

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2014 年 5 月 1 日(01.05.2014)



(10) 国際公開番号

WO 2014/065033 A1

- (51) 国際特許分類:
G06F 17/30 (2006.01) H04N 5/91 (2006.01)
H04N 5/76 (2006.01)
- (21) 国際出願番号: PCT/JP2013/074507
- (22) 国際出願日: 2013 年 9 月 11 日(11.09.2013)
- (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 2012-236680 2012 年 10 月 26 日(26.10.2012) JP
- (71) 出願人: 株式会社日立国際電気(HITACHI KOKU-SAI ELECTRIC INC.) [JP/JP]; 〒1018980 東京都千代田区外神田四丁目 1 4 番 1 号 Tokyo (JP).
- (72) 発明者: 小倉 慎矢(OGURA, Shinya); 〒1878511 東京都小平市御幸町 3 2 番地 株式会社日立国際電気内 Tokyo (JP). 平井 誠一(HIRAI, Seichi); 〒1878511 東京都小平市御幸町 3 2 番地 株式会社日立国際電気内 Tokyo (JP). 新保 直之(SHIMBO, Naoyuki); 〒1878511 東京都小平市御幸町 3 2 番地 株式会社日立国際電気内 Tokyo (JP). 高田 智巳(TAKADA, Tomomi); 〒1878511 東

京都小平市御幸町 3 2 番地 株式会社日立国際電気内 Tokyo (JP).

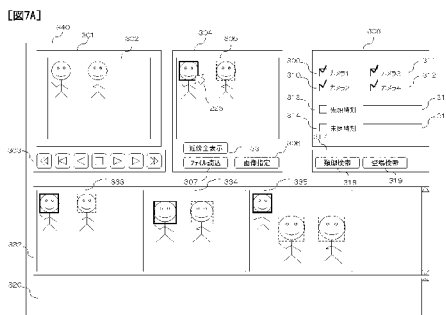
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (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, LR, NA, RW, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BG, CZ, EE, HR, HU, IS, LT, LV, MK, MT, NO, PL, RO, RS, SI, SK, SM), OAPI (KM).

添付公開書類:

— 国際調査報告 (条約第 21 条(3))

(54) Title: SIMILAR IMAGE RETRIEVAL DEVICE

(54) 発明の名称: 類似画像検索装置



306 Specify image
307 Read file
309 Camera 1
310 Camera 2
311 Camera 3
312 Camera 4
313 Start time
314 End time
318 Retrieve similar
319 Retrieve appearance
331 Display entire vicinity

(57) Abstract: In order to simplify a repetitive operation when a key image for retrieval (a face) is specified in moving images, when a face is clicked, a face detection process is performed with respect to images which are automatically advanced by one frame, and thus the relevant person can be selected successively, and a scene in which the relevant person appears continuously over time can be selected without operating a viewer. When the relevant person is not detected (for example, when the relevant person is covered by a person in the foreground), the "Next" button is pressed. As another example, when one face is specified all of the face detection results in the vicinity of the specified image (for example, from five seconds before to five seconds after the specified image, for a total of ten seconds) are displayed, and the user specifies the face of the relevant person in those results, and thus multiple images of a series of persons can be specified without operating a viewer. As another example, the results of the automatic tracking of a person are displayed on an operation terminal, and thus the user can check for a mistake in the results of the automatic tracking of a person, and if there is a mistake, the mistake can be corrected.

(57) 要約: 動画から検索キー画像(顔)を指定する際の繰返し操作を簡素化する。顔をクリックすると、自動的に1コマ進めた画像に対して顔検出処理を行うことで、次々に当該人物を選択でき、当該人物が時間的に連続に登場する場面をビューア操作せず選択できるようにした。なお、当該人物が検出されない(例えば前の人と重なる)とき『次ボタン』を押す。別の例として、顔を1つ指定すると、指定画像の近傍(例えば5秒前~5秒後までの10秒間)の顔検出結果が全

て表示され、その中から当該人物の顔をユーザが指定することでビューア操作せず一連の人物の複数画像を指定できるようにした。更に別の例として、自動的に人物追跡された結果を操作端末に表示することで、自動的に人物追跡された結果に間違いが無いかユーザが判定でき、間違いがあれば修正できるようにした。

WO 2014/065033 A1

明 細 書

発明の名称： 類似画像検索装置

技術分野

[0001] 本発明は、類似画像検索装置に係り、特に、動画像の中から所望の複数の検索キーを指定して検索を行う類似画像検索装置に関する。

背景技術

[0002] 従来、監視カメラ等で撮影され或いは記録された映像（動画像）の中から、所望の人物を画像認識技術等を用いてコンピュータに検索させる人物検索システムが知られる（例えば、特許文献1乃至4参照。）。このような、タグ付け等の外部情報に拠らずに画像そのものの特徴に基づき検索する技術は、一般にC B I R（Content-Based Image Retrieval）と呼ばれる。

特許文献1乃至4が用いる画像認識技術は、画像から人物（の顔）が映った部分を切出し、人物を個々に特定するための特徴量として色ヒストグラム等を抽出し、この特徴量が所望の人物のものと類似する場合に、同一人物であると推定するものである。

[0003] このような類似顔画像検索システムでは、検索の精度はキー画像の条件、例えば表情、顔の向き、照明の当たり方、顔の検出（切出し）のズレ具合に影響され、キー画像での条件と近いものが、検索結果として出てくることが多い。

一方で、ユーザは検索キーに指定した人物を、そのような条件に影響されずに検索できることを望む。このような場合ユーザは、指定された複数のキー画像についてキー画像を1つずつ用いて検索しそれらの結果を結合して表示する複数キー検索機能を利用することができる。すなわち、最初に、所望の人物画像（1枚でよい）をキー画像にして検索する。その検索結果を確認し、同一人物の画像が含まれていれば、次にその画像に前後する画像からなる動画の1コマずつをキー画像に指定して、上記複数キー検索機能を用いて再度検索する。これにより様々な条件の画像を見つけられる可能性が向上す

る。

先行技術文献

特許文献

[0004] 特許文献1：特開 2 0 0 9 － 0 2 7 3 9 3 号公報

特許文献2：特開 2 0 1 1 － 2 3 7 8 7 9 号公報

特許文献3：特開 2 0 1 1 － 1 0 7 7 9 5 号公報

特許文献4：特開 2 0 1 1 － 0 4 8 6 6 8 号公報

特許文献5：特開 2 0 1 0 － 1 4 0 4 2 5 号公報

非特許文献

[0005] 非特許文献1：Michael J. Swain and Dana H. Ballard, " Color indexing", International Journal of Computer Vision, 米国, Springer, 1 9 9 1 年 1 月, 第 7 巻, 第 1 号, p. 1 1 － 3 2

発明の概要

発明が解決しようとする課題

[0006] 上述したとおり、最善の 検索を期待するユーザは、検索したい人物の画像を既得の動画中からできるだけ多く特定し、それらを全てキー画像に指定し、未特定の画像を検索しようとする。このとき、既得動画中に検索対象人物が 1 回（1 コマ或いは複数コマ連続する 1 シーン）しか映っていないことが保証されている場合（例えば、一人ずつしか通過できないレーン等で通過に同期して撮影するような運用など）は、容易に動画中の当該人物の全ての画像を自動的に選択できる。

一方、通常の防犯カメラ映像のように、検索対象人物が複数回映る可能性があり、更にそれらの画像において他の人物も一緒に映る可能性がある場合は、ユーザは動画の 1 コマずつをコマ送りする再生装置の操作をマウスクリックで行った後、画像指定ボタンを押して顔検出させ、顔検出結果の含まれた複数の顔から探したい人物の顔をマウスクリックして指定するという 3 つの手順を、画像の枚数だけ繰り返す必要があり非常に手間がかかる。

[0007] この問題を図 8、図 9 を参照して詳述する。図 8 は、従来の端末装置 103 に表示される検索画面 400 を示す図である。

図 8 の検索画面 400 は、再生画像表示領域 301、画像再生操作領域 303、検索キー画像指定領域 304、検索絞込パラメータ指定領域 308、検索実行領域 317、及び、検索結果表示領域 320 よりなる。

画像再生操作領域 303 は、録画装置に記録された画像を再生操作する領域である。画像再生操作領域 303 を構成する各ボタンには、それぞれ固有の再生種類が割当てられており、例えば左から、巻戻し、コマ戻し、逆再生、再生停止、順再生、コマ送り、早送りの再生種類が割当てられている。ユーザが各ボタンをマウス 282 で適宜押下することにより、ボタンに割当てられた再生種類で再生画像表示領域 301 に動画 302 が再生される。

[0008] 検索キー画像指定領域 304 は、検索キー画像の指定と表示を行う領域である。本領域は、キー原画像 305 と、画像指定ボタン 306、ファイル読込ボタン 307 よりなる。

キー原画像表示部 305 は、類似検索のためのキー画像或いはその元となる画像（キー原画像と呼ぶ）を表示する領域である。初期状態においては検索キー原画像は未指定であるので、画像表示はされていない。

画像指定ボタン 306 は、画像再生操作領域 303 に現在表示されている動画 302 をキー原画像に指定するボタンである。例えば動画 302 を再生停止状態にし画像指定ボタン 306 を押すと、そのときの画像がキー原画像に指定され、キー原画像表示部 305 にも表示される。キー原画像が新たに表示される都度、必要に応じ顔検出処理が実行され、検出された顔を切出す際の枠が自動的に付加される。枠は、初期的には（複数あるときは全て）非選択状態となっている。枠の 1 つを選択すると、キー画像（キー顔）の指定が完了する。

ファイル読込ボタン 307 は、録画装置 102 に記録されている画像以外の画像、例えば、デジタルカメラで撮影した画像やスキャナで取込んだ画像や動画像を、画像再生操作領域 303 に表示させるボタンである。このファ

イル読込ボタン３０７を押下すると、ファイルを開くダイアログボックスが表示され、そこで指定したファイルが読み込まれ、画像再生操作領域３０３で再生可能な状態になるか、自動的に再生が始まる。ファイルが静止画であれば、そのままキー原画像に指定されたことになり、キー原画像表示部３０５に表示される。

[0009] 検索絞込パラメータ指定領域３０８は、検索の際の絞込パラメータの種類とその値（範囲）を指定する領域である。本領域は、撮像装置指定チェックボックス３０９、３１０、３１１、３１２と、時刻指定チェックボックス３１３、３１４、時刻指定欄３１５、３１６から構成する。

撮像装置指定チェックボックス３０９、３１０、３１１、３１２は、検索の際に検索対象とする撮像装置を指定するボタンである。本ボタンは、押下すると選ばれたことを示すチェックマークがそれぞれ表示される。このマークは再押下すると非表示となり、押下で表示・非表示を繰り返す。初期状態においては、全撮像装置を検索対象とするため、撮像装置指定チェックボックスは全て選択状態となる。

時刻指定チェックボックス３１３、３１４は、検索の際に検索対象とする時刻範囲を指定するボタンである。表示の態様については本ボタンも他のチェックボックスと同様である。時刻指定チェックボックス３１３を選択状態にした場合には時刻範囲に先頭時刻を与える。非選択状態にした場合には、時刻範囲に先頭時刻を与えない、すなわち、録画装置に記録された最も古い時刻の画像までを検索対象範囲とすることを意味する。時刻指定チェックボックス３１４も同様であり、これを非選択状態にした場合には、録画装置に記録された最も新しい時刻の画像までを検索対象範囲とすることを意味する。

時刻指定欄３１５、３１６は、上述の先頭時刻と末尾時刻の値を指定する入力欄である。

初期状態においては、全時間帯を検索対象とするため、時刻指定チェックボックスは全て非選択状態、時刻指定欄は空欄とする。

[0010] 検索実行領域 317 は、検索実行を指示する領域である。本領域は、類似人物検索ボタン 318 と登場イベント検索ボタン 319 よりなる。

類似人物検索ボタン 318 は、キー原画像表示部 305 による類似人物検索実行を指示するボタンである。検索絞込パラメータ指定領域 308 にてパラメータが指定されている場合には、指定されたパラメータに従って検索の実行をすることを指示する。

登場イベント検索ボタン 319 は、登場イベント検索実行を指示するボタンである。通常、監視カメラシステム等では、動き検出や人感センサ発報、入退出管理等の他システムからの通知等のイベントを映像と間接的に対応付けて記録したり、映像の記録自体もそれらの発報があったときだけ行ったりしており、これらをイベント記録などと呼んでいる。登場イベント検索は、記録された各種イベントの内、人物の顔が正面方向から撮影されることが期待できるイベントに対応付けられた映像のみを、検索対象とするものである。検索絞込パラメータ指定領域 308 にてパラメータが指定されている場合には、指定されたパラメータに従って検索の実行をすることを指示する。

[0011] 図 9 は、操作画面 400 で、動画の 1 コマずつを複数のキー画像として検索するための各手順の操作を示す図である。動画の 1 コマずつを複数のキー画像として検索したい場合、ユーザはまず、画像再生操作領域 303 を使用して、キー画像に使いたい動画の先頭の画像を再生画像表示領域 301 に表示させる（手順 1 と呼ぶ）。次に、画像指定ボタン 306 を押す（手順 2 と呼ぶ）。次に、キー原画像表示部 305 に表示されている顔画像を示す矩形の中から、検索キーにしたい顔をマウスで指定する（手順 3 と呼ぶ）。すると端末装置 103 の内部で、このときの画像（の ID）と指定した顔の領域の情報が保持される。保持した情報は、人物検索ボタン 318 が押されるまで蓄積され、次にユーザは画像再生操作領域 303 のコマ送り機能を用いて、動画 302 を 1 コマ送る（手順 4）。次に先ほどと同様に手順 2 と手順 3 を行う。以降、動画の 1 コマずつを複数のキー画像として検索するための末尾の画像になるまで、手順 2、手順 3、手順 4 を繰り返す。

この、手順 2、手順 3、手順 4 の繰り返しがユーザにとって非常に面倒な作業であるという課題がある。

[0012] 本発明は、このような問題に鑑みてなされたものであり、単調な繰り返し作業を自動化により排除し、一般ユーザに使い易いユーザインターフェース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

[0013] 本発明を概略的に述べると、本発明の一側面では、顔をクリックすると 1 コマ進めた画像に対して顔検出処理を行うことで、次々に当該人物を選択でき、当該人物が時間的に連続に登場する場面をビューア操作せず選択できるようにした。なお、当該人物が検出されない（例えば前の人と重なる）とき『次ボタン』を押す。

[0014] 本発明の他の側面では、顔を 1 つ指定すると、指定画像の近傍（例えば 5 秒前～5 秒後までの 10 秒間）の顔検出結果が全て表示され、その中から当該人物の顔をユーザが指定することでビューア操作せず一連の人物の複数画像を指定できるようにした。

[0015] 本発明の他の側面では、自動的に人物追跡された結果を操作端末に表示することで、自動的に人物追跡された結果に間違いが無いかユーザが判定でき、間違いがあれば修正できるようにした。例えば、自動判定された一連の人物の顔を太枠に、その他を破線枠で区別して表示する。

[0016] 本発明のより具体的な一側面では、カメラで撮影し記録した動画画像から、所定の被写体が映った部分画像をキーとして指定し、指定された該部分画像の特徴量に近い特徴量を有する画像を検索する類似画像検索装置において、
動画画像の中から所望の 1 コマを指定するためのプレビューを提供する手段と、

前記プレビューされた 1 コマをキー原画像に指定する操作を受付ける手段と、

前記受付ける手段で指定された 1 コマに時間的に近傍の複数コマを、自動的にキー原画像に追加指定する手段と、

前記受付ける手段或いは前記追加指定する手段でキー原画像に指定された 1 ないし複数コマを、該 1 ないし複数コマから所定のアルゴリズムで検出された被写体に対応する領域を示す図形を付加して表示する手段と、

付加された前記図形の初期状態を非選択状態とし、1 つのコマに対しいずれか 1 つの図形を選択状態とする操作を受付ける選択手段と、

選択状態となった前記図形に対応する複数の被写体をキーとする検索を要求する手段と、を有する。

[0017] 上記の類似画像検索装置において、前記被写体は人の顔であり、前記特徴量は、前記動画像から自動的に検出された前記被写体ごとに予め抽出され、該自動的に検出された元の画像に対応付けて記録されており、更に、前記要求する手段からの要求を受けて、前記選択状態となった前記図形に対応する複数の被写体は同一人物であると看做して、該複数の被写体に対応する複数の特徴量を 1 つずつキーとして検索し、それら結果を結合して応答する検索実行手段を設けてもよい。

[0018] 上記の類似画像検索装置において、前記表示する手段は、受付ける手段で指定された 1 コマに対し、前記図形を付加して、キー原画像表示領域に表示するものであり、

前記選択手段がキー原画像表示領域に表示された前記図形に対する前記選択状態とする操作を受付けるか、或いは、所定のボタンの 1 回の押下のみに応じて、前記追加指定する手段が、そのとき前記キー原画像表示領域に表示されていた 1 コマに続く次のコマをキー原画像に追加指定することで、自動的に該次のコマが前記キー原画像表示領域に表示され、再び前記選択手段が操作を受付けるという動作を繰り返し、要求する手段は、該繰り返しの間に選択状態にされた複数の前記図形に対応する複数の被写体を前記キーとするように構成してもよい。

[0019] 上記の類似画像検索装置において、追加指定する手段は、前記受付ける手段により前記 1 コマをキー原画像に指定する操作が受け付けられた後、自動的に或いは所定のボタンの 1 回の押下のみに応じて、前記受付ける手段で指定

された 1 コマに時間的に前または後に連続する複数コマのうち、少なくとも被写体が検出された複数コマをキー原画像に追加指定し、前記表示する手段は、前記追加指定する手段で指定された複数のキー原画像を、夫々に前記図形を付加して指定済キー表示領域に表示するものであり、前記選択手段が、前記指定済キー表示領域に表示された複数のキー原画像に対する、前記図形を前記選択状態とする操作を複数受け、前記要求する手段が、前記選択状態にされた複数の前記図形に対応する複数の被写体を前記キーとするように構成してもよい。

[0020] 上記の類似画像検索装置において、前記表示する手段は、受け付ける手段で指定された 1 コマに対し、前記図形を付加して、キー原画像表示領域に表示するものであり、

前記自動追加手段は、前記選択手段が前記キー原画像表示領域に表示された前記図形に対する前記選択状態とする操作を受け付けた後の所定のボタンの 1 回の押下のみに呼応して、該選択状態にされた前記図形に対応する被写体を起点として、記録された動画像の中から当該被写体に時空間的連続性を満たす被写体を追跡し、当該追跡が成功した範囲の複数のコマをキー原画像に追加指定するものであり、前記表示する手段は、前記追加指定する手段で指定された複数のキー原画像を、夫々に前記図形を付加して指定済キー表示領域に表示するように構成してもよい。

発明の効果

[0021] 本発明によれば、動画からの複数の検索キーの指定を自動もしくは半自動で行うようにしたので、簡単な操作で精度の高い検索を行うことができる。

図面の簡単な説明

[0022] [図1]類似画像検索システムの構成図（実施例 1～3）。

[図2]撮像装置 201 のハードウェア構成図。

[図3]録画装置 102 のハードウェア構成図。

[図4]端末装置 103 のハードウェア構成図。

[図5]複数キー画像検索を行う検索画面 300 を示す図（実施例 1）。

[図6]複数キー画像検索を行う検索画面330を示す図（実施例2）。

[図7A]複数キー画像検索を行う検索画面340を示す図（実施例3）。

[図7B]検索画面340でのキー画像の修正操作を示す図（実施例3）。

[図8]従来の端末装置103に表示される検索画面400を示す図。

[図9]従来の検索画面400で複数キー画像検索を行う手順を示す図。

発明を実施するための形態

[0023] 以下、本発明に係る一実施形態について図面を参照して説明する。なお、各図の説明において、実質的に同一な機能を有する構成要素には同一の参照番号を付し、説明を省略する。

[0024] まず、図1～図4を参照して、本発明の一実施形態に係る類似画像検索システムの構成について説明する。図1には、本発明の一実施形態に係る類似画像検索システムのシステム構成を例示してある。

類似画像検索システムは、図1に示すように、ネットワーク200に、撮像装置201、録画装置102、端末装置103が接続され、互いに通信可能な状態で構成される。

[0025] ネットワーク200は、データ通信を行う専用ネットワークやイントラネット、インターネット、無線LAN（Local Area Network）等の各装置を相互に接続して通信を行う通信手段である。

撮像装置201は、CCD（Charge Coupled Device）やCMOS（Complementary Metal Oxide Semiconductor）素子等で撮像した画像にデジタル変換処理を施し、変換結果の画像データを、ネットワーク200を介して録画装置へ出力するネットワークカメラや監視カメラ等の装置である。

[0026] 録画装置102は、ネットワーク200を介して撮像装置201より入力された画像データをHDD等の記録媒体に記録するネットワークビデオレコーダ等の装置である。また人物検索のための殆どの機能も搭載される。

録画装置102は、機能構成として、画像送受信部210、画像記録部211、再生制御部212、人物領域検出部213、人物特徴量抽出部214、人物特徴量記録部215、属性情報記録部216、要求受信部217、類

似人物検索部 218、登場イベント検索部 219、検索結果送信部 220 を有する。

[0027] 画像送受信部 210 は、装置外部からの画像の入出力を行う処理部であり、撮像装置 201 からの画像データの受信、端末装置 103 への画像データの送信等を行う。

画像記録部 211 は、画像データの記録媒体への書込みや読出しを行う。書込みの際には、画像データに加え、画像データを読出す際の情報となる画像 ID（画像の識別情報）も併せて記録する。

再生制御部 212 は、端末装置 103 への映像（ストリーム）再生を制御する。

[0028] 人物領域検出部 213 は、撮像装置 201 から受信した画像データに対し画像認識技術を用いた人物検出を行い、画像中の人物の存在判定をし、人物が存在する場合には、その顔を基準にして所定条件で顔の周囲まで含む、所定の縦横比を有する矩形の領域の座標算出を行う。

人物特徴量抽出部 214 は、人物領域検出部 213 で検出した領域に対して画像認識技術を用いて特徴量算出を行う。ここで算出する人物特徴量とは、静止画から抽出可能なものであり、例えば、検出した領域を一定サイズにスケーリングし、一律に分割した画素ブロック毎に色や輝度、あるいはそれらの勾配やテクスチャのヒストグラムを求め、その結果を集約した多次元ベクトルである。このような appearance-base の認識のほか、3次元形状を復元して認識するもの、例えば、顔の骨格に強く依存する輪郭や目や鼻、口に対応する特徴点の相対的配置関係を 3次元上で判断するものが挙げられるが、本実施形態においては、使用する特徴量の種類や数はいずれであってもよい。特徴量の次元を減らすため、Linde-Buzo-Gray 法などベクトル量子化の手法を用いることができる。ヒストグラムベースの特徴量であれば、画素ブロック 1 つ分を大津の 2 値化手法により最終的に 2 色で近似する Block Truncation Coding を用いても良い。

[0029] 人物特徴量記録部 215 は、人物特徴量抽出部 214 で算出した特徴量の

記録媒体への書込みと読出しを行う。大量の特徴量の中から高速に検索するために、特徴量はクラスタに分類され、それに応じ記録構造（ツリーや辞書の構成やHDD上のセクタ配置等）も最適化される。簡易的には、単純な規則のハッシュ関数で分類し、ハッシュ値と記録先を対応付ける辞書を作り、1分類に属する特徴量の数が増えすぎたときには、さらに階層を増やして細かく分類する方法がある。このほかEM法など各種の最適化アルゴリズムが知られる。

属性情報記録部216は、個々の画像データに関連する属性情報の記録媒体への書込みと読出しを行う。属性情報とは、例えば、画像の撮影時刻や撮像装置番号、各種イベントのフラグ、人物領域検出部213で検出した領域の座標等である。

[0030] 要求受信部217は、端末装置103からの検索要求の受信を行う。検索要求には、類似画像検索要求と、登場イベント検索要求がある。

類似人物検索部218は、要求受信部217にて受信した要求が類似人物検索要求であった場合に、類似画像検索を行う。基本的には、両者の特徴量ベクトルの差（ノルム）が小さいほど類似していると判断する。ヒストグラムベースの特徴量の場合、非特許文献1に記載のHistogram Intersectionにより1乃至複数ブロック分の類似度を求め、これらを重み付き加算したものを類似度とすることができ、この類似度が所定値以上のものを検索結果として出力する。

登場イベント検索部219は、要求受信部にて受信した要求が登場イベント検索要求であった場合に、登場イベント検索を行う。

検索結果送信部220は、類似人物検索部218や登場イベント検索部219から得た類似人物検索結果や登場イベント検索結果の端末装置103への送信を行う。

[0031] 端末装置103は、実際に検索を行う画像記録部211とユーザと間のインターフェースとなる装置であり、ネットワーク機能を有する一般のPC（パーソナルコンピュータ）で実現してもよい。

端末装置 103 は、機能構成として、検索要求送信部 221、検索結果受信部 222、検索結果表示部 223、再生画像表示部 224、画面操作検知部 225 の各処理部を有する。

[0032] 検索要求送信部 221 は、検索要求の録画装置 102 への送信を行う。類似人物検索の場合、キー画像（キー顔）が指定されるたびにそれを蓄積し、その後類似人物検索ボタン 318 等が押されたときに、蓄積した 1 乃至複数の検索キー画像と、絞込みパラメータを含む検索要求（クエリ）を送信するとともに、蓄積したキー画像をクリアする。なお、キー画像の蓄積や送信は、キー画像の画像データそのものではなく、その特徴量、或いはそれが抽出された原画像の ID と原画像内での位置情報の組等を用いてすることもできる。また送信したキー画像は別途、検索履歴として保存してもよい。

検索結果受信部 222 は、検索結果の録画装置 102 からの受信を行う。検索結果として受信するデータには、録画装置 102 において、類似人物検索、或いは、登場イベント検索を実施して得られた画像の集合が含まれる。集合を構成する個々の画像は、録画装置 102 に記録された映像から画像サイズ縮小処理等を施して生成される。以下、この個々の画像を「検索結果画像」、検索結果として送受信するデータを「検索結果データ」という。

検索結果表示部 223 は、検索結果受信部 222 にて受信した検索結果の画面表示を行う。表示される画面例については後述する。

再生画像表示部 224 は、DirectShow（商標）等を利用して、録画装置 102 から受信したされた画像データの復号や画面への再生（動画表示）を行う。

画面操作検知部 225 は、ユーザによる操作内容の検知及び取得を行う。

[0033] 図 2 には、本発明の一実施形態に係る類似画像検索システムに用いる撮像装置 201 のハードウェア構成を例示してある。

撮像装置 201 のハードウェア構成は、図 2 に示すように、撮像部 241、主記憶部 242、符号化部 243、ネットワークインタフェース（I/F）245 が、バス 240 で結合された形態である。

[0034] 撮像部 241 は、レンズで撮像した光信号をデジタルデータに変換する。符号化部 243 は、撮像部 241 が出力するデジタルデータを符号化して、J P E G (Joint Photographic Experts Group) などの画像データに変換する。主記憶部 242 は、撮像したデジタルデータ、符号化された画像データを記憶する。ネットワーク I / F 245 は、ネットワーク 200 を介して、主記憶部 242 上の画像データを録画装置 102 に送信するためのインタフェースである。

[0035] 図 3 には、本発明の一実施形態に係る類似画像検索システムに用いる録画装置 102 のハードウェア構成を例示してある。

録画装置 102 のハードウェア構成は、図 3 に示すように、C P U (Central Processing Unit) 251、主記憶部 252、補助記憶部 253、ネットワーク I / F 254 が、バス 250 で結合された形態である。

[0036] C P U 251 は、録画装置 102 の各部の制御と、機能を実現するためのプログラムの実行を行う。主記憶部 252 は、D R A M (Dynamic Random Access Memory) などの半導体装置で実現され、検索のための画像データや C P U 251 で実行するプログラムをロードして格納するための中間的なメモリである。補助記憶部 253 は、H D D やフラッシュメモリなどで実現され、主記憶部 252 より大容量のメモリであり、画像データやプログラムを格納する。ネットワーク I / F 254 は、ネットワーク 200 を介して、撮像装置 201 からの画像データを受信したり、端末装置 103 から検索キーワードを受信したり、端末装置 103 に画像データを送信するためのインタフェースである。

[0037] 図 4 には、本発明の一実施形態に係る類似画像検索システムに用いる端末装置 103 のハードウェア構成を例示してある。

端末装置 103 のハードウェア構成は、図 4 に示すように、C P U 261、主記憶部 262、補助記憶部 263、表示 I / F 264、入出力 I / F 265、ネットワーク I / F 266 が、バス 260 で結合された形態である。また、表示 I / F 264 は、表示装置 270 と接続され、入出力 I / F 26

5は、キーボード280やマウス282などの入出力装置と接続される。

[0038] CPU261は、端末装置103の各部の制御と、機能を実現するためのプログラムの実行を行う。主記憶部262は、DRAMなどの半導体装置で実現され、表示のための画像データやCPU261で実行するプログラムをロードして格納するための中間的なメモリである。補助記憶部263は、HDDやフラッシュメモリなどで実現され、主記憶部262より大容量のメモリであり、検索キーワード、画像データやプログラムを格納する。表示I/F264は、表示装置270と接続するためのインタフェースである。入出力I/F265は、キーボード280やマウス282などの入出力装置と接続するためのインタフェースである。ネットワークI/F266は、ネットワーク200を介して、録画装置102からの画像データを受信したり、録画装置102に検索キーワードを送信したりするためのインタフェースである。表示装置270は、例えば、LCD (Liquid Crystal Display) などの装置であり、画像や動画をその表示部に表示する装置である。ユーザは、表示装置270の表示部に表示された画像を、キーボード280やマウス282などの入出力装置を操作して、例えばGUI (Graphical User Interface) 操作することによって、端末装置103、及び、類似画像検索システムを操作する。

[0039] 次に、本発明の実施形態に係る類似画像検索システムにおける複数キー画像検索を、実施例1乃至3により説明する。

実施例 1

[0040] 本発明の実施形態である類似画像検索システムの実施例1を、図5を参照して説明する。

図5は、実施例1の端末装置103に表示される検索画面300を示す図である。本例の検索画面300は、新たに、次ボタン321を設けた点などで従来と異なる。

[0041] 本例の検索画面300を用いて、ユーザが動画の1コマずつを複数のキー画像として検索するための手順は、おおよそ以下のようになる。

まず、従来と同様に手順 1 を行う。つまり、画像再生操作領域 303 を使用して、キー画像に使いたい動画の先頭の画像（或いは任意の動画中の画像）を再生画像表示領域 301 に表示させる。なお、記録装置 201 に記録された動画の中からカメラや時刻を指定して再生するには、監視カメラシステムの分野で一般的な技術を使用すればよい。

次に、従来と同様に手順 2 を行う。つまり、画像指定ボタン 306 を押すことで、そのとき表示されている動画 302（先頭の画像である）が、キー原画像として取り込まれ、検索キー画像指定領域 304 にキー原画像 305 として表示される。

[0042] 次に、手順 3 として、表示されているキー原画像 305 中で、検索キーにしたい顔があれば、それを従来の手順 3 同様にマウス 226 で指定し、検索キーにしたい顔がなければ、次ボタン 321 を押す。

顔を指定した瞬間、このときの画像と指定した顔の領域の情報が端末装置 103 内に保持される。あるいは、顔画像の特徴量を（適宜端末装置 103 内で算出して）保持してもよい。保持したのち、端末装置 103 は、キー原画像 305 を 1 コマ進めたものに更新して検索キー画像指定領域 304 に表示する。次ボタン 321 を押した場合は、保持をせずにキー原画像 305 を 1 コマ進めたものに更新する。

[0043] ユーザはこの後、所望の人物が画像に映っている間、手順 3 のみを繰り返せばよく、手順 2 と 4 が省略できる。なお、本例を最も簡易に実装する方法は、ユーザのする操作と同じキーコードやマウスイベントを発生させる方法（エミュレーション）である。

[0044] 以下、本例の処理の一例を詳細に説明する。本例の実現に当り、顔検出や顔画像の特徴量の計算を何時何処で行うか、コマ送り中の画像を伝送し復号するか、等により幾つかの最適実装が考えられ、最も簡易な実装は、ユーザのする操作と同じキーコードやマウスイベントを発生させる方法（エミュレーション）である。以下の説明では、顔検出や特徴量は画像（動画）に埋め込まれた状態で端末装置 103 にストリーム送信されるものとする。

[0045] 従来同様の手順１での処理は次のようになる。MPEGのような動画であれば、RTSPやMRCP (Media Resource Control Protocol) 等プロトコルで、録画装置１０２と端末装置１０３間でセッションが開始され、端末装置１０３の再生画像表示部２２４から再生位置を指定したPLAYメッセージ等を送信して、所望のコマの画像が再生画像表示領域３０１に表示されているものとする。このとき再生画像表示部２２４の受信バッファには、所望のコマの後のコマも蓄えられた状態となる。静止画を一枚ずつ伝送する方法であっても、プリフェッチを行うことで同様に蓄えられた状態とする。

動画の場合、再生画像表示部２２４のDirectShowの入力フィルタ（スプリッタ）が、カメラ名、撮影時刻、画像ＩＤ、その画像内での顔領域の通し番号と座標（と特微量）を取り出して、描画フィルタ及びアプリケーションソフトに渡す。静止画の場合、画像ファイルのヘッダ等に埋め込まれたものを同様に取り出す。入力フィルタはMPEGビデオのエレメンタリストリームを復号フィルタに渡す。

描画フィルタ（レンダラ）は、カメラ名、撮影時刻を文字化および画像化し、それらを復号フィルタで復号された動画（静止画）に重畳して、表示Ｉ／Ｆ２６４への描画処理を行う。これにより再生画像表示領域３０１に動画３０２が表示される。書出しフィルタ（グラバ）は復号された動画（静止画）を１コマ単位で所定のフォーマットでアプリケーションソフトに渡す。巻戻し等の再生制御はIMediaControlやIMediaSeeking インターフェイスにより行う。巻戻しはSet_Rateメソッドで負の値を与えれば良いが、サポートしていない場合は、SetPositionsメソッドで再生コマを１つずつ指定する。またIBasicVideo インターフェイスのGetCurrentImageメソッドやMultimedia Streaming API等を使うことで書出しフィルタと同様の機能を実現できる。

[0046] 従来同様の手順２での処理は次のようになる。つまり、アプリケーションソフトは、画像指定ボタン３０６が押されたことを画面操作検知部２２５から通知されると、書出しフィルタ等から受取った１コマ分の画像（動画３０２として表示されている画像である）に、受取った顔領域の座標に応じた枠

の画像を重畳して、表示 1 / F 2 6 4 へ描画処理を行う。これによりキー原画像 3 0 5 として、動画 3 0 2 に表示されているものと同じ画像が表示される。

[0047] 本例の手順 3 での処理は次のようになる。つまり、アプリケーションソフトは、検索キー画像指定領域 3 0 4 内でマウス操作があったことを画面操作検知部 2 2 5 から通知されると、マウス操作の座標と、顔領域の枠を表示した座標とを比較し、該当する枠があるか判断する。該当する枠があれば、その枠が選択されたものとしてその枠に対応する顔領域の通し番号と座標（と特微量）を画像 ID とともに配列に保持する。これにより、選択された枠内の部分画像が、指定済キー画像となる。

次に IVideoFrameStep インターフェイスの Step メソッドによりコマを 1 つ進め、画像、その画像での各顔領域の通し番号と座標（と特微量）を入力フィルタ等から受取り、上記手順 2 と同様にキー原画像 3 0 5 として表示する。

[0048] なお、本例では操作を極力シンプルにするため、指定済キー画像の一覧や修正（原画像で枠を選択し直すこと）、削除等のインタフェースを明示的には提供しないが、エキスパート向けにこれらを提供することを妨げるものではない。

実施例 2

[0049] 本発明の実施形態である類似画像検索システムの実施例 2 を、図 6 を参照して説明する。なお、実施例 1 と同様の部分は説明を省略する。

図 6 は、実施例 2 における端末装置 1 0 3 で複数キー画像検索を行う際の検索画面 3 3 0 を示す図である。本例の検索画面 3 3 0 は、次ボタン 3 2 1 に代えて近傍全表示ボタン 3 3 1 を備え、また自動的に追加されたキー原画像に対し顔の選択を行う指定済キー表示領域 3 3 2 を更に備えた点などで、実施例 1 と異なる。

[0050] 本例の検索画面 3 3 0 を用いて、ユーザが動画の 1 コマずつを複数のキー画像として指定し検索するための手順は、おおよそ以下のようになる。

まず、実施例 1 同様に、従来の手順 1 と手順 2 を行う。つまり、目的の人

物が映った所望の画像（ただし、連続する画像中の最先のものでなくても良い）を再生画像表示領域301に表示させた状態で、画像指定ボタン306を押す。

[0051] 次に、手順3として、近傍全表示ボタン331を押す。すると手順2で画像指定ボタン306を押した時の画像の近傍（例えば5秒前～5秒後までの10秒間）の動画302のうち1つ以上の顔検出結果を有する画像が、指定済キー表示領域332に全て表示される。図6には3つのキー原画像333、334、335が表示された例を示している。端末装置103は、キー原画像333～335に、その画像内での顔領域の通し番号と座標に基づいて、顔領域の境界に相当する枠を重畳して表示する。

[0052] 次に、手順4として、指定済キー表示領域332に表示された画像1つずつについて、実施例1の手順3と同様に、目的の人物の顔を選択する。顔を選択すると、その顔に付された枠が、選択状態を示す枠（例えば太枠）に描画しなおされ、また画像IDと指定した顔の領域の情報が端末装置103内に保持される。

なお、目的の人物の顔がない画像は、そのままの顔も選択せずによい。また一旦顔選択した画像において、別の顔をクリックすると、その新たにクリックした顔が選択状態、元の顔は非選択状態となり、また新たな顔の領域の情報が保持内容に上書きされる。これにより、類似人物検索ボタン318を押したときに指定済キー表示領域332内で選択状態となっていた顔が、キー画像（キー顔）となる。

[0053] 本例によれば、従来の手順2と手順4の繰り返しを省くことができる。

本例の手順3の処理の詳細を以下、補足する。

アプリケーションソフトは、近傍全表示ボタン331の押下があったことを画面操作検知部225から通知されると、SetPositionsメソッドにより再生位置を例えば5秒前に戻す。またフィルタグラフを操作し、出力ピンをNullレンダラに接続する。また再生レートを可能な限り最高にする。それにより順次、コマの画像（GDIビットマップオブジェクト）、その画像での顔領

域の通し番号と座標を受取り、メモリに保持すると共に指定済キー表示領域 332 に並べて表示する。画像をメモリに保持するのは枠の再選択や指定済キー表示領域 332 のスクロールで再描画が必要になるからである。もしそうしたほうが処理が速くなるのなら、元のサイズではなく指定済キー表示領域 332 での表示サイズに縮小して保存してよい。画像 ID、顔領域の通し番号と座標は配列などに格納する。

そして 10 秒分の画像が取り込まれると、再生を停止し、出力ピンを元のレンダラに戻す。

Nullレンダラに接続したことで、この間、再生画像表示領域 301 への動画 301 の表示は更新されない。また通常の DirectShow フィルタは上流フィルタのスレッド内で動くので、アプリケーションのスレッドとデッドロックを起こしやすいので、Multimedia Streaming API を使用して、フィルタグラフから画像データを取り出すことが望ましい。

[0054] なお本例では、指定済キー表示領域 332 にキー原画像（の候補）が多数表示され、ユーザはその中で適宜スクロールしながら、キーにしたい顔を 1 つずつ選択する。そのため、顔が確認し易いよう、検索キー画像指定領域 304 に表示する場合と同じサイズで表示するようにしてある。なおキー原画像は、デフォルトを縮小表示（アイコン）としマウス 226 をそれに合わせることで元のサイズに表示させてもよく、顔を選択し終えた画像から順次、非表示にしても良い。

実施例 3

[0055] 本発明の実施形態である類似画像検索システムの実施例 3 を、図 7 A、図 7 B を参照して説明する。なお、実施例 1 や 2 と同様の部分は説明を省略する。

図 7 A は、実施例 3 の端末装置で複数キー画像検索を行う際の検索画面 340 を示す図である。本例は、指定済キー表示領域 332 内でのキー画像の指定を更に自動化した点などで、実施例 2 と異なる。

[0056] 本例の検索画面 340 を用いて、ユーザが動画の 1 コマずつを複数のキー

画像として指定するための手順は、おおよそ以下のようになる。

まず、実施例 2 同様に、従来の手順 1 と手順 2 を行う。つまり、目的の人物が映った所望の画像を再生画像表示領域 301 に表示させた状態で、画像指定ボタン 306 を押す。

[0057] 次に、手順 3 として、実施例 1 同様に、キー原画像表示部 305 に表示されている原画像中で、顔を示す矩形の中から、検索キーにしたい顔をマウス 226 で指定する。

次に、手順 4 として、近傍全表示ボタン 331 を押す。すると、手順 3 で指定した顔の人物について、手順 3 で指定した画像の時間的近傍において人物追跡処理を行い、その結果同一人物とされた顔を含む画像が、時刻順に全て表示される。更に顔検出結果を示す枠のうち、同一人物とされた顔の枠が、自動的に選択状態となる。

[0058] 次に、手順 5 として、指定済キー表示領域 332 に表示されたキー原画像において、間違って別の人物の顔が選択されていないか確認し、もし間違いがあれば、正しい顔を選択しなおす。それにより、図 7 B に示すようにその顔に付された枠が、選択状態を示す枠（例えば太枠）に描画しなおされ、新たな顔の領域の情報が保持内容に上書きされる。これにより、類似人物検索ボタン 318 を押したときに指定済キー表示領域 332 内で選択状態となっていた顔が、キー画像（キー顔）となる。

[0059] ここで、本例の手順 4 の詳細を補足する。アプリケーションソフトは、近傍全表示ボタン 331 の押下の通知を受けると、キー原画像 305 において選択されている枠があるか否か検査する。選択されている枠があれば、その情報（画像 ID、通し番号、座標、特徴量）を配列の先頭などに保存する。そして SetPositions メソッド等を用い 1 つ前のコマを指定し、画像とその画像での顔領域の通し番号と座標（と特徴量）を入力フィルタから受取る。もし顔領域が 1 つもなければ、そのデータは破棄してさらに 1 つ前のコマを指定する。もし顔領域が 1 つ以上あれば、すでに配列（の最後）に保持された顔領域の座標（重心座標）に最も近い座標の顔領域を見つけ、領域の位置（

重心座標）や大きさが、所定の連続性条件を満たすか（例えばカルマンフィルタの出力との差が所定値以内か）検査する。連続性条件を満たす場合、同一人物と推定されるので、それを配列に保存すると共に、更に前のコマを指定し、同様の処理を同一人物を見失う（ロスト）まで繰り返す。なお複数コマ連続して連続性条件を満たせなかったときにロストしたと判断する。また近傍全表示ボタン 331 の押下時に表示されていたキー原画像 305 より時間的に後の画像についても同様にコマを進めながら繰り返す。

なお、特徴量も取得しているので、配列の最初に保持された特徴量（ユーザにより手順 3 で指定された顔の特徴量）と今のコマの各顔領域の特徴量との類似度も用いて検査するほうが、精度が期待できる。

[0060] このように、本例での人物追跡処理断は、端末装置 103 側で顔領域の時空間的連続性にに基づき行うことができ、例えば特許文献 5 に記載の方法を利用できる。更に顔検出の結果以外のものによる追跡を併用することもできる。最も容易なのは、手順 3 で指定した顔の部分画像を用いたテンプレートマッチングである。初期テンプレートには、指定画像の時間的近傍で差分法による動体検知を行い、指定した顔に相当する人物全体の像（複数コマの中から大きさがそれらの中央値に近いものを選ぶと良い）を用いてもよい。また人物全体の像から顔以外の体の部分を切出し、その体画像から抽出した特徴量の類似度を顔の類似度と併用してもよい。なお画像処理による動体追跡手法は様々なものが知られており、本例に適用できる。複数の追跡手法を併用し、ある手法でロストしても他の手法での追跡結果をロストした手法に与え続けることで再度補足することが期待でき、全手法でロストするまで追跡できることとなり、堅牢性が向上する。或いはレーザセンサ等、距離或いは 3 次元形状を直接計測する手段をカメラに併設し、その計測結果から物体検出し追跡してもよい。時空間連続性における空間とは、2 次元でもよく、ピクセル座標のような画像空間に限らず、経緯度のような地理的空間でもよい。複数のカメラ等の位置検出手段の計測結果を共通の座標系で扱うことで、個々のカメラを意識することなく追跡を行うことができる。

[0061] 以上説明したように、本発明の実施形態は、時系列画像（動画）に連続的に映り込んでいることが期待できる所望の人物を、一括して検索キー画像に指定する場合に、好適である。なお検索対象は記録画像に限らず、撮影中の画像からリアルタイムに検索（照合）するシステムにも適用可能である。また動画は再生画像表示領域から画像指定するものに限らず、ファイル読込ボタン307を押したときのダイアログボックスにおいて複数の静止画ファイル（時系列に撮影されたもの）を選択することで、それらをキー原画像に一括指定できるようにしても良い。

ここで、本発明に係るシステムや装置などの構成としては、必ずしも以上に示したものに限られず、種々な構成が用いられてもよい。また、本発明は、例えば、本発明に係る処理を実行する方法或いは方式や、このような方法や方式をコンピュータに実現させるためのプログラムや当該プログラムを記録する有体の媒体などとして提供することもできる。

産業上の利用可能性

[0062] CCTV（Closed-Circuit Television）システム、顔認証システム、犯罪者データベース等のほか、テレビ番組制作システムや個人向けの電子アルバム等、カメラに映った人物や車両などの映像を扱うシステムに利用可能である。

符号の説明

[0063] 102：録画装置、 103：端末装置、 113：複数検索キー選択部、
200：ネットワーク、 201：撮像装置、 210：画像送受信部、
211：画像記録部、 212：再生制御部、 213：人物領域検出部、
214：人物特徴量抽出部、 215：人物特徴量記録部、 216：属性情報記録部、
217：要求受信部、 218：登場人物検索部、 219：登場イベント検索部、
220：検索結果送信部、
240：バス、 241：撮像部、 242：主記憶部、 243：符号化部、 245：ネットワークI/F、

250 : バス、 251 : CPU、 252 : 主記憶部、 253 : 補助記憶部、 254 : ネットワーク I / F、

260 : バス、 261 : CPU、 262 : 主記憶部、 263 : 補助記憶部、 264 : 表示 I / F、 266 : ネットワーク I / F、 270 : 表示装置、 280 : キーボード、 282 : マウス、

300, 330, 340 : 検索画面、 301 : 再生画像表示領域、 302 : 動画、 303 : 画像再生操作領域、 304 : 検索キー画像指定領域、 305 : 検索キー画像、 306 : 画像指定ボタン、 307 : ファイル読込ボタン、 308 : 検索絞込パラメータ指定領域、 309 ~ 312 : 撮像装置指定チェックボックス、 313, 314 : 時刻指定チェックボックス、 315, 316 : 時刻指定欄、 317 : 検索実行領域、 318 : 類似人物検索ボタン、 319 : 登場イベント検索ボタン、 320 : 検索結果表示領域、

331 : 全近傍表示ボタン、 332 : 指定済キー表示領域、 333 ~ 335 : キー原画像。

請求の範囲

- [請求項1] カメラで撮影し記録した動画像から、所定の被写体が映った部分画像をキーとして指定し、指定された該部分画像の特徴量に近い特徴量を有する画像を検索する類似画像検索装置において、
- 動画像の中から所望の1コマを指定するためのプレビューを提供する手段と、
- 前記プレビューされた1コマをキー原画像に指定する操作を受付ける手段と、
- 前記受付ける手段で指定された1コマに時間的に近傍の複数コマを、自動的にキー原画像に追加指定する手段と、
- 前記受付ける手段或いは前記追加指定する手段でキー原画像に指定された1ないし複数コマを、該1ないし複数コマから所定のアルゴリズムで検出された被写体に対応する領域を示す図形を付加して表示する手段と、
- 付加された前記図形の初期状態を非選択状態とし、1つのコマに対していずれか1つの図形を選択状態とする操作を受付ける選択手段と、
- 選択状態となった前記図形に対応する複数の被写体をキーとする検索を要求する手段と、を有する類似画像検索装置。
- [請求項2] 前記被写体は人の顔であり、
- 前記特徴量は、前記動画像から自動的に検出された前記被写体ごとに予め抽出され、該自動的に検出された元の画像に対応付けて記録されており、
- 前記要求する手段からの要求を受けて、前記選択状態となった前記図形に対応する複数の被写体は同一人物であると看做して、該複数の被写体に対応する複数の特徴量を1つずつキーとして検索し、それら結果を結合して応答する検索実行手段を設けたことを特徴とする請求項1に記載の類似画像検索システム。
- [請求項3] 前記表示する手段は、受付ける手段で指定された1コマに対し、前

記図形を付加して、キー原画像表示領域に表示するものであり、

前記選択手段がキー原画像表示領域に表示された前記図形に対する前記選択状態とする操作を受け付けるか、或いは、所定のボタンの1回の押下のみに応じて、前記追加指定する手段が、そのとき前記キー原画像表示領域に表示されていた1コマに続く次のコマをキー原画像に追加指定することで、自動的に該次のコマが前記キー原画像表示領域に表示され、再び前記選択手段が操作を受け付けるという動作を繰り返し、

要求する手段は、該繰り返しの間に選択状態にされた複数の前記図形に対応する複数の被写体を前記キーとすることを特徴とする請求項1又は2に記載の類似画像検索装置。

[請求項4]

追加指定する手段は、前記受け付ける手段により前記1コマをキー原画像に指定する操作が受け付けられた後、自動的に或いは所定のボタンの1回の押下のみに応じて、前記受け付ける手段で指定された1コマに時間的に前または後に連続する複数コマのうち、少なくとも被写体が検出された複数コマをキー原画像に追加指定し、

前記表示する手段は、前記追加指定する手段で指定された複数のキー原画像を、夫々に前記図形を付加して指定済キー表示領域に表示するものであり、

前記選択手段が、前記指定済キー表示領域に表示された複数のキー原画像に対する、前記図形を前記選択状態とする操作を複数受け、

前記要求する手段が、前記選択状態にされた複数の前記図形に対応する複数の被写体を前記キーとすることを特徴とする請求項1又は2に記載の類似画像検索装置。

[請求項5]

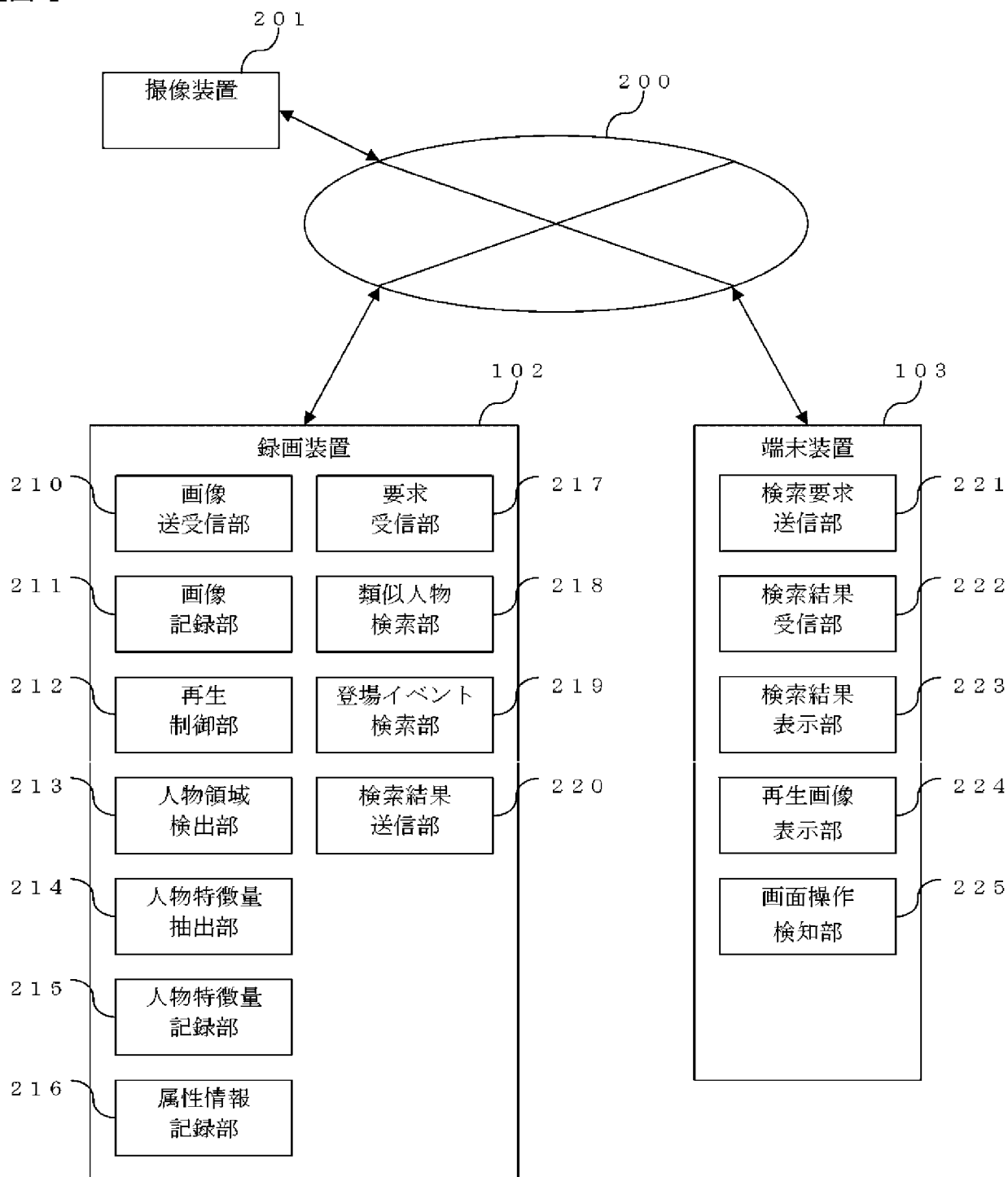
前記表示する手段は、受け付ける手段で指定された1コマに対し、前記図形を付加して、キー原画像表示領域に表示するものであり、

前記自動追加手段は、前記選択手段が前記キー原画像表示領域に表示された前記図形に対する前記選択状態とする操作を受け付けた後の所

定のボタンの 1 回の押下のみに呼応して、該選択状態にされた前記図形に対応する被写体を起点として、記録された動画像の中から当該被写体に時空間的連続性を満たす被写体を追跡し、当該追跡が成功した範囲の複数のコマをキー原画像に追加指定するものであり、

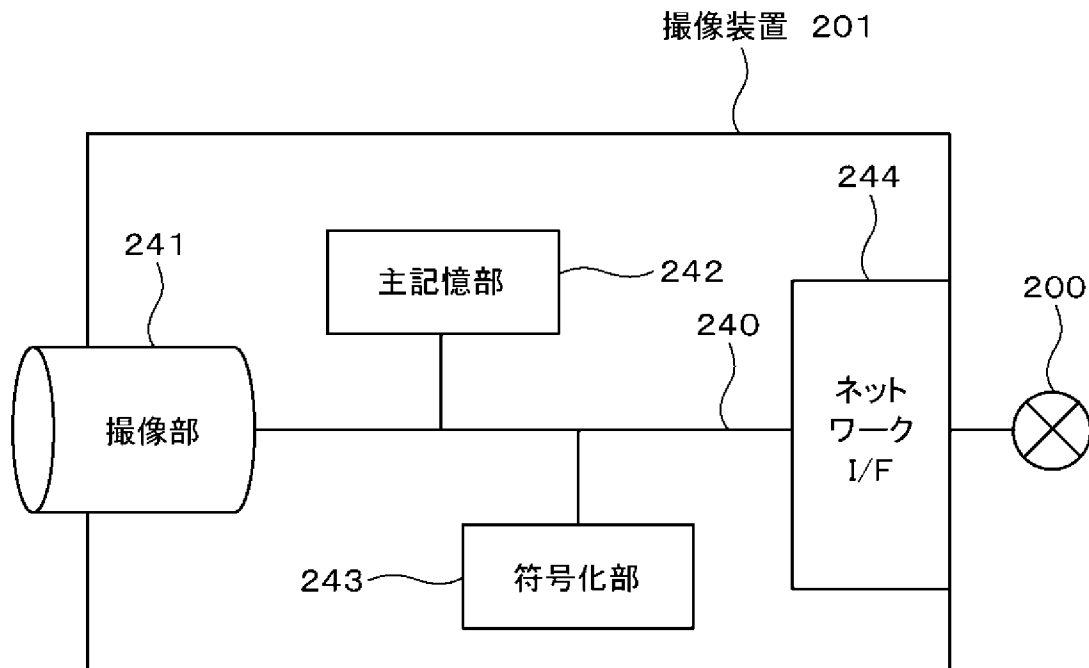
前記表示する手段は、前記追加指定する手段で指定された複数のキー原画像を、夫々に前記図形を付加して指定済キー表示領域に表示することを特徴とする請求項 2 又は 4 に記載の類似画像検索装置。

[図1]

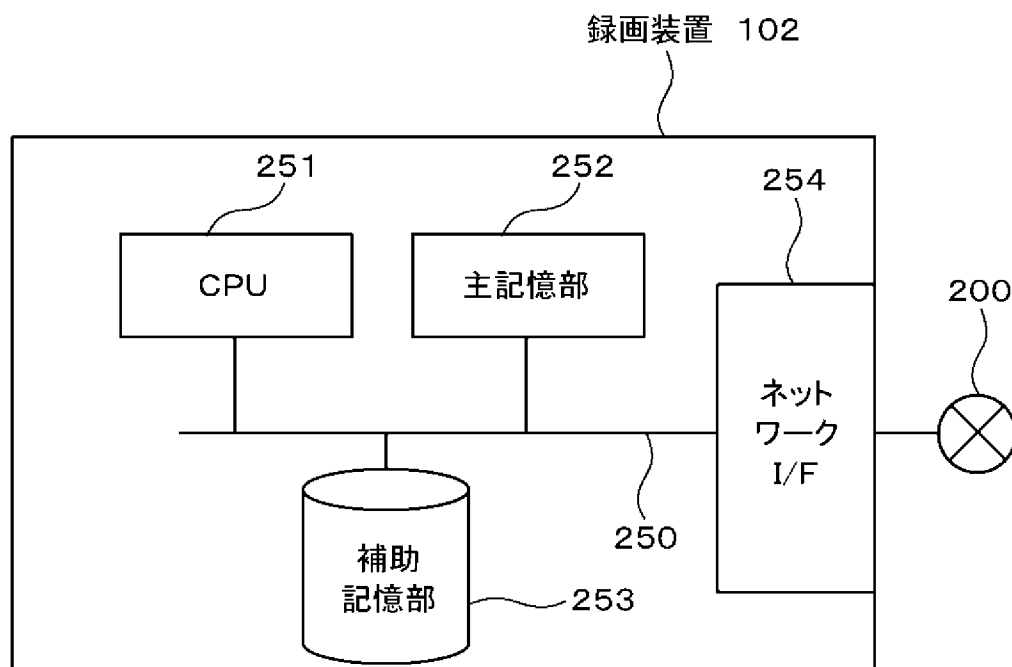


[図2]

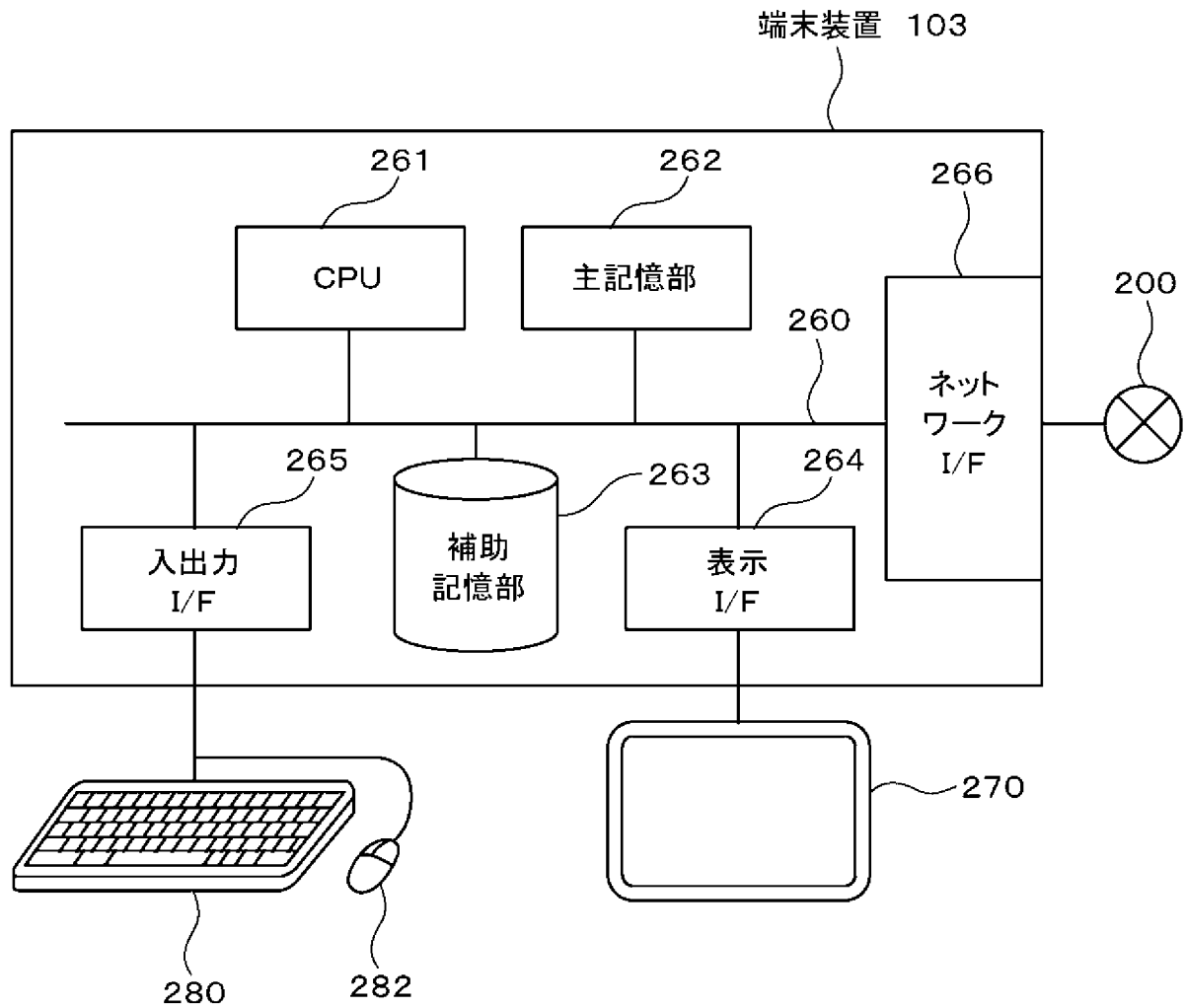
図 2



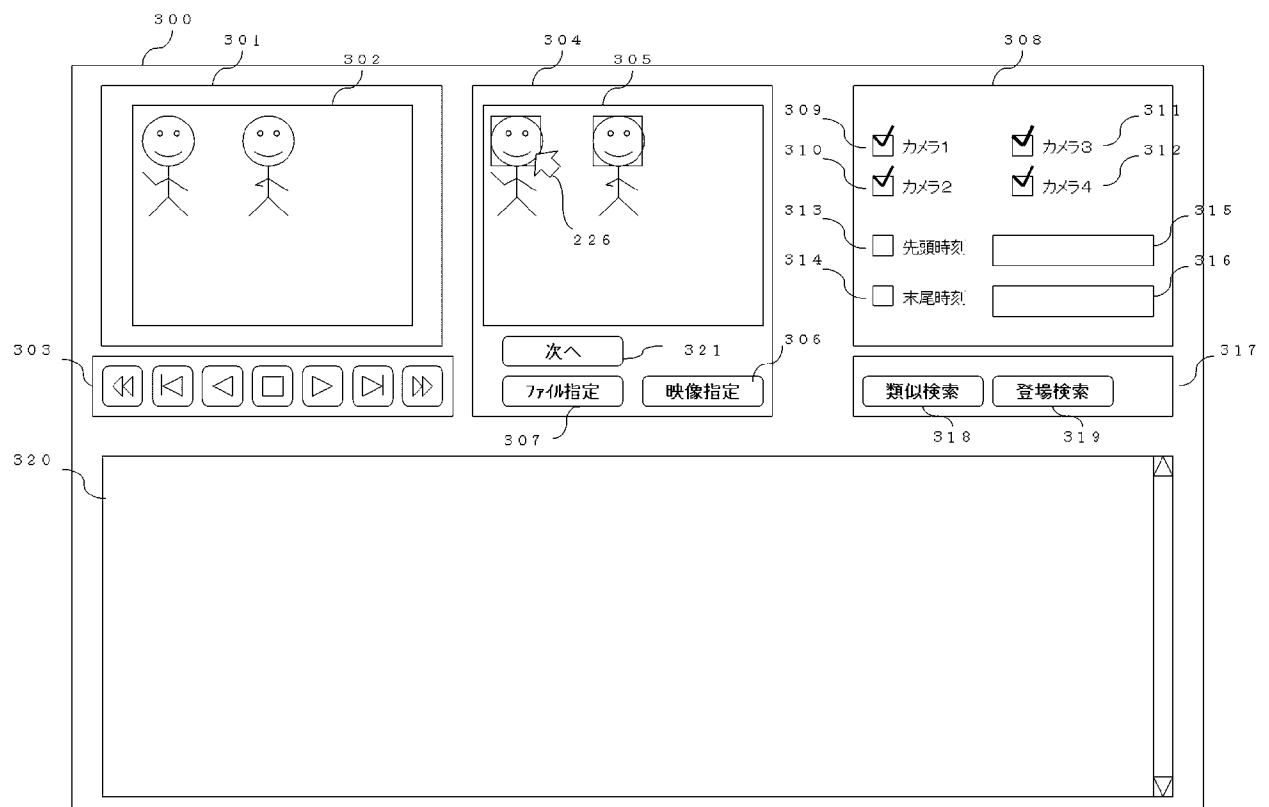
[図3]



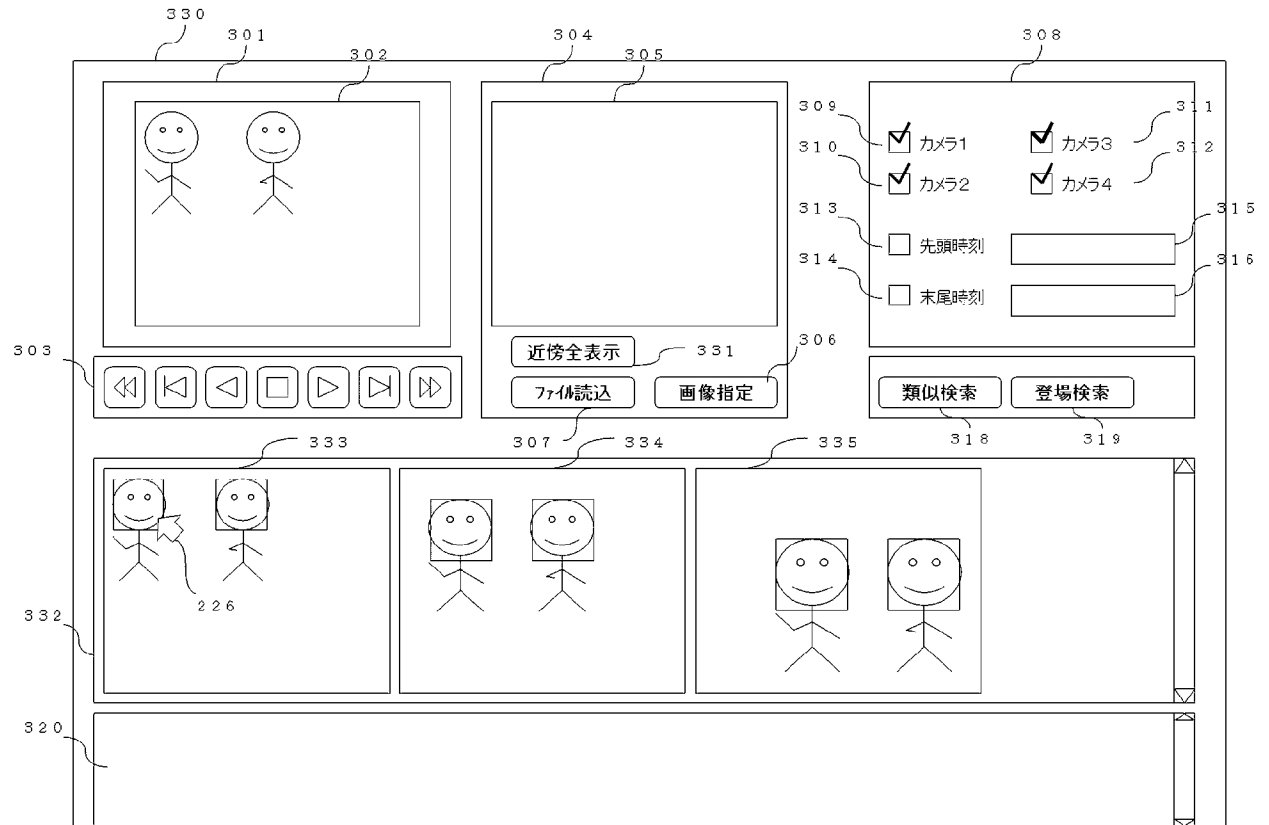
[図4]



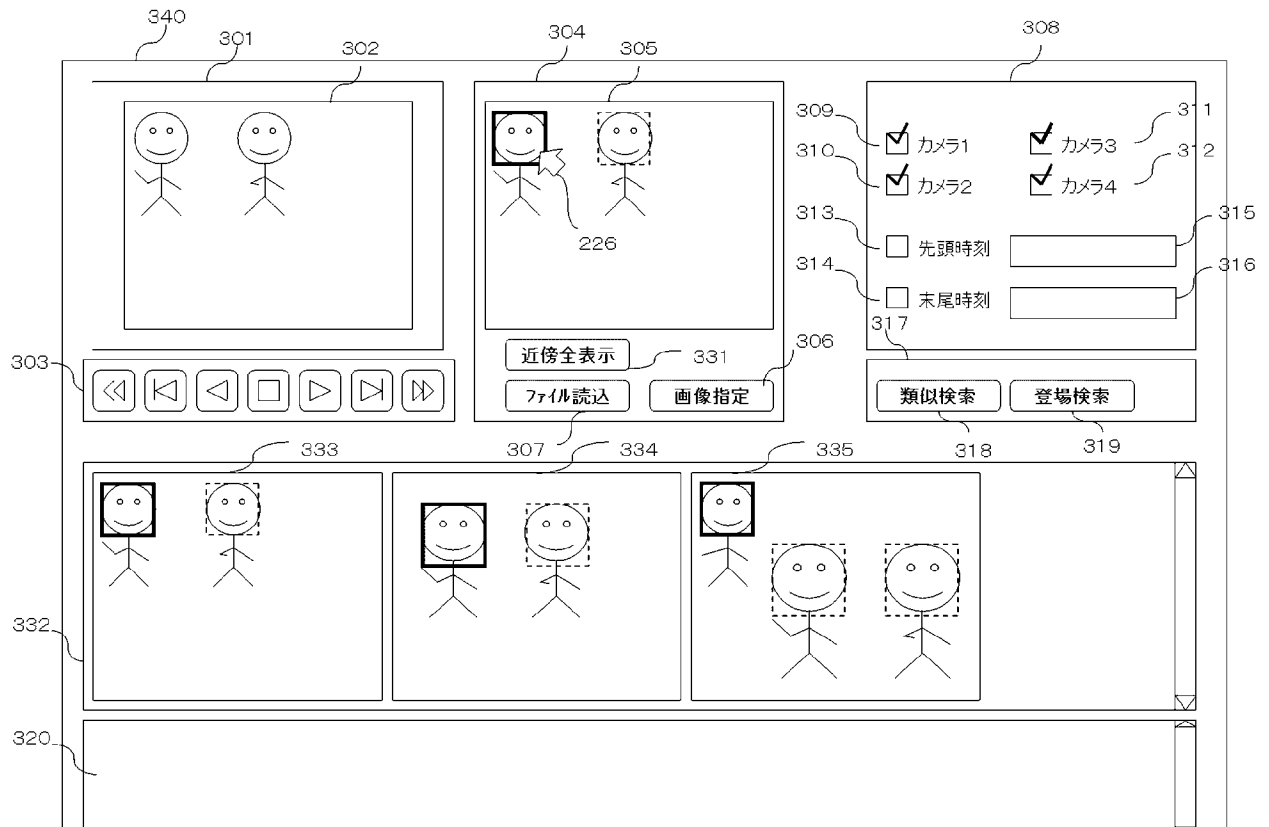
[図5]



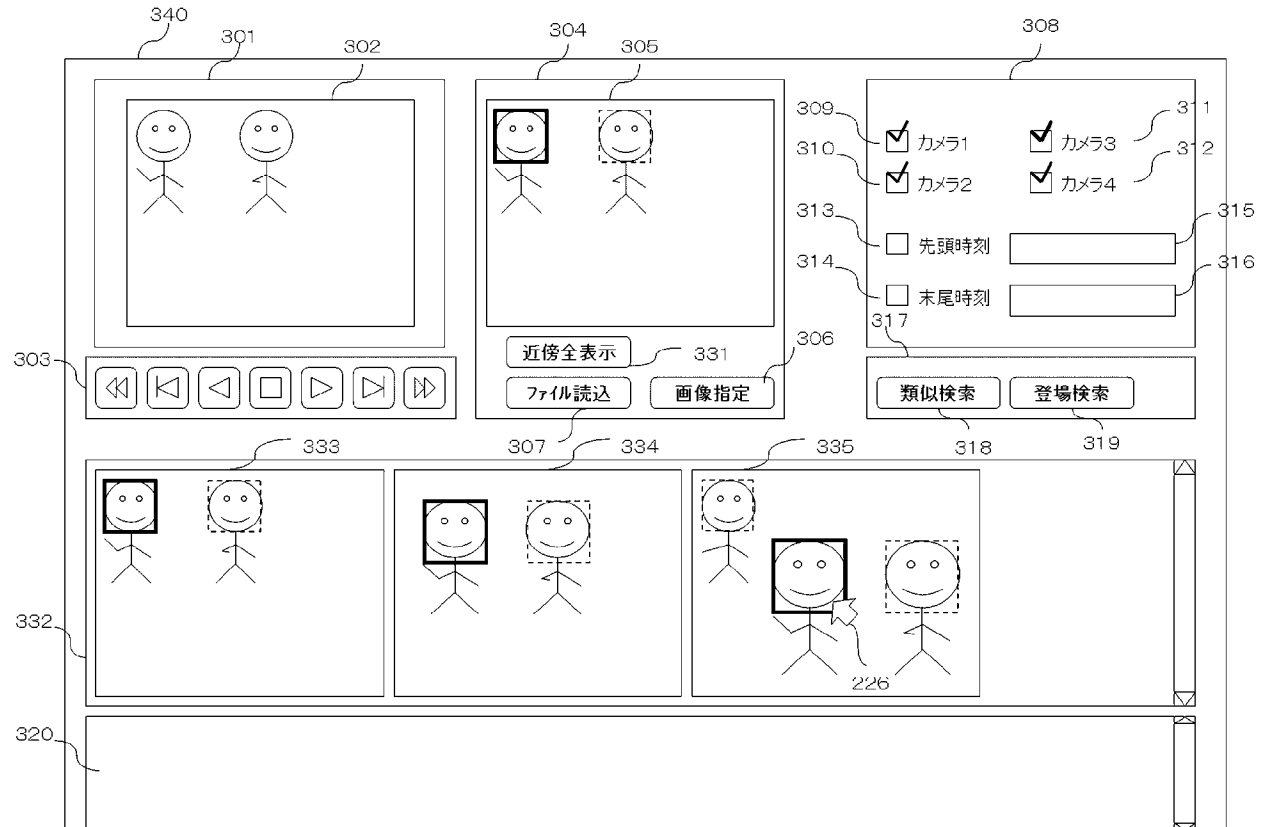
[図6]



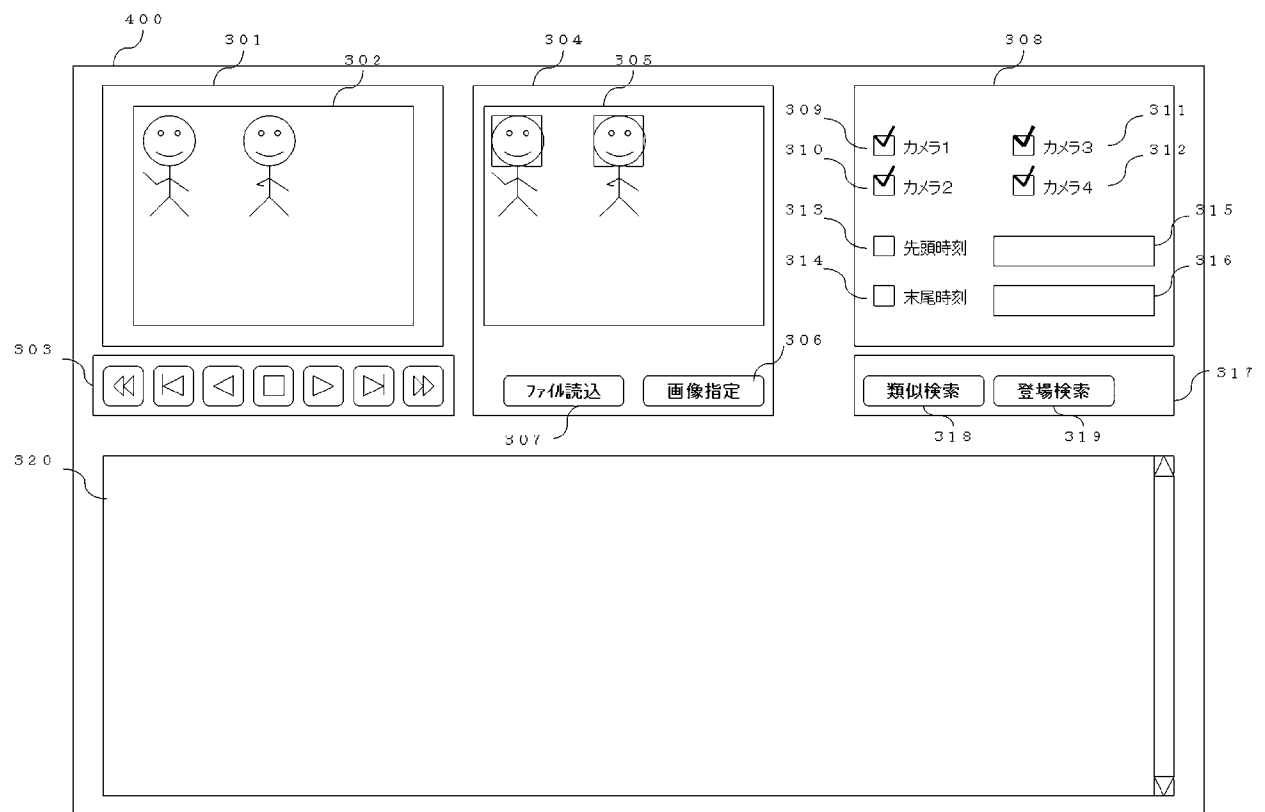
[図7A]



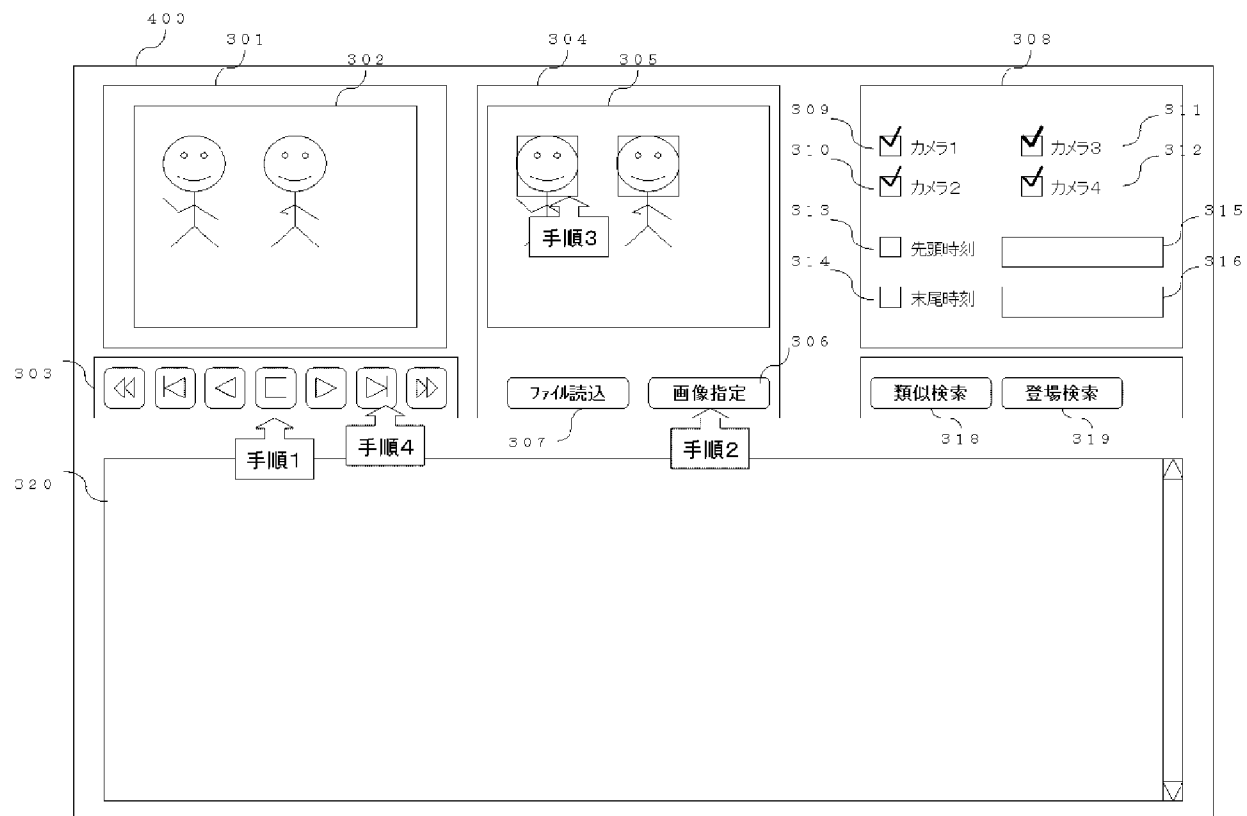
[図7B]



[図8]



[図9]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/074507

A. CLASSIFICATION OF SUBJECT MATTER

G06F17/30(2006.01)i, H04N5/76(2006.01)i, H04N5/91(2006.01)i

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)

G06F17/30, H04N5/76, H04N5/91

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	1996-2013
Kokai Jitsuyo Shinan Koho	1971-2013	Toroku Jitsuyo Shinan Koho	1994-2013

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.
Y	JP 2009-27393 A (Hitachi, Ltd.), 05 February 2009 (05.02.2009), paragraphs [0012] to [0028]; fig. 3, 5, 7 (Family: none)	1-5
Y	JP 2010-226687 A (Sony Corp.), 07 October 2010 (07.10.2010), paragraphs [0023] to [0026]; fig. 3 & US 2010/0220194 A1 & CN 101820530 A	1-5
Y A	JP 2011-48668 A (Hitachi Kokusai Electric Inc.), 10 March 2011 (10.03.2011), paragraphs [0049] to [0054] & US 2011/0052069 A1	2-5 1

☐ 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
25 September, 2013 (25.09.13)

Date of mailing of the international search report
08 October, 2013 (08.10.13)

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（I P C））

Int.Cl. G06F17/30(2006.01)i, H04N5/76(2006.01)i, H04N5/91(2006.01)i

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（I P C））

Int.Cl. G06F17/30, H04N5/76, H04N5/91

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	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 3 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 3 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 3 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 請求項の番号
Y	JP 2009-27393 A（株式会社日立製作所） 2009.02.05, 【0012】～【0028】，第3,5,7図（ファミリーなし）	1-5
Y	JP 2010-226687 A（ソニー株式会社） 2010.10.07, 【0023】～【0026】，第3図 & US 2010/0220194 A1 & CN 101820530 A	1-5

☒ 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

2 5 . 0 9 . 2 0 1 3

国際調査報告の発送日

0 8 . 1 0 . 2 0 1 3

国際調査機関の名称及びあて先

日本国特許庁（I S A / 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）

野崎 大進

5 M

5 3 7 7

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 9 9

C (続き) . 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 請求項の番号
Y A	JP 2011-48668 A (株式会社日立国際電気) 2011.03.10, 【0049】～【0054】 & US 2011/0052069 A1	2-5 1