された場合、操作されたキューに対応するオブジェクトが動画中に出現する箇所（タイミング）を、後述の伸縮タイムライン 6 0 0 に対応させて表示する領域である。

【 0 0 4 5 】

20

伸縮タイムライン 6 0 0 は、動画の時間経過を示す目盛を表示しており、動画の再生に伴ってスライダ 6 0 1 が目盛に沿って移動可能となっている。図 1 0 に示すように、キュー OC 1 ~ OC 1 0 のスライダ SD 1 ~ SD 1 0 が操作されていない状態（すなわち、重要度が 0.00 である状態）では、伸縮タイムライン 6 0 0 の目盛は等間隔で表示され、再生画面領域 4 0 0 に表示された動画が全てのフレームで同じ再生速度で再生されることとなる。

【 0 0 4 6 】

再生速度設定領域 7 0 0 は、再生画面領域 4 0 0 に表示された動画全体を何倍の速度で再生するかをスライダ 7 0 1 によって設定する領域である。

【 0 0 4 7 】

30

次に、図 1 1 及び図 1 2 を参照して、ユーザが実際にインタフェース 1 0 0 を用いて動画を閲覧する際に行われる閲覧設定処理の一例について説明する。

【 0 0 4 8 】

まず、キュー OC 1 ~ OC 1 0 のうち少なくとも 1 つのキューのスライダがユーザによって操作されると、操作されたキューに対応するオブジェクトの重要度が 0.00 以外の値に指定される（ステップ S 1 0）。図 1 2 に示すインタフェース 1 0 0 では、オブジェクトが “car”、“bird” 及び “bottle” に対応するキュー OC 2、OC 4 及び OC 5 のスライダ SD 2、SD 4 及び SD 5 がそれぞれ、0.56、1.00 及び 1.00 の位置に操作された例、すなわち、“car”、“bird” 及び “bottle” の重要度がそれぞれ、0.56、1.00 及び 1.00 に指定された例を示している。

40

【 0 0 4 9 】

次に、表示制御装置 1 0 は、ステップ S 1 0 で重要度が 0.00 以外の値に指定されたオブジェクトが検出された箇所をオブジェクトタイムライン 5 0 0 上に表示する（ステップ S 1 1）。例えば、オブジェクトが “car”、“bird” 及び “bottle” に対応するキュー OC 2、OC 4 及び OC 5 が、それぞれ、緑色、水色、ピンク色で表示されていた場合、オブジェクトタイムライン 5 0 0 では、“car” が出現する区間 TL 2 が緑色でハイライトされ、“bird” が出現する区間 TL 4 が水色でハイライトされ、“bottle” が出現する区間 TL 5 がピンク色でハイライトされる。このように、重要度が 0.00 以外の値に指定されたオブジェクトが出現する区間をオブジェクトタイムライン 5 0 0 上でキューの色に対応させて表示させることで、ユーザが関心をもったオブジェクトが動画全体のどのタイミン

50

グで出現するかが一目で確認可能となる。

【0050】

さらに、再生速度設定領域700のスライダ701の移動により、動画全体を何倍の速度で再生するかが指定される(ステップS12)。表示制御装置10は、ステップS10及びS12におけるユーザの指定に基づいて、動画の各フレームの再生速度を算出する(ステップS13)。各オブジェクトの重要度及び動画全体の再生速度に基づくフレーム毎の再生速度(表示時間)の算出方法は、非特許文献1に開示された方法を用いることができる。

【0051】

そして、表示制御装置10は、ステップS13で算出した各フレームの再生速度を伸縮タイムライン600上に反映させる(ステップS14)。具体的に、表示制御装置10は、各フレームの再生速度に応じて伸縮タイムライン600上の目盛の幅を変更する。再生速度が遅くなるにしたがって伸縮タイムライン600の目盛の幅が広くなる。また、表示制御装置10は、ステップS10において重要度が0.00以外の値に指定されたオブジェクトが出現する区間を伸縮タイムライン600に沿ってハイライト表示させる。例えば、図12の伸縮タイムライン600において、“car”が出現する区間ET1、“bottle”が出現する区間ET2、及び“bird”が出現する区間ET3が赤くハイライトされる。伸縮タイムライン600において赤くハイライトされた区間E1~E3は他の区間NTよりも遅く再生される。このように、ユーザによって重要度が0.00以外の値に指定されたオブジェクトが出現するシーンとその他のシーンの再生速度の違いを伸縮タイムライン600上で明確に区別させることが可能となる。また、表示制御装置10は、伸縮タイムライン600の区間E1~E3を、それぞれ、オブジェクトタイムライン500上の区間TL2、TL5、及びTL4に対応させる。これにより、ユーザは、動画を閲覧する前に、関心のあるオブジェクトが出現するシーンの再生速度と出現するタイミングの双方を把握することができる。

【0052】

次に、本実施形態のインタフェース100の評価実験について説明する。

一人称視点映像は見回り(監視)や伝統技能の保存、ライフログなど、幅広く応用することができ、それぞれの映像において重要なシーンが異なる。そこで、様々なシーンをウェアラブルカメラで撮影した一人称視点映像を用意し、その中から特定のシーンを、本実施形態のインタフェース100(以下、インタフェースと呼ぶ。)及び非特許文献1に開示されたインタフェース(以下、インタフェースと呼ぶ。)を用いて発見するタスクを実験参加者に与えることで、インタフェースに対する評価実験を行った。

【0053】

評価実験では、図13Aに示すように、8種類のシーンをウェアラブルカメラで撮影した動画(映像1~映像8)をそれぞれ2つのデータセットA及びBに分け、計16本の映像を用意した。実験参加者は、インターネット上で公開されている一般的な映像閲覧システムを使用した経験のある大学生16名である。図13Bに示すように、実験参加者を8名ずつ2つのグループXとYに分け、グループXは、データセットAをインタフェースで、データセットBをインタフェースで閲覧し、グループYは、データセットAをインタフェースで、データセットBをインタフェースで閲覧した。

【0054】

図14は、映像1~映像8のそれぞれについて、検出回数の多い順でのオブジェクトの名称と、本実施形態のインタフェースを用いて選択されたキュー(オブジェクトの名称)を示している。インタフェースを用いた場合、映像の種類に応じてキューが動的に生成されることで、図14に示すように、単純に検出回数の多い順のオブジェクトとは異なるオブジェクトのキューが選択された。例えば、“person”はどの種類の映像に対しても検出されているが、映像3(市街地散策)と映像8(テーマパーク観光)では映像の大部分で歩行者が撮影されるため、“person”(人物)のキューは有効なキューとして選択されなかった。このように、インタフェースを用いることで、映像全体を指定するような

10

20

30

40

50

冗長なキューを排除して有効なキューの組み合わせを選択することが可能となる。

【 0 0 5 5 】

評価実験では、各映像から特定のシーンを2秒程度抜き出して実験参加者に提示し、インタフェース及びを用いて、提示されたシーンを見つけるタスクを実験参加者に与え、その完了時間(タスク完了時間)を測定した。そして、タスク完了時間から平均閲覧速度を計算した。平均閲覧速度は、非特許文献1においてインタフェースの高速閲覧性能を図る尺度として導入されたもので、図15に示すように、特定のシーンが映像内で位置している時間(目的シーンの位置(s))をそのシーンを発見するタスクの完了時間(タスク完了時間(s))で割ったものである。平均閲覧速度が大きいシステムほど、効率的に目的のシーンを発見可能なシステムとなる。

10

【 0 0 5 6 】

そして、インタフェース及びのそれぞれについて、8種類の映像それぞれに対するタスクとそれらを合計したタスク全体における平均閲覧速度から95%信頼区間を計算した。図16は、インタフェース及びのそれぞれについて、映像毎に平均閲覧速度の実験参加者平均と95%信頼区間を示したものである。図16に示すように、ほとんどの映像について、本実施形態のインタフェースの方が優位な有意差(図16において「*」が付された箇所を参照。)が得られた。

【 0 0 5 7 】

図17に、インタフェース及びインタフェースに対する実験参加者(16名)の主観評価の結果を7段階評価(ローマ数字のI~VII)で示す。図17では、インタフェースが優位の回答ほどローマ数字の値が大きくなり、インタフェースが優位の回答ほどローマ数字の値が小さくなっている。また、ローマ数字に関連付けて付されているアラビア数字は、実験参加者の数である。例えば、「VII 10」は10名の実験参加者がインタフェースの方が優位であると回答していることを示す。図17に示すように、4つの全ての質問に対し、過半数の参加者が本実施形態のインタフェースの方を高く評価していることがわかる。

20

【 0 0 5 8 】

このように、本実施形態によれば、動画に含まれる各オブジェクトの重要度を指定するためのキュー(指定部)を動画の内容に応じて動的に生成しインタフェースに提示することにより、より効率的に動画から特定のシーンを発見することが可能となる。

30

【 0 0 5 9 】

本発明は、上述の実施形態に限定されるものではなく、本発明の趣旨を逸脱しない範囲内で種々の変更が可能である。

【 0 0 6 0 】

例えば、本実施形態では、動画からオブジェクト(物体、人物など)のみを検出して動画の閲覧に有効なキューを動的に自動生成するようにしたが、動画からオブジェクト以外の情報(例えば、撮影場所、被写体の動作など)も検出してキューを自動生成するようにしてもよい。

【 符号の説明 】

【 0 0 6 1 】

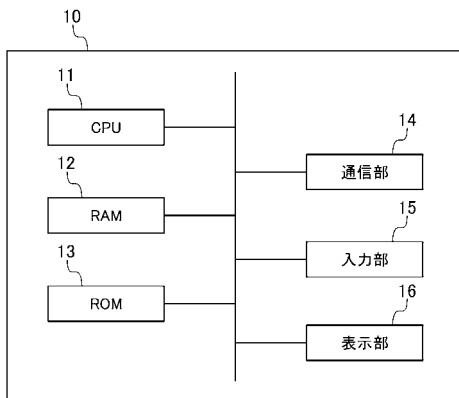
40

- 1 0 表示制御装置
- 1 1 CPU
- 1 2 RAM
- 1 3 ROM
- 1 4 通信部
- 1 5 入力部
- 1 6 表示部
- 1 0 0 インタフェース
- 2 0 0 オブジェクトキュー領域
- 2 0 1 オブジェクト名表示領域

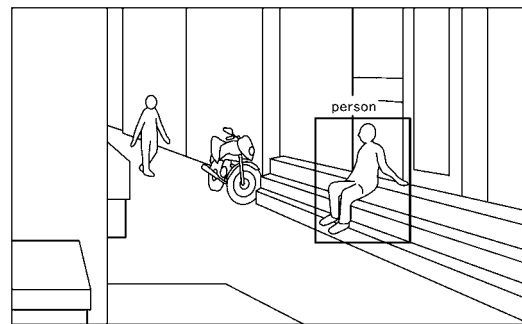
50

- 2 0 2 キュー表示領域
- 3 0 0 動画選択領域
- 4 0 0 再生画面領域
- 5 0 0 オブジェクトタイムライン
- 6 0 0 伸縮タイムライン
- 6 0 1 スライダ
- 7 0 0 再生速度設定領域
- 7 0 1 スライダ
- OC 1、OC 2、OC 3、OC 4、OC 5、OC 6、OC 7、OC 8、OC 9、OC 1 0
- キュー
- SD 1、SD 2、SD 3、SD 4、SD 5、SD 6、SD 7、SD 8、SD 9、SD 1 0
- スライダ

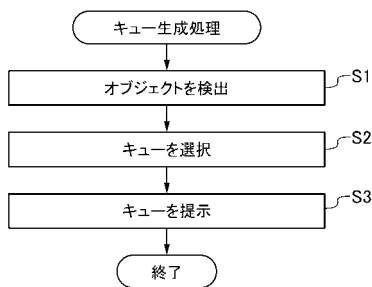
【 図 1 】



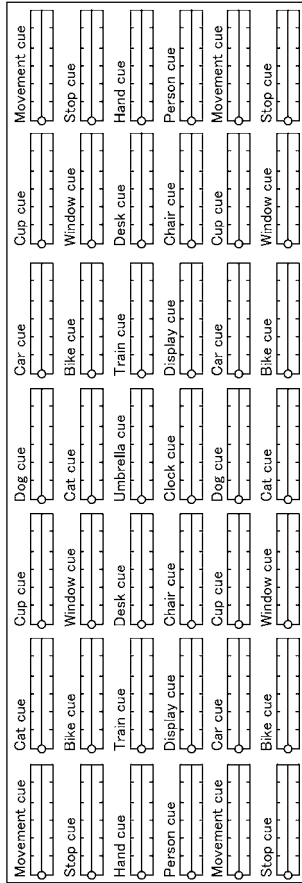
【 図 3 】



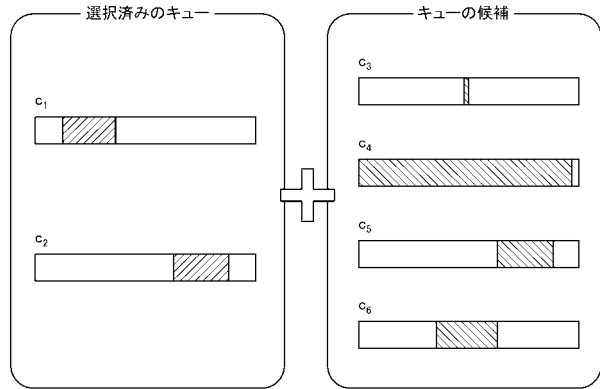
【 図 2 】



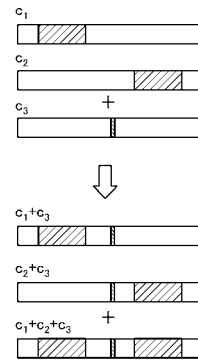
【 図 4 】



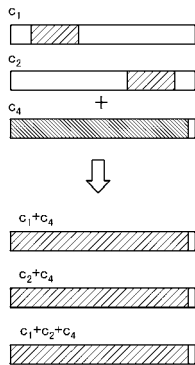
【 図 5 】



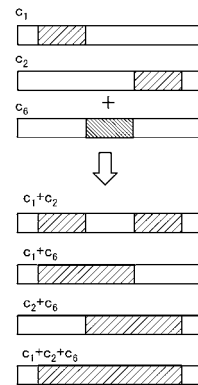
【 図 6 A 】



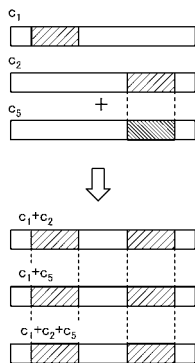
【 図 6 B 】



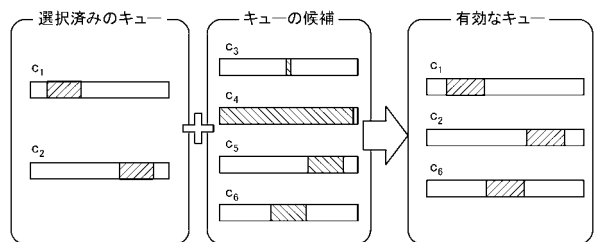
【 図 6 D 】



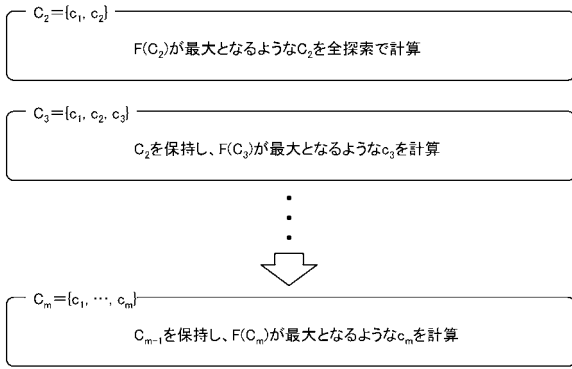
【 図 6 C 】



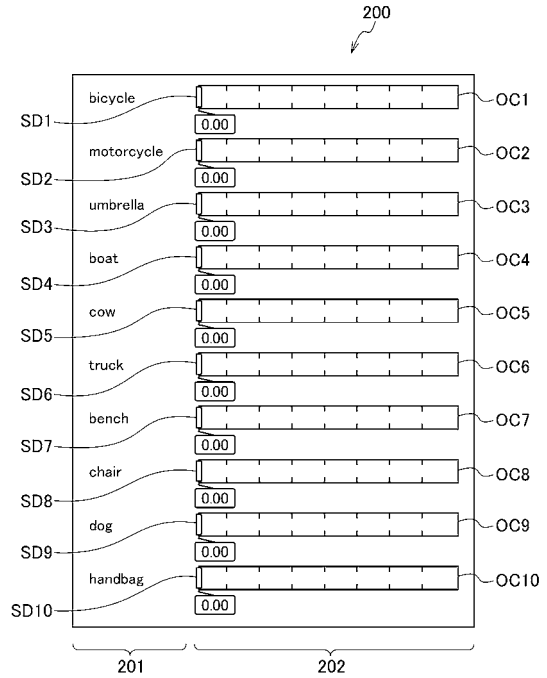
【 図 7 】



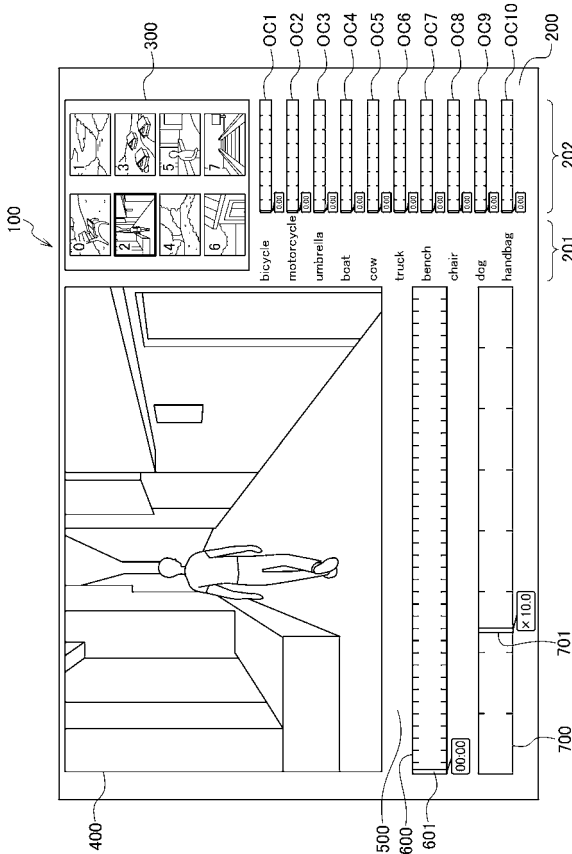
【 図 8 】



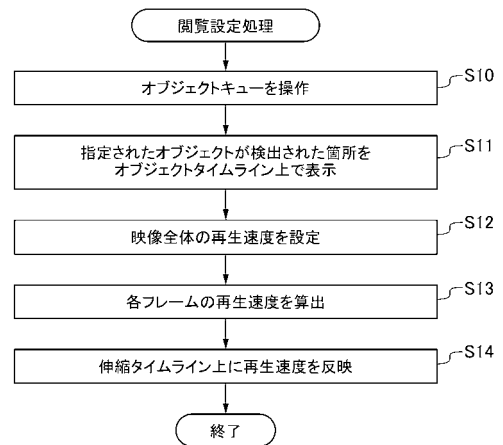
【 図 9 】



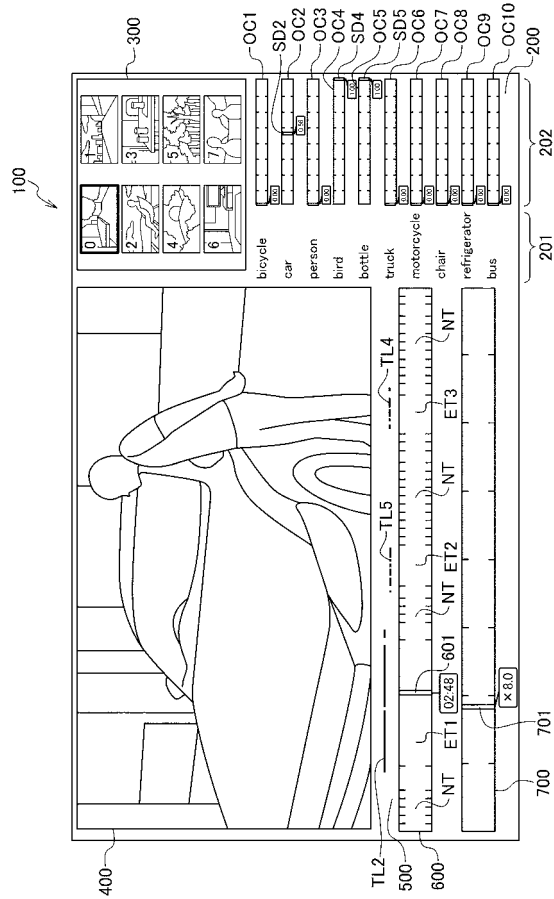
【 図 10 】



【 図 11 】



【図 1 2】



【図 1 3 A】

	映像の種類(シーン)	データセット
映像1	公園の散歩	A B
映像2	自転車レース	A B
映像3	市街地散策	A B
映像4	買い物	A B
映像5	犬に装着した映像	A B
映像6	バレーボール観戦	A B
映像7	料理	A B
映像8	テーマパーク観光	A B

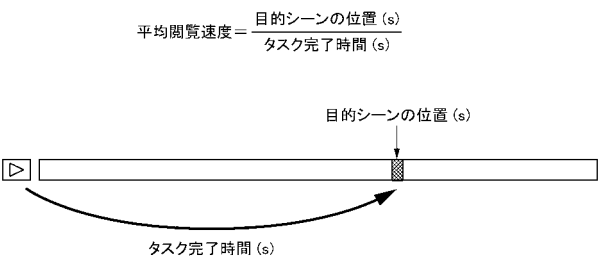
【図 1 3 B】

	データセットA	データセットB
グループX	インタフェースβ	インタフェースα
グループY	インタフェースα	インタフェースβ

【図 1 4】

	検出回数の多い オブジェクト	選択された キュー		検出回数の多い オブジェクト	選択された キュー
映像1	person car bicycle truck bottle bird	bicycle car person bird bottle truck	映像5	person dog car chair train backpack	person dog car chair train backpack
映像2	person bicycle car traffic light motorcycle truck	person bicycle car traffic light motorcycle umbrella	映像6	person car airplane train bench sink	person TV train traffic light umbrella airplane
映像3	person motorcycle bicycle umbrella cow boat	bicycle motorcycle umbrella boat cow truck	映像7	bottle person bowl dining table cup sink	person bottle sink bowl refrigerator TV
映像4	person refrigerator bottle cup toothbrush cell phone	person refrigerator bottle cup toothbrush cell phone	映像8	person potted plant backpack dining table chair cup	backpack potted plant dining table chair suitcase cell phone

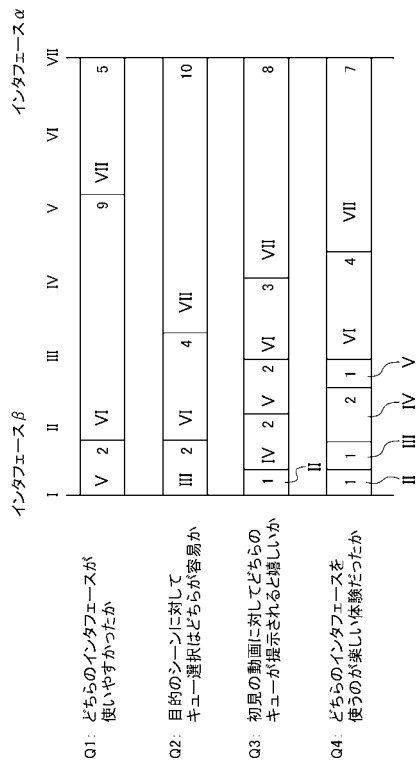
【図 1 5】



【図 1 6】

映像	インタフェースα		インタフェースβ	
	平均	95%信頼区間	平均	95%信頼区間
映像1	29.8	23.6~36.0*	11.7	10.3~13.2
映像2	18.1	13.9~22.2*	8.31	5.94~10.7
映像3	39.6	28.2~51.1*	11.8	7.01~16.5
映像4	18.6	13.3~23.9*	9.22	6.95~11.5
映像5	22.8	10.8~34.7*	6.92	4.01~9.83
映像6	2.03	1.48~2.57	3.26	2.77~3.75*
映像7	14.3	12.4~16.3*	8.78	7.33~10.2
映像8	71.6	45.2~98.0*	23.8	19.7~28.0
Total	13.7	10.9~16.5*	7.84	7.01~8.68

【 図 17 】



フロントページの続き

特許法第30条第2項適用申請有り https://ipsj.ixsq.nii.ac.jp/ej/index.php?active_action=repository_view_main_item_detail&page_id=13&block_id=8&item_id=184119&item_no=1、平成29年11月1日(抄録) https://ipsj.ixsq.nii.ac.jp/ej/index.php?active_action=repository_view_main_item_detail&page_id=13&block_id=8&item_id=184119&item_no=1、平成29年11月1日(予稿) 情報処理学会 第209回 コンピュータビジョンとイメージメディア研究会(CVIM209) 北海道札幌市北区北21条西11丁目 北海道大学 フード&メディカルイノベーション国際拠点、平成29年11月8日 <https://www.wiss.org/WISS2017/Program.html>、平成29年10月30日 <http://www.wiss.org/WISS2017Proceedings/>、平成29年12月6日 <https://www.youtube.com/watch?v=qo3IXAdIi4o>、平成29年12月6日 第25回インタラクティブシステムとソフトウェアに関するワークショップ(WISS2017) 山梨県北杜市大泉町西井出8240-1039八ヶ岳ロイヤルホテル、平成29年12月6日 <https://www.youtube.com/watch?v=FyUPaFXFWvA>、平成30年2月16日 <https://chi2018.acm.org/technical-program/> <https://dl.acm.org/citation.cfm?doid=3170427.3189085>、平成30年2月27日 <https://www.youtube.com/watch?v=77ueFtn66-M>、平成30年3月3日 <https://www.youtube.com/watch?v=dFHfvuhWETo>、平成30年4月7日

(出願人による申告)平成29年度 国立研究開発法人科学技術振興機構 戦略的創造研究推進事業、研究題目「鑑賞・創作支援のための音楽動画分析合成技術に関する研究開発」、産業技術力強化法第19条の適用を受ける特許出願 平成26年度 国立研究開発法人科学技術振興機構 戦略的創造研究推進事業、研究題目「複数ウェアラブルカメラ映像の統合によるグループの注意・行動解析」、産業技術力強化法第19条の適用を受ける特許出願

- (72)発明者 中村 優文
東京都新宿区戸塚町1丁目104番地 学校法人早稲田大学内
- (72)発明者 佐藤 洋一
東京都文京区本郷七丁目3番1号 国立大学法人東京大学内
- (72)発明者 樋口 啓太
東京都文京区本郷七丁目3番1号 国立大学法人東京大学内
- (72)発明者 米谷 竜
東京都文京区本郷七丁目3番1号 国立大学法人東京大学内
- Fターム(参考) 5C053 HA21 HA29 LA06