

(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



(10) 国際公開番号

WO 2013/186962 A1

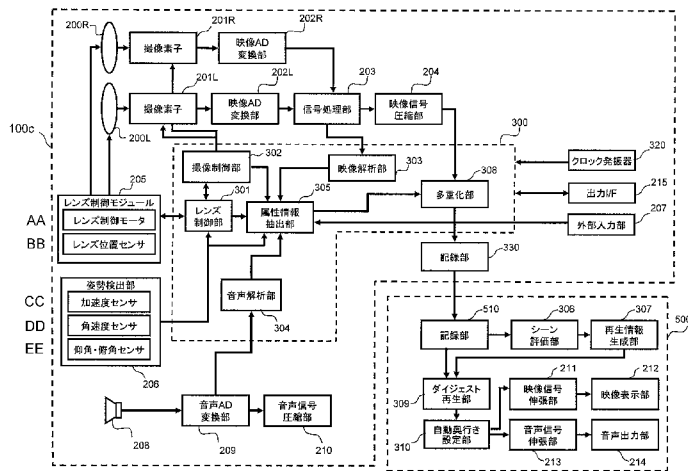
(43) 国際公開日
2013年12月19日(19.12.2013)

- (51) 国際特許分類:
H04N 13/04 (2006.01) H04N 5/91 (2006.01)
H04N 5/225 (2006.01) H04N 5/92 (2006.01)
H04N 5/232 (2006.01) H04N 5/93 (2006.01)
- (21) 国際出願番号: PCT/JP2013/001069
- (22) 国際出願日: 2013年2月25日(25.02.2013)
- (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 2012-132252 2012年6月11日(11.06.2012) JP
- (71) 出願人: パナソニック株式会社 (PANASONIC CORPORATION) [JP/JP]; 〒5718501 大阪府門真市大字門真1006番地 Osaka (JP).
- (72) 発明者: 森岡 芳宏 (MORIOKA, Yoshihiro), 山内 栄二 (YAMAUCHI, Eiji), 松浦 賢司 (MATSUURA, Kenji).
- (74) 代理人: 奥田 誠司 (OKUDA, Seiji); 〒5410041 大阪府大阪市中央区北浜一丁目8番16号 大阪証券取引所ビル10階 奥田国際特許事務所 Osaka (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, KM, 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, 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, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT,

[続葉有]

(54) Title: VIDEO PROCESSING DEVICE, IMAGING DEVICE, AND PROGRAM

(54) 発明の名称: 映像処理装置、撮影装置、およびプログラム



- 201L, 201R Imaging element
- 202L, 202R Video AD converting unit
- 203 Signal processing unit
- 204 Video signal compression unit
- 205 Lens control module
- AA Lens control motor
- BB Lens position sensor
- 206 Orientation detecting unit
- CC Acceleration sensor
- DD Angular velocity sensor
- EE Angle of elevation/angle of depression sensor
- 207 External input unit
- 209 Audio AD converting unit
- 210 Audio signal compression unit
- 211 Video signal expansion unit
- 212 Video display unit
- 213 Audio signal expansion unit
- 214 Audio signal output unit
- 215 Output I/F
- 301 Lens control unit
- 302 Imaging control unit
- 303 Video analysis unit
- 304 Audio analysis unit
- 305 Attribute information extracting unit
- 306 Scene evaluating unit
- 307 Playback information generating unit
- 308 Multiplexing unit
- 309 Digest generating unit
- 310 Automatic depth setting unit
- 320 Clock oscillator
- 330, 510 Recording unit

(57) Abstract: A video processing device according to an embodiment of the present invention is provided with an interface that acquires information specifying a plurality of scenes, to be used for digest playback, extracted from stereoscopic video acquired by means of imaging, and information indicating an amount of disparity for an object included in the plurality of scenes. The video processing device is also provided with a digest video generating unit that corrects, on the basis of the information specifying the plurality of scenes and the information indicating the amount of disparity, the amount of disparity in a plurality of successive image frames that include image frames immediately before or immediately after the boundary between two successive scenes from among the plurality of scenes, and then generates a digest video.

(57) 要約: ある実施形態における映像処理装置は、撮影によって取得された立体映像の中からダイジェスト再生のために抽出された複数のシーンを特定する情報、および前記複数のシーンに含まれる被写体の視差量を示す情報を取得するインタフェースと、前記複数のシーンを特定する情報および前記視差量を示す情報に基づいて、前記複数のシーンのうち、連続する2つのシーンの境界の直前または直後の画像フレームを含む連続する複数の画像フレームにおける視差量を補正してダイジェスト映像を生成するダイジェスト映像生成部とを備える。

WO 2013/186962 A1

NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI 添付公開書類:
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, — 國際調查報告 (條約第 21 條(3))
NE, SN, TD, TG).

明 細 書

発明の名称：映像処理装置、撮影装置、およびプログラム

技術分野

[0001] 本願は、映像の中からその一部を抜き出すことによってダイジェスト映像を生成する技術に関する。

背景技術

[0002] 従来より、映像の中からその一部を抜き出してダイジェスト（要約）映像を生成する技術が知られている。撮影された映像について、その全内容を再生するためには少なからず時間を要する。そのため、映像の概要を短時間で把握したいというニーズがある。また、映像の中から重要な部分だけを抽出して視聴したいというニーズもある。

[0003] このようなニーズに応えるべく、特許文献1には、映像のメタデータ（属性情報）に基づいてシーンを評価し、その評価結果に基づいて、映像のシーンやクリップの数を絞り込んだダイジェスト映像を生成する撮影装置が開示されている。

[0004] また、視聴者の嗜好により、好ましいダイジェスト映像の内容が大きく異なることに着目し、ユーザの多様な嗜好に対応したダイジェスト再生を実現できる撮影装置が特許文献2に開示されている。この撮影装置は、シーン抽出の判断基準となる属性情報をユーザが任意に入力できるようにすることにより、個々のユーザに対応したダイジェスト映像の作成を可能にしている。

[0005] 一方、立体（3D）映像を撮影可能な撮影装置において、上記のようなダイジェスト再生を可能にする技術も存在する。例えば特許文献3には、被写体距離情報を撮影画像の属性情報として記録し、当該距離情報を用いて、ハイライトシーンを選択する撮影装置が開示されている。

先行技術文献

特許文献

[0006] 特許文献1：特開2008-227860号公報

特許文献2：国際公開第2011/099299号

特許文献3：特開2011-15256号公報

特許文献4：国際公開第2012/029298号

発明の概要

発明が解決しようとする課題

[0007] 本開示は、従来よりも好ましいダイジェスト映像を生成するための技術を提供する。

課題を解決するための手段

[0008] 本開示の一態様による映像処理装置は、撮影によって取得された立体映像の中からダイジェスト再生に用いられる複数のシーンを特定する情報、および前記複数のシーンに含まれる被写体の視差量を示す情報を取得するインタフェースと、前記複数のシーンを特定する情報および前記視差量を示す情報に基づいて、前記複数のシーンのうち、連続する2つのシーンの境界の直前または直後の画像フレームを含む連続する複数の画像フレームにおける視差量を補正してダイジェスト映像を生成するダイジェスト映像生成部とを備えている。

発明の効果

[0009] 本開示の実施形態によれば、シーンの繋ぎ部の前後で過度な飛び出し量の変化が生じない立体映像のダイジェスト映像を生成することができる。また、他の実施形態では、臨場感を増すために飛び出し量を一時的に増加させるような演出を伴うダイジェスト映像を生成することも可能である。このため、視聴者の眼精疲労や頭痛等の発生を抑え、立体映像の安全性を確保することができ、さらに、迫力のあるダイジェスト映像を生成することもできる。

図面の簡単な説明

[0010] [図1]実施形態1におけるビデオカメラの外観図

[図2]実施形態1におけるビデオカメラの概略構成を示すブロック図

[図3]実施形態1における撮影した動画像のクリップ、シーン、フレームの関

係を示す図

[図4]実施形態1におけるシーンを識別する情報を示す図

[図5]実施形態1における属性情報と評価値との対応関係を示す図

[図6]実施形態1における属性情報と評価値との対応関係に基づいて、各シーンを評価した例を示す図

[図7A]実施形態1における人物を撮影した映像の第1の例を示す図

[図7B]実施形態1における人物を撮影した映像の第2の例を示す図

[図7C]実施形態1における人物を撮影した映像の第3の例を示す図

[図7D]実施形態1における人物を撮影した映像の第4の例を示す図

[図8]実施形態1における再生すべきシーンを抽出した再生情報の例を示す図

[図9]実施形態1における撮影した映像から再生情報を生成する動作を示すフローチャート

[図10A]実施形態1における生成された撮影情報に基づいてダイジェスト再生を行う動作を示すフローチャート

[図10B]実施形態1における左右両映像およびそれらの視差量を示すデプスマップの例を示す図

[図10C]実施形態1における補正後の左右両映像およびそれらの視差量を示すデプスマップの例を示す図

[図10D]実施形態1における視差量の補正対象となる画像フレームの選定の例を示す図

[図10E]実施形態1におけるトランジション映像の挿入の例を示す図

[図11]実施形態1における属性情報と評価値との対応関係の他の例を示す図

[図12]実施形態1における属性情報と評価値との対応関係に基づいて、各シーンを評価した他の例を示す図

[図13]実施形態1における再生すべきシーンを抽出した再生情報の他の例を示す図

[図14]実施形態1における属性・評価対応データをユーザが選択する選択画面の例を示す図

[図15]実施形態2におけるビデオカメラ100bの概略構成を示すブロック図

[図16]実施形態2における撮影した映像、音声、および多重化データを記録する動作を示すフローチャート

[図17]実施形態2における多重化された属性情報に基づいて、再生情報を生成し、ダイジェスト再生する動作を示すフローチャート

[図18]他の実施形態におけるビデオカメラ100cおよび再生装置500の概略構成を示すブロック図

[図19]他の実施形態におけるビデオカメラ100dおよび再生装置501の概略構成を示すブロック図

[図20A]被写体が比較的遠い場合の撮影の状況を模式的に示す図

[図20B]被写体が比較的近い場合の撮影の状況を模式的に示す図

[図21]他の実施形態における撮影時の処理の例を示すフローチャート

[図22]被写体が比較的近いシーンのみを抽出してダイジェスト映像を生成する処理の例を示す図

[図23]再生時に遠くに写っている被写体を拡大表示する実施形態を示す図

発明を実施するための形態

[0011] 以下、添付の図面を参照しながら、例示的な実施形態を説明する。但し、必要以上に詳細な説明は省略する場合がある。例えば、既によく知られた事項の詳細説明や実質的に同一の構成に対する重複説明を省略する場合がある。これは、以下の説明が不必要に冗長になるのを避け、当業者の理解を容易にするためである。

[0012] なお、発明者らは、当業者が本開示を十分に理解するために添付図面および以下の説明を提供するのであって、これらによって特許請求の範囲に記載の主題を限定することを意図するものではない。

[0013] (実施形態1)

立体映像についてダイジェスト再生を行う場合、シーンの繋ぎ部の前後で主要な被写体である人物の顔などの飛び出し量（視差量）が急激に変化する

ことがある。そのような急激な飛び出し量の変化を伴う映像は、視聴者を驚かせることになるとともに、眼精疲労や頭痛等の障害を引き起こし得るため、安全性に問題がある。

[0014] また、従来技術では、立体映像の安全性を保ちながらも、一時的に立体性を強調するような臨場感のある映像表現を行うことができなかった。

[0015] 本実施形態によれば、シーンの繋ぎ部の前後で視差量が適切に調整された立体映像のダイジェスト映像を生成することができる。

[0016] [1-1. 構成]

図1は、映像を撮影するビデオカメラ（映像撮影装置）100aの外観を示す斜視図である。ビデオカメラ100aは、立体映像を撮影、記録、再生する機能を備えている。ビデオカメラ100aは、さらに、立体映像の中からダイジェスト再生するための複数の部分を抽出し、再生する機能、およびそのようなダイジェスト再生が可能な映像信号を記録する機能も備えている。ここで「立体映像」とは、視差を有する一对の映像の組を意味する。本明細書では、立体映像を示すデータも「立体映像」と呼ぶことがある。一对の映像の組を好適に組み合わせて処理することにより、立体映像を再生することができる。また、「ダイジェスト再生」とは、映像の一部のシーンを複数抽出して連続的に再生することを意味する。

[0017] ビデオカメラ100aは、2種類のレンズ群200R、200Lを備え、これらのレンズ群を用いて撮影を行うことによって立体映像を撮影する。本実施形態では、レンズ群200Lは、レンズ群200Rと比較して小型のレンズである。レンズ群200Rはズームレンズを有しているが、レンズ群200Lはズームレンズを有していない。レンズ群200Lを通して撮影した映像は電子ズーム処理がなされ、レンズ群200Rを通して撮影された映像と同一範囲の領域が抽出される。これらによって立体映像を構成する2つの映像が生成される。

[0018] レンズ群200Rとレンズ群200Lとの距離は、撮影する立体映像の視差の大きさに影響するため、例えば人の左右両眼の距離と同程度に設定され

得る。さらに、レンズ群200Rおよびレンズ群200Lは、ビデオカメラ100aを地面に平行に置いたときに、ほぼ同一水平面上に位置するように配置され得る。これは、人が対象物を見る際には、左右の目がほぼ水平な状態で見ることが一般的であるためである。

[0019] また、レンズ群200Rおよびレンズ群200Lは、その光学中心がビデオカメラ100aにおける撮像素子の撮像面に平行な同一平面上に位置するように配置されて得る。これは、自然な立体映像を取得するために、レンズ群200Rおよびレンズ群200Lが被写体からほぼ同一距離に位置するようにするためである。なお、より厳密には、レンズ群200R、200Lは、それらの後段に配置される撮像素子との位置関係を考慮して設計される。

[0020] これらのレンズ群200R、200Lの相対的な位置関係が理想的な関係に近いほど、撮影された映像から立体映像を生成する際の信号処理量を軽減することができる。レンズ群200R、200Lが撮像面に平行な同一平面上にある場合、立体映像を構成する左右の画像フレームにおける同一被写体の位置が、エピポーラ拘束 (Epipolar Constraint) 条件を満たす。このため、立体映像を生成するための信号処理において、一方の映像面上での被写体の位置が確定すると、他方の映像における被写体の位置も比較的容易に算出することが可能となる。

[0021] なお、レンズ群200Rおよびレンズ群200Lの配置関係が上記のような理想的な関係にない場合であっても、信号処理によって適切な立体映像を生成することが可能であるため、必ずしも上記の条件が満たされている必要はない。

[0022] レンズ群200Rは、ビデオカメラ100aの本体前部に設けられ、レンズ群200Lは、撮影した映像を確認するための映像表示部 (ディスプレイ) 212の背面に設けられている。モニター部104は、被写体がある側とは反対側 (ビデオカメラ100aの後部側) に、撮影した映像を表示する。ビデオカメラ100aは、レンズ群200Rを用いて撮影した映像は右眼視点の映像として、レンズ群200Lを用いて撮影した映像は左眼視点の映像

として処理する。

[0023] なお、本実施形態において、レンズ群200R、200Lを異なる構成とする理由は、レンズ群200Lを簡易な構成にすることによって装置全体の小型化および低コスト化を図るためである。ただし、この点は本開示における技術の本質ではなく、2つのレンズ群200R、200Lが全く同一の構造を有していてもよい。

[0024] 図2は、ビデオカメラ100aの概略構成を示すブロック図である。ビデオカメラ100aは、レンズ群200R、200L、撮像素子（イメージセンサ）201R、201L、映像AD変換部（Analog-to-Digital Converter）202R、202L、信号処理部203、映像信号圧縮部204、レンズ制御モジュール205、姿勢検出部206、外部入力部207、マイクロフォン208、音声AD変換部209、音声信号圧縮部210、映像信号伸張部211、映像表示部212、音声信号伸張部213、音声出力部214、出力I/F（Interface）215、制御部300、クロック発振器320、記録部330を備えている。

[0025] レンズ群200R、200Lは、それぞれに対応する撮像素子201R、201Lの撮像面上に被写体像を形成するために、被写体から入射する光を調整する光学系である。レンズ群200Rは、異なる特性をもつ複数のレンズを有し、レンズ間の距離を変化させることによって焦点距離やズーム倍率（映像の拡大倍率）を変化させることができる。焦点距離やズーム倍率の調整は、ビデオカメラの撮影者によって手動で実行されてもよいし、後述するレンズ制御モジュール205を通じて制御部300によって自動的に実行されるものであってもよい。一方、レンズ群200Lは単焦点レンズであり、焦点距離やズーム倍率の調整の機能を有していない。

[0026] 撮像素子201R、201Lは、レンズ群200R、200Lをそれぞれ通過して入射する光を電気信号に変換する。撮像素子201R、201Lは、典型的にはCCDまたはC-MOSセンサである。撮像素子201R、201Lは、光電変換によって受光量に応じた電気信号を出力する複数の光感

知セルを有している。これにより、撮像面上に形成された像に応じた画像信号が出力される。また、撮像素子201R、201Lは、画像に関する情報以外にも、3原色点の色度空間情報、白色の座標、および3原色のうち少なくとも2つのゲイン情報、色温度情報、 Δuv （デルタ uv ）、および3原色または輝度信号のガンマ情報等の情報を出力するように構成されていてもよい。撮像素子201R、201Lによって生成されるこれらの情報は、制御部300における後述する属性情報抽出部305に入力されてもよい。

[0027] 映像AD変換部202R、202Lは、それぞれ、撮像素子201R、201Lが出力するアナログの電気信号をデジタル信号へ変換する回路である。その結果、撮像素子201Rによって取得された右眼用映像および撮像素子201Lによって取得された左眼用映像がデジタル信号としてそれぞれ出力される。

[0028] 信号処理部203は、映像AD変換部202R、202Lから出力された左右両映像における画角および画素数を合わせる処理を行う。その結果、同一の撮影範囲についての同一の画素数の左右映像が生成される。画角および画素数を合わせる処理の詳細は、例えば特許文献4に開示されている。特許文献4の記載内容全体を本願明細書に援用する。なお、レンズ群200R、200Lおよび撮像素子201R、201Lを同一の構成にする場合、画角合わせおよび画素数合わせ処理は不要である。

[0029] 信号処理部203は、さらに、左右の両映像信号を、NTSC (National Television System Committee) や PAL (Phase Alternating Line) といった所定の映像信号フォーマットに変換する。信号処理部203は、映像AD変換部202からのデジタル信号を、例えばNTSCで規定された水平線の数、走査線の数、フレームレートに準拠したデジタル映像信号（映像データ）に変換する。信号処理部203からの出力は、制御部300における映像解析部303へ入力される。信号処理部203は、例えば、映像信号変換用のICで構成され得る。あるいは、ICと信号処理を規定したプログラムとの好適な

組み合わせによって上記の処理が実現され得る。なお、映像信号フォーマットとしては、1映像フレームの有効画素数が水平方向に1920、垂直方向に1080である所謂フルハイビジョン方式や、1映像フレームの有効画素数が水平方向に1280、垂直方向に720である方式等があり得る。

[0030] 映像信号圧縮部204は、信号処理部203から出力されるデジタル映像信号に所定の符号化変換を施し、データ量を圧縮する。符号化変換の具体例としては、MPEG (Moving Picture Experts Group) 2、MPEG4、H264といった符号化方式があり得る。映像信号圧縮部204の出力は制御部300における多重化部308へ入力される。映像信号圧縮部204は、例えば、信号圧縮伸張用のICで構成され得る。

[0031] レンズ制御モジュール205は、レンズ群200R、200Lの状態を検出したり、レンズ群200R、200Lを動作させるように構成されている。レンズ制御モジュール205は、レンズ制御モータとレンズ位置センサとを有している。レンズ位置センサは、レンズ群200R、200Lを構成する複数のレンズ間の距離または位置関係等を検出するものである。レンズ位置センサは、その検出信号を制御部300におけるレンズ制御部301に出力する。レンズ制御モジュール205は、2種類のレンズ制御モータを有している。第1のレンズ制御モータは、制御部300におけるレンズ制御部301からの制御信号に基づいて、レンズ群200Rを光軸方向に移動させる。これにより、レンズ群200Rに含まれる複数のレンズ間の距離が変更され、レンズ群200Rの焦点距離およびズーム倍率が調整される。第2のレンズ制御モータは、制御部300におけるレンズ制御部301からの制御信号に基づいて、レンズ群200R、200Lに含まれる少なくとも1つのレンズ（手振れ補正用のレンズ）を光軸に直交する面内で移動させる。これにより、像振れを補正することができる。

[0032] 姿勢検出部206は、ビデオカメラ100a本体の姿勢を検出する。姿勢検出部206は、加速度センサ、角速度センサ、および仰角・俯角センサを

有する。これらのセンサにより、撮影時にビデオカメラ100aがどのような姿勢となっているかを認識することができる。加速度センサおよび角速度センサは、ビデオカメラ100aの姿勢を詳細に検出するために、直交3軸方向（ビデオカメラ100aの上下方向、前後方向、左右方向）の姿勢をそれぞれ検出できる。姿勢検出部206からの信号は、制御部300におけるレンズ制御部301へ入力される。なお、姿勢検出部206は、上記の構成に限らず、上記センサの一部だけで構成されていてもよいし、上記センサ以外のセンサを有していてもよい。

[0033] 外部入力部207は、ビデオカメラ100aに外部からの情報を入力するためのインタフェースである。外部入力部207からの信号は、制御部300における属性情報抽出部305へ入力される。なお、図2では、外部入力部207からの信号は、制御部300における属性情報抽出部305だけに入力されているが、レンズ制御部310等の入力操作に対応する部分にも入力され得る。撮影時には、この外部入力部207を介してビデオカメラ100aに外部からの各種情報が入力される。例えば、外部入力部207は、ユーザからの情報の入力を受け付ける入力インタフェースの1つである入力ボタンや、外部から通信によって入力される撮影インデックス情報等を受信する受信部や、ビデオカメラ100aが三脚に設置されているか否かを検出する三脚センサを含み得る。例えば、ユーザが入力ボタンを操作することによって、撮影の開始、終了、撮影中の映像にマーキングを挿入する動作や、後述する属性情報およびその評価を入力、設定する、といったユーザからの各種要求をビデオカメラ100aに伝えることができる。すなわち、外部入力部207は、ユーザの入力操作に応じて、後述する属性情報および評価値の少なくとも一方を記録部330に入力する入力部を構成する。ここで、撮影インデックス情報とは、例えば、映画撮影時における撮影場面を識別する番号や、撮影回数を示す番号といった各撮影を識別するために用いられる識別子である。三脚センサは、ビデオカメラ100aの三脚が固定される部分に設けられたスイッチで構成されている。三脚センサにより、三脚を使用して

いるか否かを判定できる。

- [0034] マイクロフォン208は、ビデオカメラ100aの周囲の音を電気信号に変換して、音声信号として出力する。音声信号は音声AC変換部209に入力される。
- [0035] 音声AD変換部209は、マイクroフォン208が出力するアナログの音声信号をデジタル音声信号（音声データ）へ変換する。音声AD変換部209は、例えばAD変換回路で構成される。変換されたデジタル音声信号は、制御部300における音声解析部304および音声信号圧縮部210に入力される。
- [0036] 音声信号圧縮部210は、音声AD変換部209が出力したデジタル音声信号を所定の符号化アルゴリズムで変換するものである。符号化にはMP3（MPEG Audio Layer-3）やAAC（Advanced Audio Coding）といった方式が用いられ得る。音声信号圧縮部210は、例えば、圧縮用のICで構成される。
- [0037] 映像信号伸張部211は、制御部300における後述するダイジェスト再生部309から出力され、自動奥行き設定部310によって視差量が調整された映像信号を復号する。映像信号伸張部211からの出力は、映像表示部212に入力される。映像信号伸張部211は、例えば、映像信号伸張用のICで構成される。
- [0038] 映像表示部212は、ビデオカメラ100a内の記録媒体に記録した映像や、ビデオカメラ100aがリアルタイムに撮影している映像を表示するディスプレイである。映像表示部212には、上記の映像の他、撮影に関する情報や機器情報のような各種情報が表示される。映像表示部212は、例えばタッチセンサを有する液晶ディスプレイで構成される。このようなタッチセンサを有する映像表示部212は、外部入力部207としても機能する。
- [0039] 音声信号伸張部213は、制御部300におけるダイジェスト再生部309から出力される音声信号を復号するものである。音声信号伸張部213によって復号された音声信号は、音声出力部214に入力される。音声信号伸

張部 213 は、例えば、音声信号伸張用の IC で構成される。

[0040] 音声出力部 214 は、映像に付随する音声を出力する。音声出力部 214 はまた、ビデオカメラ 100a からユーザへ報知する警告音を出力する。音声出力部 214 は、例えばスピーカまたはイヤホン端子である。

[0041] 出力 I/F 215 は、映像信号をビデオカメラ 100a から外部へ出力するためのインタフェースである。具体的には、出力 I/F 215 は、ビデオカメラ 100a と外部機器とをケーブルで接続するためのケーブルインタフェースや、映像信号をメモリカードに記録する場合のメモリカードインタフェースであり得る。

[0042] 制御部 300 は、ビデオカメラ 100a の全体を制御する機能部である。制御部 300 は、撮像素子 201R、信号処理部 203、映像信号圧縮部 204、レンズ制御モジュール 205、姿勢検出部 206、外部入力部 207、音声 AD 変換部 209、音声信号圧縮部 210、映像信号伸張部 211、音声信号伸張部 213、出力 I/F 215 および記録部 330 との間で信号の授受が可能に構成されている。本実施形態では、制御部 300 は、CPU で構成されている。制御部 300 は、記録部 330 に格納されたプログラムを読み込んで実行することによってビデオカメラ 100a の各種制御を実行する。制御部 300 が実行する制御の例としては、例えば、レンズ群 200R、200L の焦点距離やズームの制御や、姿勢検出部 206 や外部入力部 207 からの入力信号の処理や、信号処理部 203、映像信号圧縮部 204、音声信号圧縮部 210、映像信号伸張部 211 および音声信号伸張部 213 等の IC の動作制御が挙げられる。また、図示は省略するが、制御部 300 とレンズ制御モジュール 205 等との間では、適宜、信号が AD 変換または DA 変換され得る。なお、制御部 300 は、集積回路などのハードウェアのみで構成することもできる。

[0043] クロック発振器 320 は、ビデオカメラ 100a 内で動作する制御部 300 等に処理動作の基準となるクロック信号を出力するものである。なお、クロック発振器 320 は、利用する集積回路や扱うデータによって、単一また

は複数のクロックを用いることも可能である。また、ひとつの発振子のクロック信号を任意の倍数に乗じて使用してもよい。

[0044] 記録部330は、ROM (Read Only Memory)、RAM (Random Access Memory) およびHDD (Hard Disk Drive) を含み得る。ROMは、制御部300が処理するプログラムやプログラムを動作させるための各種データを格納する記録媒体である。RAMは、制御部300がプログラム実行時に使用するメモリ領域として使用される。また、RAMは、ICのメモリ領域としても使用され得る。HDDは、映像信号圧縮部204により符号化変換された映像データや静止画データといった各種データを保存する。また、HDDには、制御部300が実行するプログラムが記録されている。なお、このプログラムは、HDDに限られず、半導体メモリに記録されていてもよいし、CD-ROMやDVD等の可搬型記録媒体に格納されていてもよい。

[0045] 以下、制御部300の構成および機能について詳細に説明する。ここでは、制御部300のうち、ダイジェスト再生を行う際に必要な構成を中心に説明する。ただし、制御部300は、一般的なビデオカメラと同様の機能を発揮させるための構成、例えば、撮影した映像を記録したり、再生したりする構成も有する。

[0046] 制御部300は、レンズ制御モジュール205を制御するレンズ制御部301と、撮像素子201R、201Lを制御する撮像制御部302と、信号処理部203から出力される立体映像を解析する映像解析部303と、音声AD変換部209からの出力を解析する音声解析部304と、映像中の属性情報を抽出する属性情報抽出部305と、シーンを評価するシーン評価部306と、ダイジェスト再生のための再生情報を生成する再生情報生成部307と、映像データおよび音声データと多重化する多重化部308と、ダイジェスト再生のための映像を生成するダイジェスト再生部309と、映像の奥行きを調整する自動奥行き設定部310とを有している。制御部300は、記録部330に記録されたプログラムを読み込んで実行することにより、後

述する各種の処理を実現する。映像解析部303は、映像に含まれる被写体の検出、被写体と背景の分離、また、被写体および背景の輪郭、カラー、テクスチャの解析やオブジェクト認識、さらに、映像の構図の解析を行う。なお、本実施形態では、ダイジェスト再生部309および自動奥行き設定部310がインタフェースおよびダイジェスト映像生成部としての機能を有する。

[0047] レンズ制御部301には、レンズ制御モジュール205のレンズ位置センサの検出信号および姿勢検出部206の各種センサの検出信号が入力される。レンズ制御部301は、これらの検出信号および撮像素子201R、201L等の他の構成要素によって生成された情報に基づいて、レンズ群200R、200Lを適正に配置させるための制御信号をレンズ制御モータに出力する。このようにして、レンズ制御部301は、ズーム制御、フォーカス制御、手ぶれ補正制御を行う。また、レンズ制御部301は、これらの制御信号を属性情報抽出部305にも送付する。なお、姿勢検出部206の各種センサの検出信号は、属性情報抽出部305にも入力される。

[0048] 撮像制御部302は、撮像素子201R、201Lの動作、例えば撮影時の露出量、撮影速度、および感度を制御する。撮像制御部302から出力される制御信号は、撮像素子201R、201Lだけでなく、属性情報抽出部305にも送付される。

[0049] 映像解析部303は、信号処理部203からの映像データに基づいて映像の特徴を抽出する。映像解析部303は、映像の色情報（例えば、映像に含まれる色の分布）およびホワイトバランス情報を検出する。なお、色分布の検出は、デジタル映像信号を形成するデータに含まれる色情報を確認することで実現可能である。また、映像解析部303は、映像に人物の顔が含まれている場合には、映像の中から顔検出を行う。顔検出については、パターンマッチング等の公知の方法を用いることによって実現可能である。

[0050] 音声解析部304は、音声AD変換部209からの音声データを解析して、特徴のある音を抽出する。ここで特徴のある音とは、例えば、撮影者の声

、特定の単語の発音、歓声、および銃声等があり得る。これらの音（音声）が持つ特有の周波数を予め登録しておけば、それらの音に近い音が含まれているか否かを判別できる。また、これ以外にも、例えば、音の入力レベルが所定レベル以上のときにその音が特徴のある音であると判定するようにしてもよい。

[0051] 属性情報抽出部305は、映像に関する属性情報を抽出する。属性情報は、映像の属性を示す情報であり、例えば撮影に関する情報（以下、「撮影情報」ともいう）、外部から入力された情報、およびその他の情報を含み得る。属性情報抽出部305には、撮像素子201R、201L、姿勢検出部206、外部入力部207、レンズ制御部301、撮像制御部302、映像解析部303、および音声解析部304から信号が入力される。属性情報抽出部305は、これらの信号に基づいて属性情報を抽出する。

[0052] 属性情報には、映像中の被写体の距離情報が含まれる。距離情報は、当該被写体とビデオカメラ100aとの距離を示す情報であり、被写体の飛び出し量または奥行き量を示す。この距離情報は、個々の画像フレームにおいて合焦している主要被写体（人物の顔など）とビデオカメラ100aとの距離として定められる。この距離情報は、例えばレンズ群200R、200Lの焦点距離から求められ得る。あるいは、以下のいずれかの方法によって距離情報を求めてもよい。

[0053] 一般に、被写体距離を測定する方法には、複数のカメラを用いる方法と測距専用のカメラを用いる方法とがある。複数のカメラを用いる方法では、例えば、エピポーラ線上に2台のカメラを互いの光軸が平行になるように（平行法）左右に設置し、撮影した左右映像中の同一オブジェクトについて生じる視差量に基づいて、奥行き情報、すなわち距離情報を求めることができる。本実施形態では、2組の撮像系（光学系200R/L、撮像素子201R/L、および映像AD変換部202R/L）を上記の2台のカメラとして用いることにより、距離情報を求めることができる。

[0054] 一方、測距専用カメラを用いる方法では、TOF（Time of Flight）

light) 方式のように、特定の発光パターンを有する光源からオブジェクトに光を照射し、そのオブジェクトからの反射光が検出されるまでの時間を測定することによって測距が行われ得る。このような測距専用カメラを用いて被写体距離を測定する場合、ビデオカメラ100aには、図2に示す構成要素に加え、測距専用カメラが付加される。

[0055] また、ロボットの分野等で利用されているレーザーレンジファインダのようにレーザーを使った測距方式や、超音波やミリ波を利用した測距方式など、距離を測定する方法には様々な方式がある。距離情報の取得方法は、特定の方法に限定されない。

[0056] 属性情報には、距離情報の他、例えば、撮影時の撮影装置の状態やカメラワークなどの撮影に関する情報、CGなどによる映像に関する情報、映像そのものが含む被写体や背景に関する距離情報以外の情報、映像に付帯する音声に関する情報、映像の編集内容に関する情報が含まれ得る。

[0057] 撮影時の撮影装置に関する属性情報の他の例としては、焦点距離、ズーム倍率、露出、撮影速度、感度、3原色点の色空間情報、ホワイトバランス、3原色のうち少なくとも2色のゲイン情報、色温度情報、 Δuv (デルタ u v)、3原色または輝度信号のガンマ情報、色分布、顔認識情報、カメラ姿勢(加速度、角速度、仰角・俯角等)、撮影時刻(撮影開始時刻、終了時刻)、撮影インデックス情報、ユーザ入力、フレームレート、サンプリング周波数等が挙げられる。属性情報抽出部305は、例えばレンズ制御部301から出力される制御信号に基づいて、被写体の距離情報の他、焦点距離やズーム倍率を属性情報として抽出し得る。また、属性情報抽出部305は、姿勢検出部206から出力される信号に基づいて、カメラ姿勢(加速度、角速度、仰角・俯角等)を検出し、当該カメラ姿勢から、パン、ティルト等の撮影時のカメラワークを属性情報として抽出することができる。さらには、これらのカメラワークに基づいて、カメラワーク後のフィックス撮影部分(ビデオカメラ100aを静止させて撮影した部分)を属性情報として抽出することができる。このように、属性情報抽出部305は、入力された信号その

ものから属性情報を抽出してもよいし、入力された信号と組み合わせたり、分析したりすることによって属性情報を抽出してもよい。

[0058] シーン評価部306は、属性情報抽出部305が抽出した属性情報に基づいて、それぞれの属性情報を含む部分の映像を評価して、その評価（値）を当該部分に付与する。この評価の詳細については後述する。

[0059] 再生情報生成部307は、シーン評価部306によって付与された評価に基づいて、再生すべき部分（シーン）を選択し、ダイジェスト再生すべき部分を特定する情報（以下、「再生情報」という）を生成する。再生情報の詳細についても後述する。

[0060] 多重化部308は、映像信号圧縮部204から出力された符号化映像データ、音声信号圧縮部210から出力された符号化音声データ、および再生情報生成部307から出力された再生情報を多重化して出力する。多重化部308により多重化されたデータは、記録部330に格納される。多重化の方式としては、例えば、MPEGのTS（Transport Stream）等の技術がある。ただし、これに限定されるものではない。なお、本実施形態では、多重化する場合を例として示しているが、必ずしも多重化しなくてもよい。

[0061] 属性情報抽出部305、シーン評価部306、再生情報生成部307、および多重化部308の処理は、撮影中または撮影直後に逐次、実行される。

[0062] ダイジェスト再生部309は、撮影の終了後に、ユーザの入力に基づいてダイジェスト再生のための立体映像を生成する。ダイジェスト再生部309は、記録部330に記録された多重化されたデータを読み出し、再生情報に従って、ダイジェスト再生させるべきシーンの符号化映像データおよび符号化音声データを自動奥行き設定部310に出力する。

[0063] 自動奥行き設定部310は、立体映像の安全性の確保および臨場感の向上を図るため、必要に応じて立体映像の奥行き量（飛び出し量）を調整する処理を行う。自動奥行き設定部310は、ダイジェスト映像のシーンの切り替わり前後で被写体の飛び出し量が本来の値から変化するようにシーンの切り

替わり前後の少なくとも一部の画像フレームの視差量を調整する。この調整方法の具体例については後述する。自動奥行き設定部 310 は、調整後の符号化映像データおよび符号化音声データをそれぞれ映像信号伸張部 211 および音声信号伸張部 213 に出力する。

[0064] 出力された符号化映像データおよび符号化音声データは、それぞれ映像信号伸張部 211 および音声信号伸張部 213 によって復号され、映像表示部 212 および音声出力部 214 から出力される。このようにして、映像の中から特定の部分だけを抽出したダイジェスト再生が実行される。なお、このダイジェスト映像は、再生する代わりに記録部 330 に記録するようによい。

[0065] 上述の奥行き量（飛び出し量）の調整の目安として、人が両眼視差（または網膜像差）を検出できる最小値（最小視差弁別閾）から融像できる最大値までの範囲内で視差を調節することが適切と考えられる。視差の検出が可能な最小値は約 1.2 秒（1 秒は 1 度の 3,600 分の 1）であり、この値よりも小さいと視差を検出できない。また、パーナム（Panum）の融像限界によって決まると言われている立体の融像限界値は約 2 度であり、この値より大きくなれば不安定な二重像ができると言われている。特に 2 度から 5 度までは負荷が強い不安定な領域であり、安定した融像領域は約 70 分（1 分は 1 度の 60 分の 1）以下といわれている。

[0066] これらのことを考慮すると、安全で臨場感のある立体映像の奥行き量（飛び出し量）の調整の目安としては、約 1 分から 70 分の間で調整すればよい。また、高速で飛び出してくる物体のように、人がピントを調節できない物体については、短時間であれば、2 度を超える調整をしても安全で迫力のある立体映像が得られると考えられている。シーンを説明する文字や図形の CG を 3 次元空間に貼り付けて表示する場合においても、この調整範囲内で行われることが一般的である。本実施形態における奥行き調整においても、この調整範囲内で奥行き調整を行うことにより、安全で、より臨場感のある立体映像のダイジェスト再生が可能になる。

[0067] [1-2. 動作]

[1-2-1. シーン評価および再生情報の生成]

図3は、ビデオカメラ100aが撮影した映像の構成を示す模式図である。図4は、クリップ内を複数のシーンに分割した例を示す図である。図4では、各シーンを「開始時間」と「終了時間」で特定しているが、他の情報、例えばフレーム番号（シーンの開始フレーム番号および終了フレーム番号）によって各シーンを特定してもよい。

[0068] 本明細書では、ユーザが撮影の開始を指示し、撮影の終了または撮影の一時停止を指示するまでに撮影された映像の単位を「クリップ」と呼ぶ。すなわち、ユーザが撮影の開始、および撮影の終了または一時停止を何度も繰り返すと、クリップが複数生成される。1つのクリップは、少なくとも1つの「シーン」から構成される。「シーン」は、論理的につながりのある一続きの映像であり、少なくとも1つの「フレーム」から構成される。「フレーム」とは、映像を構成する最小単位である個々の画像である。本明細書では、「フレーム」のことを「画像フレーム」と称することがある。

[0069] 1つの「クリップ」を1つの「シーン」として扱うこともできる。また、画面が大きく変わるフレームを境として「シーン」を設定してもよい。例えば、映像解析部303が、連続するフレーム間の動きベクトルを算出し、動きの大きさ（即ち、動きベクトルの変化量）が所定の値より大きいときを、「シーン」の境界としてもよい。つまり、そのようにして設定された2つの境界の間に含まれる映像が1つの「シーン」に設定される。あるいは、その他の撮影情報に基づいて「シーン」を区切ってもよい。例えば、撮影者からのボタン入力により「シーン」を区切ってもよい。この場合、撮影者の明確な意図に基づいて「クリップ」内の「シーン」が構成される。以上の方法で「シーン」を抽出した場合には、図4に示すように、「クリップ」には複数の連続する「シーン」が含まれることになる。

[0070] 一方、「クリップ」内の特定の部分だけを「シーン」として扱うこともできる。例えば、映像の中から、映像として重要な部分だけを「シーン」とし

て扱うこともできる。具体的には、特定の属性情報を含む部分を1つの「シーン」として扱ってもよい。例えば、重要と想定される属性情報を含む所定の時間幅の映像を「シーン」とすることができる。そのようにすれば、重要な部分のみが「シーン」として抽出されることになる。その結果、「クリップ」内には、離散的に複数の「シーン」が含まれることになる。このように、「シーン」は、任意に設定することが可能である。本実施形態では、映像に含まれる重要な部分を「シーン」として取り扱う。

- [0071] 続いて、シーン評価部306によるシーン評価の動作を説明する。
- [0072] 図5は、映像を評価する際に用いられる各種の属性情報と各属性情報に対する評価値との対応を示すテーブルの例を示す図である。このテーブルは、記録部330に記録されている。シーン評価部306は、このテーブルを参照することによって映像を評価する。
- [0073] 図5に示すように、個々の属性情報には評価値が個別に設定されている。図5に示す例では、評価値が高いほど評価が高い（好ましい）ことを表している。例えば、クリップイン（撮影の開始部分）およびクリップアウト（撮影の終了直前部分）は、それぞれ映像の導入部分および終了部分であり、映像が持つ論理的な意味が高いと推定される。そのため、クリップイン（A）には、評価値「100」が、クリップアウト（F）には、評価値「90」が設定されている。撮影時のカメラワークとしてズームアップ（D）やズームダウン（G）は、特定の被写体への注目度を高めるものであるため、評価値「30」が設定されている。
- [0074] さらに、主要被写体の飛び出し量を大きくして臨場感や迫力を増したい場合には、飛び出し量の大きいシーンが優先して抽出されるようにすればよい。図5に示す例では、被写体距離一近（K）は、主要な被写体の飛び出し量が大きいことを示しており、評価値「70」が設定されている。一方、被写体距離一中（L）は、主要な被写体の飛び出し量が中程度であることを示しており、評価値「10」が設定されている。このような設定により、迫力のある立体映像のダイジェスト映像が生成される。ここで、被写体距離一近（

K) または被写体距離一中 (L) に該当するか否かは、距離情報が予め定められた範囲内にあるか否かによって判断される。例えば、距離情報から定まる被写体距離が第1の閾値よりも短ければ被写体距離一近 (K) に該当し、当該被写体距離が第1の閾値よりも長く、第2の閾値よりも短い場合は被写体距離一中 (L) に該当する、とすることができる。

[0075] また、一般に人物を対象に撮影が行われることが多いことから、顔が検出されたこと (Z) には評価値「50」が設定されている。特に、顔検出の中でも、特定の人物Aの顔が検出されたこと (X) は、「100」の評価値が設定され、特定の人物Bの顔が検出されたこと (Y) は、評価値「80」が設定されている。

[0076] この特定の人物の顔およびそれに対する評価値は、詳しくは後述するが、ユーザが適宜設定することができる。つまり、単に人物を撮影しているというだけでなく、特定の人物を撮影している映像にはユーザの意思で高い評価値を付与することができる。なお、評価は、プラスの評価、即ち、好ましい評価だけでなく、マイナスの評価、即ち、好ましくない評価もあり得る。例えば、像ぶれは視聴者にとって見づらい映像となる可能性があるので、このような属性情報を有するシーンには、負の評価値が付与される。

[0077] なお、図5に示す例では、評価が数値化されているが、このような例に限られるものではない。例えば、評価としてA, B, C, …のような符号を用いてもよい。この場合、評価として用いられる符号には、予め優劣が定められていればよい (例えば、Aが最も評価が高い等)。また、A, B, Cなどの符号の評価は、ユーザの意思で自由に設定可能にしてもよい。

[0078] シーン評価部306は、属性情報抽出部305によって属性情報が抽出された映像の部分に、上記テーブルに基づいて、当該属性情報に対応する評価値を付与する。当該評価値に基づいて、所定の個数のシーンを抽出する。このようにして、シーン評価部306は、ダイジェスト再生に用いられ得る特徴的な映像の部分シーンをシーンとして、ダイジェスト再生されるシーンの個数よりも多めに予め抽出しておく。例えば、シーン評価部306は、評価値が高

い属性情報を有する部分を含む所定時間幅の映像を1つのシーンとして抽出する。そして、評価値が高い順に、所定の個数のシーンを抽出する。この所定の個数は、ユーザが任意に設定可能としてもよいし、固定値として予め設定しておいてもよい。ここで、評価値が高い属性情報を有する部分であっても、それよりも前の映像に同じ属性情報が含まれる部分をシーンとして抽出している場合には、シーンとして抽出しないようにしてもよい。このようにすることで、同じ属性情報を有するシーンばかりが抽出されることを防止することができる。あるいは、特定の属性情報（例えば、人物Aの顔検出や人物Bの顔検出等）を有する部分を優先的にシーンとして抽出してもよい。

[0079] 所定の個数のシーンを抽出した後、シーン評価部306は、抽出したシーンの中から、所定の抽出条件に基づいて、ダイジェスト再生すべきシーンを抽出する。例えば、抽出条件が「評価値が高い順に3個」という条件であれば、シーン評価部306は、評価値が上位3つのシーンを抽出する。この個数は、ユーザが任意に設定できるようにしてもよい。また、抽出条件が「評価値の高い順に抽出して、合計時間が所定時間内になる」という条件であれば、シーン評価部306は、合計時間が所定の時間内に納まるように、評価値が上位のシーンから順に抽出する。この所定の時間は、予め所定の値に設定しておいてもよいし、ユーザが任意に設定できるようにしてもよい。また、評価値が所定の値以上であることが抽出条件であれば、シーン評価部306は、評価値が所定の値以上のシーンを、個数および合計時間にかかわらず抽出する。この所定の値は、ユーザが任意に設定することができようにしてもよい。このように、シーン評価部306は、付与した評価値に基づいて、様々な観点でシーンを抽出することが可能である。なお、抽出条件は、ユーザが適宜設定できるようにしてもよいし、予めいくつかの条件が設定されていてもよい。

[0080] また、1つのシーンが複数の属性情報を有する場合には、それぞれの属性情報の内容に割り当てられている評価値を加算した値をそのシーンの評価値としてもよい。あるいは、複数の属性情報の中から最も高い評価値または複

数の属性情報の評価値の平均値をそのシーンの評価値としてもよい。

[0081] ビデオカメラ100aは、属性情報と評価値との対応関係を規定するテーブルを1つに限らず、複数有し、シーン評価に用いるテーブルを適宜選択するように構成されていてもよい。例えば、ビデオカメラ100aが、属性情報と評価値との対応関係を規定する複数のテーブルの中から撮影モード（例えば、風景の撮影、人物（ポートレート）撮影、スポーツ撮影、静物撮影等）に応じて最適なテーブルを選択するように構成してもよい。また、このように撮影の状況に応じて適切なテーブルを適宜設定する構成としては、撮影の各状況に対して1対1のテーブルを予め用意しておくのではなく、撮影状況の種類よりも少ない数のテーブルを用意しておいてもよい。この場合、撮影状況に応じて、複数のテーブルを合成（各評価値を一定の比率で加算する等）してもよい。この際、合成時の各テーブルの重み付け（例えば、加算時の比率）を変えておけば、撮影状況に応じたテーブルを設定することもできる。

[0082] 以下、シーン評価部306が行う処理を、具体例を挙げて説明する。

[0083] 図6は、シーン評価部306が、ある映像から属性情報を抽出して評価値を付与した結果の例を示す図である。図6の横軸は時間（シーン）を、縦軸は評価値を表している。

[0084] 図6において、時間0付近の部分は、撮影を開始した直後であることを意味する「クリップイン」の属性情報Aを有し、評価値「100」が付与されている。

[0085] 属性情報K、Lを有する部分は、主要被写体の距離が比較的近く、立体特性が高いと判定された部分である。属性情報Kを有する部分は評価値「70」が付与され、属性情報Lを有する部分は評価値「10」が付与されている。

[0086] 属性情報Cを有する部分は、ユーザがビデオカメラ100aをパン、ティルト等させた後に静止して撮影している部分である。パン、ティルト等のカメラワーク後の部分は映像としての価値が高いと判断できることから、この

ようなカメラワーク後の静止撮影を属性情報として設定している。属性情報Cを有する部分には、評価値「40」が付与されている。

[0087] 属性情報Dを有する部分は、ズームアップまたはズームダウン等して撮影している部分である。ズームアップまたはズームダウンは、ユーザの撮影に関する何らかの意図が反映されており、重要と判断できることから、属性情報として設定される。属性情報Dを有する部分には、評価値「30」が付与されている。ただし、ズームアップとズームダウンとで評価値を変えてもよい。例えば、ズームアップの方が、ズームダウンよりも、撮影対象を注視する意図が大きいと判断されるため、評価値を高く設定してもよい。

[0088] 属性情報Eを有する部分は、属性情報Cと異なり、ビデオカメラ100aをパン、ティルト等させながら撮影を行った部分である。パン、ティルト等のカメラワークは、撮影対象に追従しようとするユーザの撮影意図が反映されていると判断できるため、属性情報として設定している。属性情報Eを有する部分には、評価値「25」が付与されている。

[0089] 属性情報Iを有する部分は、映像が像ぶれを伴っている部分である。この場合は、映像が揺れているため、視聴者には見づらい映像となる傾向にある。そのため、負の評価値を付与している。具体的には、属性情報Iを有する部分には、評価値「-20」が付与されている。

[0090] 属性情報Jを有する部分は、地面等を撮影している部分である。これは、ユーザが撮影停止ボタンを押さずに撮影を続けたまま、ビデオカメラ100aを手を持って歩いている場合等に生じやすい現象である。この場合には、映像にユーザの特段の意図が反映されていないと判断できるため、負の評価値を付与している。具体的には、属性情報Jを有する部分には、評価値「-10」が付与されている。

[0091] 属性情報Xを有する部分は、人物Aの顔が映っている部分である。映像解析部303は、撮影された映像に写っている被写体が人物の顔であることを認識し、さらに認識された顔が予め記録部330等に記録されている特定の人物の顔と一致するか否かの判定を行う。シーン評価部306は、この結果

に基づき、特定の人物の顔が写っている部分を映像の中から抽出することができる。属性情報Xを有する部分には、評価値「100」が付与されている。

[0092] なお、人物の顔が撮影される映像の例として、図7A～7Dに示すような場合がある。図7A、7Bに示すように、それぞれの部分に予め登録されている人物の顔が撮影されている場合は、シーン評価部306は、テーブルに予め設定された評価値（図5の例では、人物Aの顔については「100」、人物Bの顔については「80」）が付与される。

[0093] 図7Cに示すように、人物Aと人物Bの両者の顔が写っている場合には、シーン評価部306は、人物A、Bの顔のうち、評価値が高い方の評価値をもって、評価値としてもよい。図5のテーブルを用いた場合には、人物Aの顔の方が人物Bの顔よりも評価値が高いため、人物Aの顔の評価値である「100」が評価値となる。なお、図6におけるX+Yの符号が付されている部分は、人物Aと人物Bの両者の顔が写っている部分である。また、別の方法として、両方の評価値を平均化したものを評価値としてもよい。この場合、図5に示す例では、 $(100 + 80) / 2$ から、評価値は「90」となる。

[0094] さらに別の方法として、それぞれの評価値に配分比率を設定して合算するものであってもよい。例えば、映像上の顔の大きさが大きいほど、配分比率が高くなるように、それぞれの評価値に配分比率を設定してもよい。図7Cの例では、人物Aの顔の大きさと人物Bの顔の大きさの比率が5：3であるとする、 $(100 \times 5 + 80 \times 3) / 8$ から、評価値は「92.5」となる。映像上の顔の大きさはカメラから被写体までの距離を反映している可能性が高く、この方法によれば、近くの被写体に大きな配分比率を設定して、影響度を大きくすることができる。また、人物A、Bの顔の中心位置と、画面中央または画面上の顕著領域（Salientな領域）からの距離に応じて、それぞれの評価値に配分比率を設定してもよい。具体的には、当該距離が近いほど、配分比率が高くなるようにしてもよい。顕著領域とは、例えば

、画像中の記念像、ペット、カンバンなどの注目される領域である。

[0095] 図7Dに示すように、大勢の（所定の人数以上の）人物が被写体として撮影されている場合は、人物Aの顔が撮影されているものの、画面上での大きさは小さく、他の多くの人物と一緒に撮影されているため、他の人物の顔との差別化が難しい場合がある。このような場合には、撮影されている人数に応じて、人物Aの顔の評価値と、他の人物の顔の評価値とに配分比率を設定して合算するものであってもよい。図7Dに示す例では、人物A以外の、識別できない人物が10人いるので、 $(100 \times 1 + 50 \times 10) / 11$ から、評価値「54.5」が付与される。

[0096] また、顔の検出については、画面における位置、顔の大きさ、顔の向き、笑顔レベル、目の開閉情報、顔の喜怒哀楽レベルの情報を評価して、これらに応じて評価値を増減させてもよい。

[0097] このようにして、映像中で属性情報が抽出された部分に評価値を付与した後、シーン評価部306は、評価値が高い順に6個のシーンを抽出する。図6では、6個のシーンに、時間が早い順に#1～#6のラベルが付されている。続いて、シーン評価部306は、「評価値が高い順に3個」という抽出条件が設定されている場合、評価値が上位3つの#1、#2、#5のシーンをダイジェスト再生すべきシーンとして抽出する。

[0098] 次に、再生情報生成部307による再生情報の生成について詳細に説明する。再生情報生成部307は、シーン評価部306が抽出したシーンに従って、ダイジェスト再生すべきシーンを特定する情報である再生情報を生成する。例えば、再生情報は、図8に示すように、再生対象となるシーンの開始時刻と終了時刻で示されるものであり得る。この場合、各シーンの中における代表的なフレーム（シーン中における最も評価の高いフレーム等）を別途記録しておくこと、参照用画面の検索に有効である。なお、再生情報は、上記の内容に限られず、例えば、再生対象となるシーンをフレーム番号で特定してもよい。それ以外にも、後述する多重化部308で生成する多重化データ中における該当シーンの位置（場所）をシーンの特定として用いてもよい。

多重化にTS等の技術を用いる場合は、PTS (Presentation Time Stamp) やDTS (Decoding Time Stamp) 等の時刻情報等を用いて再生情報を生成してもよい。一部のビデオカメラのデータ記録方式として用いられているAVCHD (Advanced Video Codec High Definition、登録商標) 等の規格を用いて映像データを記録する場合には、Playlistファイル等に再生情報を記録する方法を用いてもよい。

[0099] 次に、撮影から記録までの全体の処理の流れを説明する。

[0100] 図9は、撮影、シーン評価、再生情報の生成、記録までの処理の流れを示すフローチャートである。

[0101] まず、ビデオカメラ100aの制御部300は、ステップS101において、撮影を開始する。入力ボタン等の外部入力部207からの入力に基づいて、撮影が開始される。

[0102] 次に、ステップS102において、属性情報抽出部305は、姿勢検出部206の検出結果、レンズ制御部301の制御情報、並びに映像解析部303および音声解析部304の解析結果等に基づいて映像の属性情報を抽出する。

[0103] 続くステップS103において、シーン評価部306は、属性情報抽出部305が抽出した属性情報に基づいて映像の各部分に評価値を付与する。その後、シーン評価部306は、いくつかの特徴的なシーンを抽出し、さらに、その中からダイジェスト再生すべきシーンを抽出する。

[0104] 続いて、ステップS104において、再生情報生成部307は、シーン評価部306によって抽出されたダイジェスト再生すべきシーンに基づいて再生情報を生成する。そして、多重化部314は、生成された再生情報を、符号化映像データおよび符号化音声データと共に多重化する。

[0105] 制御部300は、ステップS105において、多重化データを記録部330に記録する。

[0106] 制御部300は、ステップS106において、外部入力部207から、撮

影終了の入力があるか否かを判定する。撮影終了の入力がない場合は、ステップS102へ戻って、撮影を継続する。一方、撮影終了の入力がある場合には、撮影を終了する。

[0107] [1-2-2. 再生情報に基づくダイジェスト再生]

続いて、記録された再生情報に基づくダイジェスト再生の動作の流れを説明する。

[0108] ダイジェスト再生部309は、記録部330に記録された再生情報を読み出し、それに基づいてダイジェスト再生のための映像を生成する。具体的には、ダイジェスト再生部309は、図8に示されるダイジェスト再生すべき個々のシーンの開始時刻、終了時刻等の情報に基づいて記録部330に記録されている映像、音声情報から該当する部分のシーンを抽出する。そして、自動奥行き設定部310は、抽出されたシーンの映像について、必要に応じて飛び出し量（視差量）の調整を行った後、当該シーンを再生する。

[0109] 図10Aは、ダイジェスト再生時の処理の流れを示すフローチャートである。

[0110] まず、ステップS201において、ダイジェスト再生部309は、記録部330に記録された多重化データを読み出す。

[0111] 次に、ステップS202において、ダイジェスト再生部309は、読み出した多重化データを分解し、再生情報を抽出する。

[0112] 続いて、ステップS203において、ダイジェスト再生部309は、再生情報に基づいてダイジェスト再生のための複数のシーンを選択する。

[0113] 次に、ステップS204において、自動奥行き設定部310は、選択した各シーン間の被写体距離の変化量を求める。この変化量は、例えば連続する2つのシーンの境界前後の複数の画像フレームにおける被写体の距離情報の差分をとることによって求められ得る。

[0114] そして、ステップS205において、直前のシーンに対する被写体距離の変化量が予め定められた閾値を超えるシーンが存在するか否かを判定する。該当するシーンが存在する場合には、ステップS206に進み、存在しない

場合には、ステップS 2 0 7に進む。

[0115] ステップS 2 0 6では、自動奥行き設定部3 1 0は、直前のシーンに対する被写体距離の変化量が予め定められた閾値以上であると判定したシーンにおいて、被写体距離が当該閾値よりも小さくなるように当該シーンの映像における視差量を補正する。視差量の補正の方法の具体例は後述する。視差量の補正が完了すると、処理はステップS 2 0 7に進む。

[0116] ステップS 2 0 7において、自動奥行き設定部3 1 0は、再生すべき符号化映像データおよび符号化音声データを、映像信号伸張部2 1 1および音声信号伸張部2 1 3に出力し、映像表示部2 1 2および音声出力部2 1 4を介して映像および音声を再生する。

[0117] このようにして、映像の中から特定のシーンだけを抽出したダイジェスト再生が実行される。

[0118] [1 - 2 - 3. 視差量の補正方法の具体例]

次に、ステップS 2 0 6における処理の具体例を説明する。左右両映像の視差量と被写体距離との間には相関関係があるため、自動奥行き設定部3 1 0は、被写体距離を調整するために左右両映像の視差量を調整する。ここで、左右両映像の視差量は、デプスマップによって表される。デプスマップは、例えば、映像中の視差のある部分が視差量に応じて値を持ち、視差のない部分は0の値をもつ画素ごとの情報である。ここで、「視差量」は、左右の一方の映像を基準としたときに、他方の映像における対応点が何画素ずれているかを示す量である。視差量が大きいほど飛び出し量が大きくなる。このように、デプスマップは、左右の映像間の視差量、すなわち距離情報を示す。デプスマップは、典型的には映像のフレームごとに作成される。

[0119] 視差量の補正方法の例として、例えば以下の方法が挙げられる。まず、デプスマップによって示される距離情報の分布に基づき、映像中で最も近い距離 (D_{near})、最も遠い距離 (D_{far})、平均距離 (D_{avr}) の各部分の発生頻度を求める。次に、(D_{near}) と (D_{far}) との差分を計算し、距離の分布幅 (D_{width}) を求める。視差量の第1の補正方法

として、平均距離 (D_{avr}) をDC的にシフトさせることにより、奥行きを前後に補正することができる。すなわち、(D_{avr}) が所定の量だけ増減するように、画像全体の距離を一律にその量だけ増減させることができる。また、視差量の第2の補正方法として、距離の分布幅 (D_{width}) を拡大または縮小させることにより、奥行きを調整してもよい。すなわち、 D_{near} に該当する部分および D_{far} に該当する部分の少なくとも一方を前後にシフトさせ、他の部分もそれに連動してシフトさせる方法を採用してもよい。これらの計算の際、距離の発生頻度による重み付けを行なうようにすれば、より安全性と迫力とを両立する3D映像が表現できる場合もある。あるいは、3Dコンソーシアム（日本における標準化団体）が規定している安全基準を満足するように、(D_{near})、(D_{far})、(D_{avr})、(D_{width}) を調整してもよい。これにより、3D映像の安全性をさらに高めることができる。

[0120] また、視差量の補正方法の他の例として、画像フレーム全体の視差量を一律に一定の割合で小さくする方法を採用することもできる。以下、図10Bおよび図10Cを参照しながら、この補正方法を説明する。

[0121] 図10Bは、撮像素子201R、201Lを通してそれぞれ取得された右側映像および左側映像の例と、それらの間の視差量を規定するデプスマップの一例を示す図である。デプスマップは、前述のように、視差のある部分が視差量に応じて値を持ち、視差のない部分は0の値をもつ情報である。図10Bに示す例では、わかり易さのため、実際よりも粗い精度で視差量を表現しているが、デプスマップは、実際には画素ごとの視差量の情報の集合であり得る。図10Bに示す例では、被写体である家屋が、左側映像では右側映像に比べて右に寄っているため、デプスマップにおける対応する箇所は、そのずれの画素数に応じた値をもつ。この例では、視差量が大きいほど、飛び出し量が大きくなる。

[0122] ここで、図10Bに示す家屋が写っている領域のうち、デプスマップにおける視差量が8である箇所が、直前の画像フレームからの被写体距離の閾値

を超えた部分であるとする。そして、当該箇所の視差量が2以下になれば閾値内に納まる場合を想定する。その場合、自動奥行き設定部310は、デプスマップ全体の値を1/4倍に補正し、補正後のデプスマップが示す視差量になるように、左側映像を補正すればよい。

[0123] 図10Cは、この例における補正後のデプスマップ、および左右の両映像を示す図である。デプスマップ全体の値が1/4倍になったことにより、左側映像における家屋が左側に寄っている。これにより、被写体の急激な飛び出し量の変化が緩和される。なお、この例では左側映像のみを水平方向にずらすことによって補正を行っているが、右側映像を水平方向にずらしてもよい。あるいは、左右の両映像を水平方向に半分ずつずらしてもよい。

[0124] 上記の例では、自動奥行き設定部310は、画像全体の視差量を一律1/4倍になるように補正を行ったが、視差量の変化が大きい領域だけを補正するように構成されていてもよい。例えば、主要被写体が写っている領域のみを補正してもよい。

[0125] また、デプスマップ全体の視差量から一定値を引いたデプスマップが示す状態になるように一方または両方の映像を水平方向に移動してもよい。ただし、視差量が負の値になる部分が発生する場合、その部分の視差量は0に維持するようにしてもよい。視差量の補正の方法は、視差量が変化するように補正する限り、どのような方法を用いてもよい。

[0126] また、シーンが切り替わる境界の後の画像フレームだけではなく、境界前の画像フレームを補正してもよい。あるいは、境界の後のシーンに含まれる最初の画像フレームについては、直前の画像フレームからの視差量の変化を小さくし、続く画像フレームから、本来の視差量に徐々に近づくように、段階的に視差量を補正してもよい。

[0127] 以下、図10Dを参照しながら、どの画像フレームが視差量の補正の対象となり得るかを簡単な例を挙げて説明する。図10Dに示されるように、ダイジェスト映像に含まれるシーン1とシーン2との境界の前後の画像フレームにおける被写体の飛び出し量の差が比較的大きい場合を想定する。図10D

には、各画像フレームの下に、主要被写体の飛び出し量を示す数値（距離情報）が例示されている。ここでは、一例として、基準位置に対して主要被写体が遠くにある画像フレームには正の数値が設定され、基準位置に対して主要被写体が近くにある画像フレームには負の数値が設定されている。すなわち、数値が大きいほど主要被写体が奥にあるように見え、数値が小さいほど主要被写体が大きく飛び出して見える。なお、距離情報の表現形式はこの例に限らず、適宜定めてよい。

[0128] 視差量の補正の対象となる画像フレームの選択の態様には、大別して以下の3つのケースがある。第1のケースでは、境界直後の画像フレーム、すなわち、シーン2の先頭の画像フレームから始まる連続する複数の画像フレームが補正の対象となる。このケースでは、境界後のいくつかの画像フレームの飛び出し量が小さくなるように、視差量が縮小する方向に補正される。ここで、どの程度の数の画像フレームを補正の対象とするかは、設計事項であり、適宜定めてよい。なお、何らかの理由でシーン2の先頭部分が削除されるような場合には、削除後の先頭の画像フレームを含む複数の画像フレームを補正すればよい。この場合、削除後の先頭の画像フレームが境界直後の画像フレームと解釈される。

[0129] 第2のケースでは、シーン1の最後の画像フレームを含む境界前の連続する複数の画像フレームが視差の補正の対象となる。このケースでは、境界前のいくつかの画像フレームの飛び出し量が大きくなるように、視差量が拡大する方向に補正される。なお、この場合も、何らかの理由でシーン1の最後の部分が削除されるような場合には、削除後の最後の画像フレームを含む複数の画像フレームを補正すればよい。この場合、削除後の最後の画像フレームが境界直前の画像フレームと解釈される。

[0130] 第3のケースでは、シーン1における最後の画像フレームを含む連続する複数の画像フレームと、シーン2における先頭の画像フレームを含む連続する複数の画像フレームとが視差の補正の対象となる。このケースでは、境界前のいくつかの画像フレームについては、飛び出し量が大きくなるように視

差量が拡大する方向に補正され、境界後のいくつかの画像フレームについては、飛び出し量が小さくなるように視差量が縮小する方向に補正される。なお、この場合も、何らかの理由でシーン1の最後の部分またはシーン2の先頭の部分が削除されるような場合には、削除後の境界前後の複数の画像フレームを補正すればよい。

[0131] このように、本実施形態では、連続する2つのシーンの境界の直前または直後の画像フレームを含む連続する複数の画像フレームが補正の対象となる。なお、この例では、境界前後の飛び出し量の変化が緩和されるように補正対象の画像フレームの視差量が補正されるが、上述したとおり、境界前後の飛び出し量の変化を大きくするように補正される場合もある。その場合でも、どの画像フレームを補正の対象とするかは、上記の3つのケースがあり得る。

[0132] 以上のように、本実施形態における自動奥行き設定部310は、飛び出し量を示す情報に基づいて、ダイジェスト再生のために抽出された複数のシーンのうち、連続する2つのシーンの境界前後の被写体の飛び出し量の変化を緩和または強調させるように、2つのシーンに含まれる少なくとも一部の画像フレームにおける視差量を補正したダイジェスト映像を生成する。なお、被写体の距離情報あるいは飛び出し量（奥行き量）の変化量を予め定められた閾値と比較するのは一例であり、必ずしも閾値と比較する必要はない。

[0133] 上記の例では、自動奥行き設定部310は、連続する2つのシーンの境界前後における一部の画像フレームの視差量を調整したが、さらに2つのシーンの境界にトランジション映像を挿入してもよい。ここで「トランジション映像」とは、連続する2つのシーンを連結するための部分映像であり、例えばワイプ、ズーム、ディゾルブといったトランジション効果を含む。このようなトランジション映像を境界に挿入することにより、急激な飛び出し量の変化によって視聴者を驚かせたり、頭痛等を引き起こすおそれを低減させることができる。この場合、自動奥行き設定部310は、トランジション映像挿入部としての機能を有する。

[0134] 図10Eは、いわゆるフェードイン、フェードアウトなどのトランジション映像の挿入の一例を示す図である。図示されるように、境界直前の画像フレームF1（シーン1の最後のフレーム）と、境界直後の画像フレームF2（シーン2の先頭のフレーム）との間に、トランジション映像が挿入される。トランジション映像の長さは、例えば0.3秒～数秒程度に設定される。トランジション映像は、フレームF1とフレームF2とを連結するための種々の演出が付加され得る。図10Eに示す例では、フレームF1およびフレームF2の画像の少なくとも一部を混合したような複数のフレームからなる部分映像がトランジション映像として挿入されている。このトランジション映像は、最初はフレームF1の内容に近く、徐々にフレームF2の内容に近づくように、フレームF1、F2の構成比率が徐々に変化する映像である。このようなトランジション映像を挿入すれば、本実施形態における急激な飛び出し量の変化の緩和効果が一層効果的になる。

[0135] 自動奥行き設定部310はまた、以下の第1のモードおよび第2のモードのいずれかで立体映像を調整する機能を有していてもよい。第1のモードは、立体映像の安全性を担保するため、個々の画像フレームにおいて、最大の飛び出し量をもつ部分をディスプレイ面と同程度の距離に設定し、その他の部分がディスプレイ面よりも奥に見えるように立体映像を調整するモードである。第2のモードは、立体映像の臨場感を高めるため、最大の飛び出し量をもつ部分をディスプレイ面よりも手前の距離に設定し、その他の部分がディスプレイ面よりも奥に見えるように立体映像を調整するモードである。自動奥行き設定部310は、例えば撮影条件や映像特性に応じて、第1のモードと第2のモードとを自動的に切り替えることができる。例えば、スポーツ観戦中や運動をしながら、または移動体に乗りながら撮影をする場合のように、画像の振れが比較的大きい場合には自動的に第1のモードに設定し、三脚を用いた安定した撮影状態においては第2のモードに設定する、といった制御が可能である。なお、このような制御は、ダイジェスト映像のシーンの境界のみならず、シーン全般にわたって行われてもよい。

[0136] [1-2-4. 属性情報および評価値の入力]

続いて、ユーザによる属性情報および評価値の入力について説明する。本実施形態におけるビデオカメラ100aでは、テーブルの属性情報および評価値をユーザが1つ1つ具体的に入力することができる。例えば、前述の説明における人物Aの顔検出および人物Bの顔検出は、ユーザが追加で設定したものである。すなわち、ユーザが、ビデオカメラ100aが予め備えていたテーブルに、人物Aの顔検出を新たな属性情報として登録し、その評価値も登録し、さらに、別の人物Bの顔検出を新たな属性情報として登録し、その評価値も登録している。これにより、単に人物の顔検出というだけでなく、さらに詳細に特定の人物の顔検出を行うことができるようになる。

[0137] ここで、記録部330には、顔認識テーブルが予め記録されている。顔認識テーブルは、IDと顔画像と人物の名前と評価値とを1セットとして、これを複数セット（例えば、6セット）だけ登録できるように構成されている。そして、IDと評価値は予め設定されており、特定の人物の顔および名前をユーザが適宜登録していく。

[0138] 具体的には、外部入力部207の操作によりビデオカメラ100aが登録モードに設定されると、制御部300は、登録すべき人物の顔を所定時間（例えば、3秒）以上撮影することをユーザに促すメッセージを映像表示部212に表示させる。それに応じて、ユーザがビデオカメラ100aで特定の人物の顔を所定時間以上撮影すると、当該人物の顔を登録すべき新たな属性情報として認識し、記録部330に一時的に保存する。その後、制御部300は、当該人物の顔を顔認識テーブルのどのIDに対応させるかを尋ねるメッセージを映像表示部212に表示させる。なお、IDごとに評価値が既に設定されており、例えば、ID1～ID6のうちでは、ID1の評価値が「100」で最も高く、IDの番号が大きくなるほど、評価値は小さくなるように設定されている。つまり、登録すべき人物の顔とIDとを対応させることは、当該人物の顔の評価値を設定することに等しい。ユーザが対応するIDを選択すると、次に、制御部300は、当該人物の名前の入力を促すメッ

ページを映像表示部 212 に表示させる。こうして、顔認識テーブルに、ID と特定の人物の顔画像と名前と評価値が設定される。なお、この例では、評価値が予め設定されているが、評価値もユーザが任意に入力できるように構成してもよい。その場合には、評価値の入力をユーザに促すメッセージを映像表示部 212 に表示させて、評価値をユーザに入力させるようにすればよい。このようにして、属性情報および評価の 1 つ 1 つの内容をユーザが任意に設定することができる。

[0139] そして、このように設定された属性情報および評価の対応データは、ユーザの選択に基づいて、様々な使われ方をする。例えば、上記の例では、人物 A の顔、人物 B の顔、それ以外の人物の顔がそれぞれ検出されたことを属性情報として設定しているが、検出された顔が誰の顔かを識別せずに、単に人物の顔が検出されたことを属性情報として抽出することもできる。すなわち、ビデオカメラ 100a は、人物を特定せずに、単に顔が検出されたことを属性情報として抽出する通常モードと、特定の人物の顔が検出されたことを属性情報として抽出する特定モードとを有する。特定モードでは、登録された人物の顔のうち、属性情報として抽出する顔を選択することもできる。具体的には、制御部 300 は、登録されている人物の顔画像、名前または ID を映像表示部 212 に表示させる。それに対して、ユーザが映像表示部 212 を操作して属性情報として抽出する人物の顔を選択する。制御部 300 は、選択された人物の顔を属性情報として抽出する。

[0140] さらに、通常モードと特定モードとでは、顔検出を属性情報として抽出する際の条件を変更してもよい。すなわち、通常モードでは、映像中に不特定のある人物の顔が所定の時間以上撮影されていることをもって属性情報として抽出する。それに対して、特定モードでは、撮影中に特定の人物（例えば、人物 A）の顔が上記の時間よりも短い所定の時間以上撮影されている（例えば、1 フレームだけに撮影されている）ことをもって属性情報として抽出する。つまり、通常モードでは、風景等に比べて人物の方が一般的に撮影対象としての重要度が高いという考えから、顔が検出されることを属性情報と

して設定している。そのため、人物の顔が一瞬だけ撮影されたというのではなく、撮影対象として人物の顔が撮影されたと判定できる程度の時間だけ、人物の顔が撮影されることを必要条件としている。それに対して、特定モードは、風景等の他の撮影対象と特定の人物の顔を比べて重要かどうかというよりも、特定の人物の顔が写っている映像を抽出したいというユーザの明確で強い意図が反映されたモードである。そのため、特定モードにおける顔検出の重要度は、通常モードに比べて高い。そこで、特定顔検出モードでは、通所の顔検出モードよりも、顔が検出されたと認定する条件を緩くしている。

[0141] なお、顔検出の条件を変えずに、特定の人物の顔の評価値を、不特定の顔の評価値よりも高くすることによって、特定の人物の顔の重要度を高くしてもよい。

[0142] 続いて、人物Aの顔検出よりも人物Bの顔検出の評価値を高く設定した場合の例を説明する。図11は、映像を評価する際に用いる、各種の属性情報と各属性情報に対する評価との対応データのテーブルである。図12は、図11のテーブルに基づいて、シーン評価部306が、ある映像から属性情報を抽出して評価値を付与した結果を示した図である。図12の横軸は時間（シーン）を、縦軸に各シーンの評価値を示す。図13は、図11のテーブルに基づく評価から生成された再生情報を示す。

[0143] 図11に示すように、人物Aの顔検出の属性情報の評価値が「60」であるのに対し、人物Bの顔検出の属性情報の評価値が「90」となっている。そして、このテーブルを用いて映像を評価すると、図12に示す結果となる。具体的には、図5のテーブルを用いた評価と比べ、シーン#2の評価が下がり、シーン#4の評価が上がっている。この評価値に基づいて再生情報を生成すると、図13に示すように、図8におけるシーン#2に代わって、シーン#4がダイジェストに追加されている。

[0144] 前述のような評価値の変更は、テーブルの評価値をユーザが書き換えてもよいし、評価値が異なるテーブル（図5のテーブルと図11のテーブル）を

予め用意しておき、それを切り替えるようにしてもよい。例えば、予め用意しておいたテーブルを切り替える方法としては、各種のテーブルに対応するモードをユーザに選択させる選択画面（図14参照）を映像表示部212に表示させ、ユーザに外部入力部207を介してモードを選択させるようにしてもよい。あるいは、各種のテーブルを映像表示部212に表示させて、ユーザにテーブルを選択させるようにしてもよい。なお、予め用意しておくテーブルは、属性情報または評価値をユーザが直接入力することによって予め作成されたものであってもよい。さらに、映像解析部303によって解析された、映像の構図情報、映像に含まれる被写体および背景、また、被写体および背景の位置、大きさ、カラー、テクスチャなどの情報を用いて映像の属性情報や評価値を作成することもできる。

[0145] このように、属性情報の評価をユーザの嗜好に応じて変更することによって、そのユーザの嗜好に応じたダイジェストを生成することができる。なお、このような属性情報の評価の変更の機能はあくまでも一例であり、必須の機能ではない。ビデオカメラ100aは、立体映像のダイジェスト再生の際、シーンの切り替わり部において主要被写体の視差量を調整する機構を備えていれば、どのように構成されていてもよい。

[0146] [3. 効果等]

以上のように、本実施形態のビデオカメラ100aは、撮影によって取得された立体映像の中からダイジェスト再生のために抽出された複数のシーンを特定する情報、および当該複数のシーンに含まれる被写体の飛び出し量を示す情報を取得するインタフェースとして機能するダイジェスト再生部309と、当該複数のシーンを特定する情報および飛び出し量を示す情報に基づいて、当該複数のシーンのうち、連続する2つのシーンの境界の直前または直後の画像フレームを含む連続する複数の画像フレームにおける視差量を補正してダイジェスト映像を生成するダイジェスト映像生成部としての機能を有する自動奥行き設定部310とを備える。

[0147] これにより、立体映像のダイジェスト再生の際、シーンの切り替わり部に

において被写体の視差量を適切に調整することができる。その結果、シーンの切り替わりで急激な飛び出しや引き込みが発生することを防いだり、一時的に立体特性を高めて映像を強調させることができるため、より安全で臨場感のある立体映像のダイジェスト再生が可能になる。また、属性情報および評価の少なくとも一方をユーザが任意に入力可能とすることによって、ユーザの嗜好を反映したダイジェスト再生を行うこともできる。特に、本実施形態では、属性情報をユーザが追加で設定できるため、よりユーザの嗜好に合った映像をダイジェストとして再生することができる。

[0148] また、本実施形態のある態様において、自動奥行き設定部310は、飛び出し量を示す情報に基づいて、連続する2つのシーンの境界直前の画像フレームから境界直後の画像フレームへ遷移させるためのトランジション映像を当該2つのシーンの境界に挿入するトランジション映像挿入部としての機能も有する。これにより、2つのシーンの境界における飛び出し量の急激な変化の緩和効果が一層効果的になる。

[0149] また、本実施形態では、ビデオカメラ100aは、ユーザの入力操作に応じて、映像に関する属性情報を入力する外部入力部207と、映像の中からダイジェスト再生する部分を抽出するために、前記属性情報を映像の中から抽出する再生情報生成部307とを備えている。これにより、映像の中からダイジェスト再生する部分を抽出するために用いられる属性情報をユーザが適宜入力できる。その結果、ユーザの嗜好に合った映像をダイジェストとして再生することができる。

[0150] また、ビデオカメラ100aは、映像に関する属性情報および当該属性情報に対する評価との対応データについて、当該属性情報および評価の少なくとも一方をユーザの入力操作に応じて入力する外部入力部207と、映像の中からダイジェスト再生をする部分を抽出するために、当該属性情報を映像の中から抽出し、前記対応データに基づいて当該属性情報を有する部分を評価する属性情報抽出部305とを備えている。これにより、映像の中からダイジェスト再生する部分を抽出するために用いられる属性情報および／また

はその評価値をユーザが適宜入力できる。その結果、ユーザの嗜好に合った映像をダイジェストとして再生することができる。

[0151] また、予め設定された顔検出（人物を特定しない顔検出）という属性情報に対して、特定の人物の顔検出という、下位概念の属性情報を設定可能とすることによって、ユーザのより深い嗜好を反映したダイジェスト再生を行うことができる。

[0152] さらに、属性情報には、クリップイン、クリップアウトおよびズームアップ等のようなユーザの嗜好が現れ難い属性情報については固定の属性情報と、特定の人物の顔検出のようにユーザの嗜好が現れ易い属性情報については可変の属性情報とすることによって、処理を簡潔にすることができる。つまり、全ての属性情報を可変とすると、様々な属性情報に対応した制御内容（属性情報の入力や属性情報の抽出等）やメモリ容量を準備しておかなければならず、処理が煩雑となってしまふ。それに対して、可変となる属性情報がある程度絞っておくことによって、予め準備しておく制御内容やメモリ容量を少なくすることができ、処理が簡潔になる。また、自身の嗜好を強く反映させたいユーザと、自身の嗜好をあまり反映させたくないユーザ（そのような操作を煩わしいと感じるユーザ）とが存在する。そこで、全ての属性情報を可変とするのではなく、いくつかの属性情報は固定として、固定の属性情報だけに基づくダイジェスト再生を選択可能とすることによって、後者のユーザに対しては、操作の煩わしさを解消し、使いやすいビデオカメラを提供することができる。一方、前者のユーザにとっては、属性情報の一部を嗜好に応じて入力（追加、変更）することができるため、嗜好を反映させたダイジェスト再生を可能とすることができる。つまり、様々なユーザに対して、使い勝手がよく且つ嗜好を反映させたダイジェスト再生が可能なビデオカメラを提供することができる。

[0153] また、ビデオカメラ100aでは、撮影時に属性情報の抽出、シーン評価、再生情報の生成を行うため、ダイジェスト再生時の処理を低減して、ダイジェスト再生を簡潔且つ迅速に実行することができる。また、ビデオカメラ

100aの姿勢等のような属性情報は、映像から事後的に判断することが難しく、または、可能であるが煩わしく、撮影時のセンサの検出信号等により容易に検出することができる。つまり、属性情報の中には、撮影時の方が検出しやすいものがある。そのため、撮影時に属性情報の抽出を行うことによって、このような属性情報を容易に抽出することができる。

[0154] さらに、ビデオカメラ100aのような撮影装置に記録された映像は、編集等がなされていない、単に撮影しただけの映像である。そのため、重要度の低い映像も多く、前述のようなダイジェスト再生は非常に有効である。

[0155] (実施形態2)

続いて、図15を参照しながら実施形態2におけるビデオカメラ100bの構成および動作を説明する。

[0156] 図15は、ビデオカメラ100bの概略構成を示すブロック図である。ビデオカメラ100bは、立体映像を記録する前に再生すべきシーンを選択するのではなく、記録した後、ダイジェスト再生を行う際に再生すべきシーンを選択する点で実施形態1のビデオカメラ100aと異なる。具体的には、ビデオカメラ100bの基本的な構成は、ビデオカメラ100aとほぼ同様であるが、データの流れ、即ち、処理の順序が実施形態1のビデオカメラ100aと異なる。そこで、実施形態1と同様の構成要素については、同様の符号を付して説明を省略し、異なる部分を中心に説明する。

[0157] 属性情報抽出部305が属性情報を抽出するまでの処理は、実施形態1と同様である。本実施形態では、属性情報抽出部305によって抽出された属性情報は、多重化部308に入力される。多重化部308は、映像信号圧縮部204から出力される符号化映像データ、音声信号圧縮部210から出力される符号化音声データ、および属性情報抽出部305から出力される属性情報を多重化して出力する。この多重化されたデータは、記録部330に記録される。ここで、属性情報は、映像の部分ごとに当該部分と関連付けられて記録される。例えば、映像データを構成するフレームごと、または連続する複数のフレームのまとまりごとに属性情報が関連付けられて記録される。

- [0158] シーン評価部306は、ダイジェスト再生が行われるときに、記録部330から多重化されたデータを読み出し、属性情報に基づいて、映像の各部分に評価を付与し、映像の中から特徴的なシーンを抽出し、その中から、ダイジェスト再生すべきシーンをさらに抽出する。
- [0159] その後、再生情報生成部307は、シーン評価部306が抽出したシーンに基づいて、再生情報を生成して、ダイジェスト再生部309へ出力する。
- [0160] ダイジェスト再生部309は、再生情報生成部307が生成した再生情報に基づいて、該当するデータを記録部330から読み出し、ダイジェスト再生用の立体映像を生成し、自動奥行き設定部310に出力する。自動奥行き設定部310は、必要に応じて立体映像の飛び出し量を調整する処理を行う。そして、映像信号伸張部211および音声信号伸張部213へ出力する。そして、ダイジェスト映像が、映像表示部212および音声出力部214によって再生される。
- [0161] 図16は、撮影、属性情報の抽出、記録までの処理の流れを示すフローチャートである。図17は、ダイジェスト再生時の処理の流れを示すフローチャートである。
- [0162] まず、ビデオカメラ100bの制御部300は、ステップS301において、撮影を開始する。入力ボタン等の外部入力部207からの入力に基づいて、撮影が開始される。
- [0163] 次に、ステップS302において、属性情報抽出部305は、姿勢検出部206の検出結果、レンズ制御部301の制御情報、並びに映像解析部303および音声解析部304の解析結果等に基づいて映像の属性情報を抽出する。そして、ステップS303において、多重化部314は、属性情報を、符号化映像データ、符号化音声データと共に多重化する。ステップS304において、制御部300は、これらの多重化データを記録部330に記録する。
- [0164] その後、制御部300は、ステップS106において、外部入力部207から、撮影終了の入力があるか否かを判定する。撮影終了の入力がない場合

は、ステップS302へ戻って、撮影を継続する。一方、撮影終了の入力がある場合には、撮影を終了する。

[0165] このように、本実施形態においては、撮影した映像に対応する再生情報を生成することなく、多重化されたデータが記録部330に記録される。

[0166] 撮影が終了した後、ダイジェスト再生を実行する際には、図17に示すように、ダイジェスト再生すべきシーンの抽出や再生情報の生成が行われる。

[0167] まず、外部入力部207を介してユーザからダイジェスト再生の入力があると、シーン評価部306は、ステップS401において、記録部330に記録された多重化データを読み出す。

[0168] 次に、シーン評価部306は、ステップS402において、読み出した多重化データを分解し、属性情報を読み出す。続いて、シーン評価部306は、ステップS403において、属性情報に基づいて映像の各部分に評価値を付与する。そして、シーン評価部306は、ステップS404において、映像の全部分の評価が終了したか否かを判定する。終了していなければ、シーン評価部306は、ステップS401に戻り、映像の評価を継続する。一方、評価が終了していれば、シーン評価部306は、ステップS405へ進む。

[0169] ステップS405では、シーン評価部306は、評価値に基づいて映像の中からいくつかの特徴的なシーンを抽出し、さらに、その中からダイジェスト再生すべきシーンを抽出する。そして、再生情報生成部307は、シーン評価部306により抽出されたダイジェスト再生すべきシーンに基づいて再生情報を生成する。

[0170] 続いて、ダイジェスト再生部309は、ステップS406において、再生情報に基づいてダイジェスト再生のための複数のシーンを選択する。

[0171] 次に、自動奥行き設定部310は、ステップS407において、選択した各シーン間の被写体距離の変化量を求める。そして、ステップS408において、直前のシーンに対する被写体距離の変化量が予め定められた閾値を超えるシーンが存在するか否かを判定する。該当するシーンが存在する場合に

は、ステップS409に進み、存在しない場合には、ステップS410に進む。

[0172] ステップS409では、自動奥行き設定部310は、直前のシーンに対する被写体距離の変化量が予め定められた閾値を超えると判定されたシーンにおいて、当該閾値を超えないように当該シーンの映像における視差量を補正する。視差量の補正の方法としては、実施形態1と同様の方法が用いられる。視差量の補正が完了すると、処理はステップS410に進む。

[0173] 自動奥行き設定部310は、ステップS410において、再生すべき符号化映像データおよび符号化音声データを映像信号伸張部211および音声信号伸張部213に出力し、映像表示部212および音声出力部214を介して映像および音声を再生する。

[0174] 以上の処理により、撮影した映像、音声を記録部330に一度記録した後に、ダイジェスト再生に必要な再生情報を生成し、ダイジェスト再生を行うことが可能となる。本実施形態によれば、撮影後、ダイジェスト再生を実行する際に、属性情報に対する評価値を変更することができる。さらに、実施形態1と同様、安全で臨場感のある立体映像のダイジェスト再生が可能になる。

[0175] (他の実施形態)

以上のように、本出願において開示する技術の例示として、実施形態1、2を説明した。しかしながら、本開示における技術は、これに限定されず、適宜、変更、置き換え、付加、省略などを行った実施の形態にも適用可能である。また、上記実施形態1、2で説明した各構成要素を組み合わせ、新たな実施の形態とすることも可能である。

[0176] そこで、以下、他の実施の形態を例示する。

[0177] 図18は、他の実施形態におけるビデオカメラ100cおよび再生装置500の概略構成を示すブロック図である。この実施形態では、シーン評価、再生情報の生成、およびダイジェスト再生の機能を、ビデオカメラ100cとは異なる再生装置500に実行させる点が実施形態1、2と異なっている

。再生装置は、例えばパーソナルコンピュータ（PC）やクラウドサービスを提供するサーバコンピュータ等の映像編集装置であり得る。

[0178] 図18に示すように、本実施形態におけるビデオカメラ100cは、撮影した映像から、属性情報を抽出するところまでを実行する。そして、再生装置500は、属性情報が付与された映像データに基づいて、シーン評価およびダイジェスト再生を実行する。このように構成すれば、再生装置500側でユーザがシーン評価値の設定ができるため、よりユーザの嗜好に沿ったダイジェスト再生が可能になる。

[0179] なお、図18に示す構成の他、ビデオカメラがシーン評価を行うところまでを実行し、再生装置が評価値に基づいてダイジェスト再生を実行するように構成してもよい。

[0180] 上記の各実施形態では、立体映像は左右の2チャンネル映像から構成され、飛び出し量の調整は、左右の2チャンネル映像の視差量を調整することによって実現されるが、本開示における技術はこのような形態に限定されない。例えば、立体映像は、3個以上の多視点映像から構成されていてもよい。立体映像が多視点映像から構成される場合、多視点映像の演算生成ユニットでDIBR（Depth Image Based Rendering）などの方法による飛び出し量の制御を行うようにすれば、上記の各実施形態と同様、飛び出し量が適正に調整された好ましい立体映像を生成することができる。

[0181] 上記の実施形態では、人物の顔検出を上位の属性情報および固定の属性情報とし、特定の人物の顔検出を下位の属性情報および可変の属性情報としているが、これに限られるものではない。例えば、犬などの動物の顔検出を上位の属性情報および固定の属性情報とし、特定の犬の顔検出を下位の属性情報および可変の属性情報としてユーザが入力可能としてもよい。または、電車、車または飛行機などの交通手段の検出を上位の属性情報および固定の属性情報とし、特定の電車、車または飛行機の検出を下位の属性情報および可変の属性情報としてユーザが入力可能としてもよい。または、人物の声の検

出を上位の属性情報および固定の属性情報とし、特定の人物の声の検出を下位の属性情報および可変の属性情報としてユーザが入力可能としてもよい。この場合、特定の人物の声をマイクロフォン208を介してビデオカメラに入力することができる。

[0182] 図19は、さらに他の実施形態におけるビデオカメラ100dおよび再生装置501の概略構成を示すブロック図である。この実施形態では、ビデオカメラ100dは、2組ではなく1組の撮像系（レンズ群200、撮像素子201、および映像AD変換部202）を有している点で、図18に示すビデオカメラ100cとは異なっている。また、再生装置501は、自動奥行き設定部310を有していない点を除いて、図18に示す再生装置500と同様の構成を有している。

[0183] ビデオカメラ100dは、被写体までの距離を測定するための測距部220を備えている。測距部220は、特定の発光パターンで被写体に光を照射する光源223と、その被写体からの反射光を検出するセンサ221とを有している。測距部220は、撮像制御部302からの制御信号に基づき、映像を構成する個々のフレームの生成に同期して被写体の距離を測定し、測定した距離を示す情報（距離情報）を撮像制御部302に送る。ここで、距離情報は、フレームごとに関連付けられて送られる。撮像制御部302は、送られてきた距離情報を、属性情報抽出部305に送る。これにより、属性情報抽出部305は、被写体の距離情報に基づく属性情報を生成することができる。

[0184] 図20A、図20Bは、本実施形態におけるビデオカメラ100dによる撮影中の状況を模式的に示す図である。このビデオカメラ100dは、監視カメラとして用いられ得る。図20Aに示されるように、ビデオカメラ100dから所定の距離Lの位置に禁止線420が設けられている場合を想定する。これは、例えば美術館のように、展示物にある一定の距離よりも近づくことが禁止されている場所において、展示物の付近に監視カメラが設置されている場合に相当する。あるいは、監視対象の建造物に近づく不審な人物や

車両を監視するための監視カメラが当該建造物に設置されている場合に相当する。ビデオカメラ100dが映像の記録を行っている最中に、図20Bに示すように人物400が禁止線420を超えて近づいたことを検出すると、ビデオカメラ100dの属性情報抽出部305は、そのことを示す属性情報を生成し、対応するフレームに関連付けて多重化部308に送る。これにより、人物400が近づいてきたシーンのみからなるダイジェスト映像を事後的に生成することができる。ビデオカメラ100dを監視カメラとして利用する場合、撮影中に図20Bに示すように、被写体が禁止線420を超えて近づいてきたときにスピーカー208が警告音を鳴らすようにしてもよい。

[0185] 図21は、本実施形態における処理の一例を示すフローチャートである。まずステップS211において、ビデオカメラ100dは、フレームのデータの生成および被写体までの距離の測定を行う。次に、ステップS212において、被写体までの距離が所定の閾値（上記の例では禁止線までの距離L）よりも短いかなかを判定する。この判定結果がYesである場合、ステップS213に進み、そのフレームのデータに被写体が近づきすぎていることを示す属性情報を付与する。ステップS212における判定結果がNoである場合、ステップS213は省略される。次に、ステップS214において撮影の終了の指示が出されているかを判定する。この判定結果がNoである場合、次のフレームについてステップS211からステップS214までの動作を繰り返す。以降、ステップS214の判定結果がYesになるまで同様の処理を繰り返す。

[0186] 一方、再生装置500は、ビデオカメラ100dによって記録された属性情報に基づいて、被写体が近くに写っているシーンのみを抽出してダイジェスト映像を生成することができる。図22は、ビデオカメラ100dによって記録された元映像から、被写体が禁止線420を超えたシーンのみを抽出してダイジェスト映像を生成したことを示す図である。図22の例では、時系列に並んだ映像を構成するフレームF1～F7のうち、被写体までの距離が閾値よりも短いと判定されたフレームF3～F6が抽出されている。再生

装置500がこのような処理を行うことにより、被写体が近いシーンのみを再生することができる。

[0187] 監視カメラでは、通常、長時間の撮影が行われるため、記録される映像の全てを確認するのに多くの時間を要する。そのような長い映像から不審な人物等が写ったシーンを人手を介して抽出することは困難である。本実施形態によれば、長時間の映像から、自動的に被写体が比較的近いシーンのみを抽出することができるため、従来にない利便性の高い監視カメラを実現することができる。

[0188] なお、図20A、20B、22では、被写体の例として人物が描かれているが、車両や動物のように、人物以外の動く被写体がカメラに近づいてきた場合も同様の処理が行われ得る。動く被写体の検出は、公知の動き検出技術を用いて実行され得る。特に人物を検出する場合は、公知の顔認識技術を利用してもよい。また、被写体が近いシーンのみならず、被写体が遠いシーンや、被写体までの距離が所定の条件に合致するシーンのみを抽出するようにしてもよい。図5を参照して説明したように、複数の閾値を設けて被写体距離の程度を多段階に評価できるようにしてもよい。

[0189] 本実施形態における再生装置501は、図23に示すように、被写体が小さく写っているシーンにおいて、被写体を拡大して表示するようにしてもよい。図23では、画面内に別の枠440を設け、そこに被写体を拡大表示する例を示しているが、別枠440を設けることなく当該被写体を拡大表示してもよい。この機能を実現するためには、属性情報抽出部305が、図21に示す処理とは逆に、被写体までの距離が所定の閾値よりも大きい場合に被写体が遠いことを示す属性情報をフレームデータに付与するように構成されていればよい。再生装置501は、拡大表示とともに、または拡大表示に代えて、被写体が目立つように赤色等の枠450で被写体を囲む処理を行ってもよい。

[0190] なお、本実施形態では、再生装置501が、シーン評価部306、再生情報生成部307、ダイジェスト再生部309を有しているが、実施形態1、

2のように、ビデオカメラ100dの制御部300がこれらを有していてもよい。その場合、ビデオカメラ100d単独でダイジェスト動画の生成を行うことが可能になる。

[0191] また、本実施形態のビデオカメラ100dは、監視カメラに限らず、一般のビデオカメラであってもよい。その場合、例えば公知の顔認識技術を利用して、特定の人物が比較的近いシーンのみを含むようなダイジェスト映像を生成してもよい。

[0192] 上記の各実施形態における処理は、ビデオカメラのような撮影装置に限られず、例えばPCやサーバコンピュータなどの映像編集装置や、ハードディスク(HDD)レコーダ、デジタル多用途ディスク(DVD)／ブルーレイディスク(BD)レコーダのような映像記録再生装置に適用することもできる。このような映像編集装置や映像記録再生装置が、入力された映像の中から属性情報を抽出し、属性情報に基づいて映像の各部分を評価し、その評価結果に基づいてダイジェスト再生を行うようにしてもよい。あるいは、放送局等において放送用映像の制作に用いられるスイッチャーに上記の各実施形態における処理を適用することもできる。映像を編集するスイッチャーに上記実施形態の処理を適用した場合、人の手を介することなく、自動で好ましいダイジェスト映像の制作が可能となるため、編集時間を大幅に削減することができる。

[0193] また、記録機能、シーン抽出機能、再生機能を有しない映像処理装置を用いて、ダイジェスト映像の生成を行ってもよい。そのような映像処理装置の一例は、ダイジェスト再生に用いられる複数のシーンを特定する情報と、それらのシーンに含まれる被写体の視差量を示す情報とを、記録媒体または電気通信回線を通じて取得することにより、視差量が適切に補正されたダイジェスト映像を生成することができる。また、映像処理装置の他の例は、撮像系から被写体までの距離が予め定められた条件を満たすことを示す属性情報に基づいて、映像の中からダイジェスト再生に用いられる複数のシーンを特定する情報を生成する再生情報生成部と、複数のシーンを特定する情報に基

づいて、当該複数のシーンから構成されるダイジェスト映像を生成するダイジェスト映像生成部とを備える。これにより、例えば被写体が近いシーンのみ、または被写体が所定の距離の範囲内に位置するシーンのみを抽出することができる。これらの映像処理装置によって生成されたダイジェスト映像は、可搬型の記録媒体に記録され、あるいは電気通信回線を通じて他の再生装置に送られることにより、他の再生装置によって再生することが可能である。

[0194] また、ビデオカメラの構成は、上記の実施形態に限定されるものではない。例えば、映像AD変換部202R、202L、信号処理部203、映像信号圧縮部204、音声AD変換部209、音声信号圧縮部210、映像信号伸張部211および音声像信号伸張部213の全部またはいくつかを単一の集積回路として実現することも可能である。また、制御部300が実行する処理の一部を別途、FPGA (Field Programmable Gate Array) を用いてハードウェアとして実現することも可能である。

[0195] また、上記の内容を実現する方法やソフトウェアプログラムとしても実現できることは言うまでもない。

[0196] 以上の実施形態は、あくまでも例示であって、本発明、その適用物、あるいはその用途の範囲を制限することを意図するものではない。映像処理装置、撮影装置、およびこれらの装置に用いられるプログラムは、ダイジェスト再生のために抽出した複数のシーンの切り替わりの前後で視差量を変化させる機能を備えている限り、どのように構成されていてもよい。

[0197] 以上のように、本開示における技術の例示として、実施の形態を説明した。そのために、添付図面および詳細な説明を提供した。

[0198] したがって、添付図面および詳細な説明に記載された構成要素の中には、課題解決のために必須な構成要素だけでなく、上記技術を例示するために、課題解決のためには必須でない構成要素も含まれ得る。そのため、それらの必須ではない構成要素が添付図面や詳細な説明に記載されていることをもっ

て、直ちに、それらの必須ではない構成要素が必須であるとの認定をするべきではない。

[0199] また、上述の実施の形態は、本開示における技術を例示するためのものであるから、特許請求の範囲またはその均等の範囲において種々の変更、置き換え、付加、省略などを行うことができる。

産業上の利用可能性

[0200] ここに開示された技術は、映像の中からダイジェスト再生する部分を抽出する映像抽出装置およびそれを備えた撮影装置および再生装置に有用である。

符号の説明

[0201] 100 a、100 b、100 c、100 d ビデオカメラ
200、200 R、200 L レンズ群
201、201 R、201 L 撮像素子
202、202 R、202 L 映像AD変換部
203 信号処理部
204 映像信号圧縮部
205 レンズ制御モジュール
206 姿勢検出部
207 外部入力部
208 スピーカー
209 音声AD変換部
210 音声信号圧縮部
211 映像信号伸長部
212 映像表示部（ディスプレイ）
213 音声信号伸長部
214 音声出力部
215 出カインタフェース
300 制御部

- 3 0 1 レンズ制御部
- 3 0 2 撮像制御部
- 3 0 3 映像解析部
- 3 0 4 音声解析部
- 3 0 5 属性情報抽出部
- 3 0 6 シーン評価部
- 3 0 7 再生情報生成部
- 3 0 8 多重化部
- 3 0 9 ダイジェスト再生部
- 3 1 0 自動奥行き設定部 3 1 0
- 3 2 0 クロック発振器
- 3 3 0 記録部
- 5 0 0 再生装置

請求の範囲

- [請求項1] 撮影によって取得された立体映像の中からダイジェスト再生に用いられる複数のシーンを特定する情報、および前記複数のシーンに含まれる被写体の視差量を示す情報を取得するインタフェースと、
- 前記複数のシーンを特定する情報および前記視差量を示す情報に基づいて、前記複数のシーンのうち、連続する2つのシーンの境界の直前または直後の画像フレームを含む連続する複数の画像フレームにおける視差量を補正してダイジェスト映像を生成するダイジェスト映像生成部と、
- を備える映像処理装置。
- [請求項2] 前記ダイジェスト映像生成部は、前記視差量を示す情報に基づいて、前記連続する2つのシーンの境界直前の画像フレームから境界直後の画像フレームへ遷移させるためのトランジション映像を前記2つのシーンの境界に挿入するトランジション映像挿入部を有している、請求項1に記載の映像処理装置。
- [請求項3] 前記ダイジェスト映像生成部は、連続する前記2つのシーンの境界で被写体の視差量の変化が予め定められた閾値以上である場合、前記視差量の変化が前記閾値よりも小さくなるように前記視差量を補正する、請求項1または2に記載の映像処理装置。
- [請求項4] 前記ダイジェスト映像生成部は、連続する前記2つのシーンの境界で被写体の視差量の変化が予め定められた閾値以上である場合、前記トランジション映像を挿入する、請求項2に記載の映像処理装置。
- [請求項5] 前記ダイジェスト映像生成部は、前記2つのシーンのうち、後のシーンにおける先頭から所定枚数の画像フレームにおける視差量を補正する、請求項1から4のいずれかに記載の映像処理装置。
- [請求項6] 前記ダイジェスト映像生成部は、前記2つのシーンのうち、先のシーンにおける最後から所定枚数の画像フレームにおける視差量を補正する、請求項1から5のいずれかに記載の映像処理装置。

- [請求項7] 前記立体映像を取得した撮像系から前記立体映像中の被写体までの距離に基づいて、前記立体映像の中からダイジェスト再生に用いられる前記複数のシーンを特定する情報を生成する再生情報生成部をさらに備える、請求項1から6のいずれかに記載の映像処理装置。
- [請求項8] 前記被写体は、人間の顔を含む、請求項1から7のいずれかに記載の映像処理装置。
- [請求項9] 前記ダイジェスト映像生成部によって生成された前記ダイジェスト映像を再生する再生部をさらに備える請求項1から8のいずれかに記載の映像処理装置。
- [請求項10] 請求項1から9のいずれかに記載の映像処理装置と、
撮影によって前記立体映像を取得する2つの撮像系と、
を備える撮影装置。
- [請求項11] 前記撮影装置から被写体までの距離を測定する測距部をさらに備え、
前記映像処理装置は、前記距離に基づいて、前記立体映像の中からダイジェスト再生に用いられる前記複数のシーンを特定する情報を生成する再生情報生成部を有している、
請求項10に記載の撮影装置。
- [請求項12] 前記距離が予め定められた条件を満たす場合に、前記条件を満たすことを示す属性情報を前記立体映像に含まれる対応する画像フレームと関連付けて記録する記録部と、
を備える請求項11に記載の撮影装置。
- [請求項13] コンピュータに対し、
撮影によって取得された立体映像の中からダイジェスト再生に用いられる複数のシーンを特定する情報、および前記複数のシーンに含まれる被写体の視差量を示す情報を取得するステップと、
前記複数のシーンを特定する情報および前記視差量を示す情報に基づいて、前記複数のシーンのうち、連続する2つのシーンの境界前後

の少なくとも一部の画像フレームにおける視差量を補正するステップと、

前記視差量が補正された前記複数のシーンによって構成されるダイジェスト映像を生成するステップと、
を実行させるプログラム。

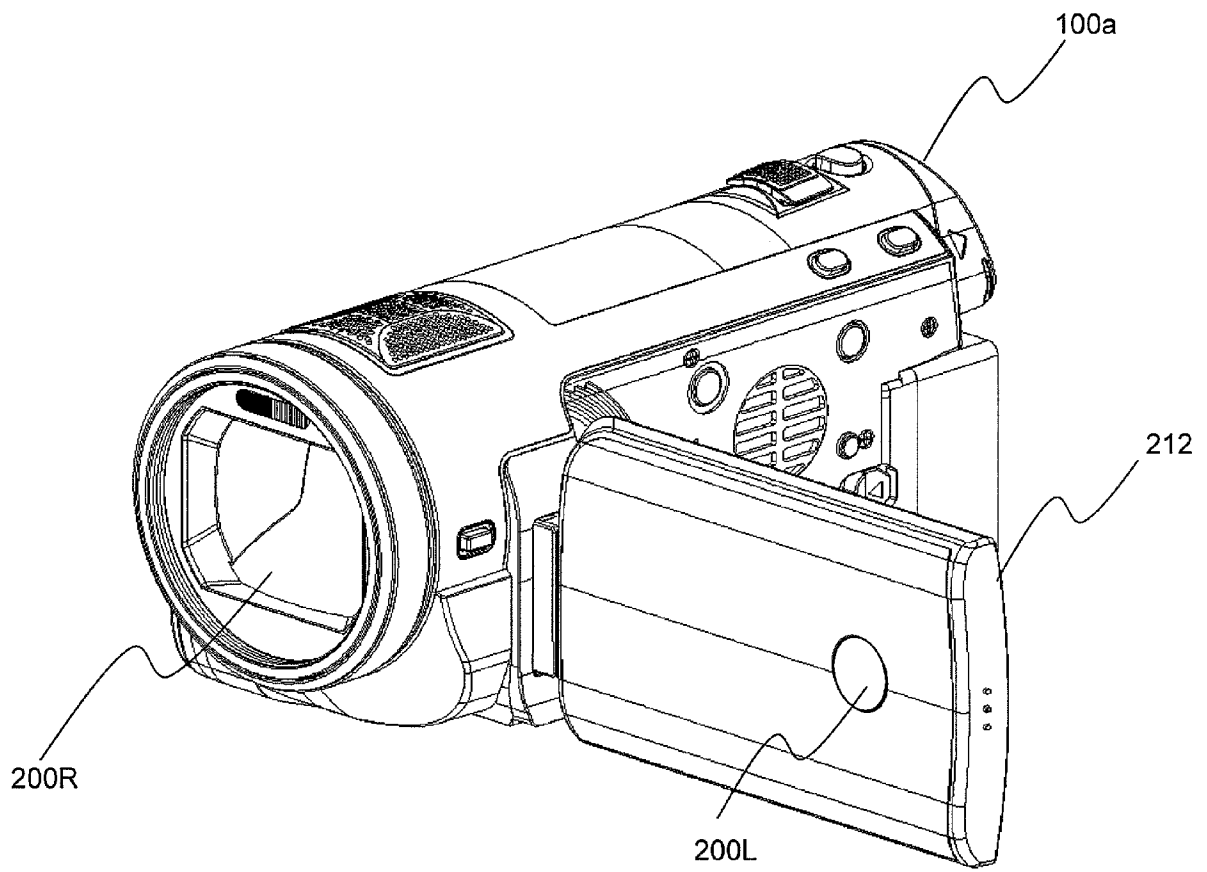
[請求項14]

撮影によって映像を取得する撮像系と、
前記撮像系から被写体までの距離を示す情報を取得する測距部と、
前記距離が予め定められた条件を満たす場合に、前記条件を満たすことを示す属性情報を前記映像に含まれる対応する画像フレームと関連付けて記録する記録部と、
を備える撮影装置。

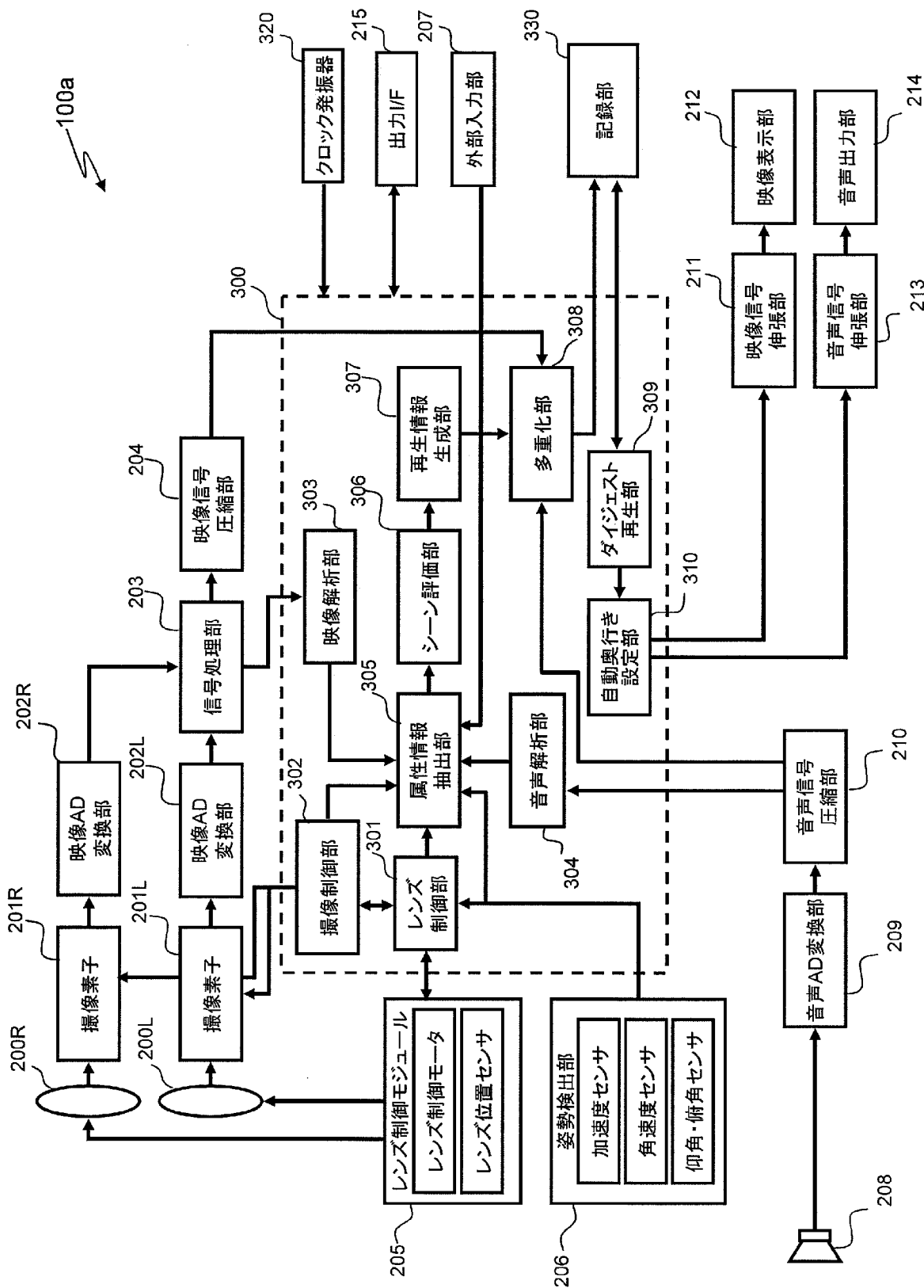
[請求項15]

請求項14に記載の撮影装置によって記録された前記属性情報に基づいて、前記映像の中からダイジェスト再生に用いられる複数のシーンを特定する情報を生成する再生情報生成部と、
前記複数のシーンを特定する情報に基づいて、前記複数のシーンから構成されるダイジェスト映像を生成するダイジェスト映像生成部と、
、
を備える映像処理装置。

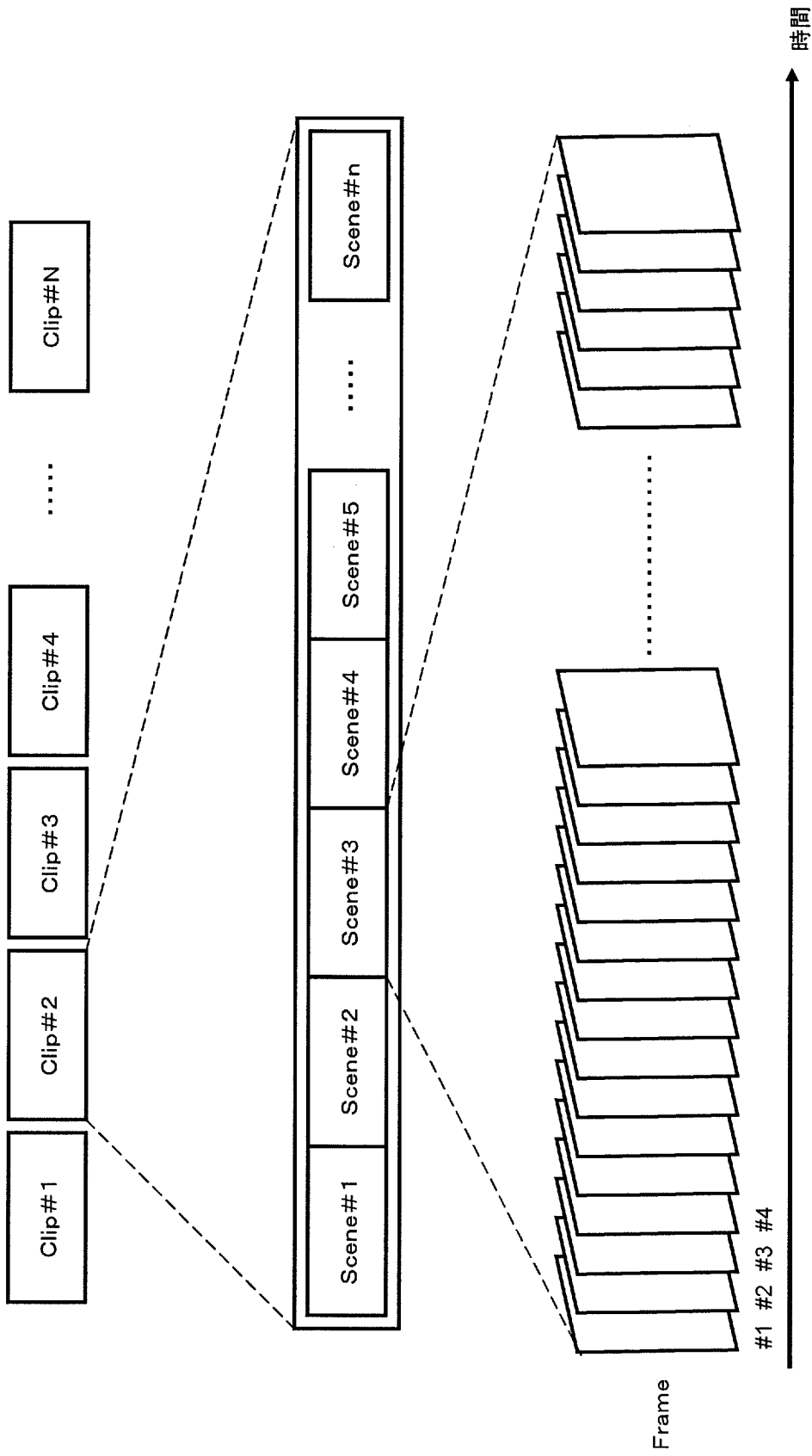
[図1]



[図2]



[図3]



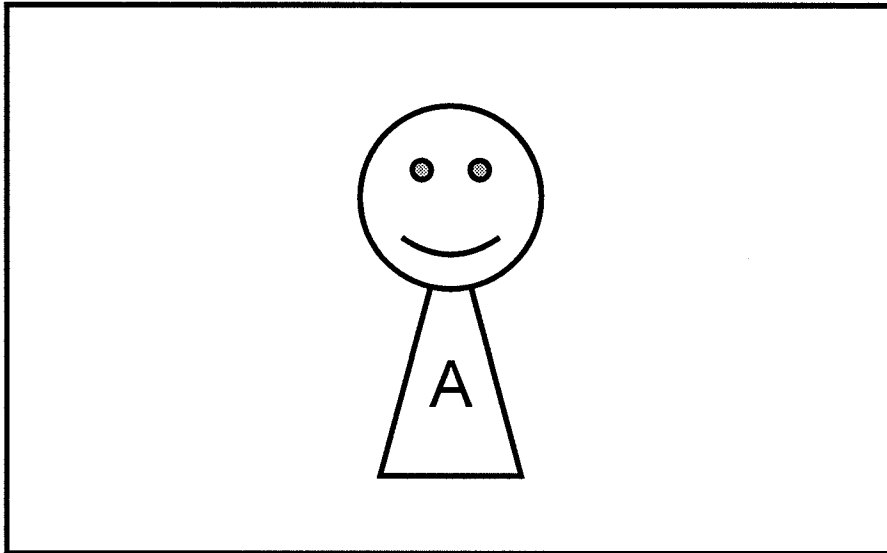
[図4]

シーン番号	開始時間	終了時間	代表時間
1	0:00:00	0:00:15	0:00:10
2	0:00:15	0:00:30	0:00:15
3	0:00:30	0:01:01	0:00:45
4	0:01:01	0:03:00	0:02:30
5	0:03:00	0:05:30	0:05:00
6	0:05:30	0:06:00	0:05:40
.			
.			
.			

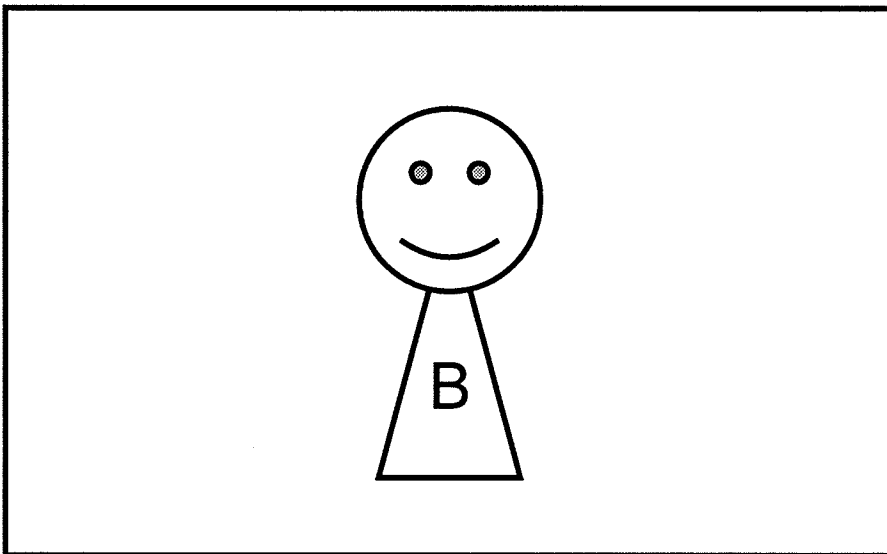
[図5]

識別	属性情報の内容	評価値
A	クリップイン	100
B	特定音声の抽出	50
C	(カメラワーク後に)静止して撮影	40
D	ズームアップ	30
E	パン、ティルト	25
F	クリップアウト	90
G	ズームダウン	30
H	特定色の抽出	50
I	像ぶれ	-20
J	地面(垂直下方向)撮影	-10
K	被写体距離-近	70
L	被写体距離-中	10
.	.	.
.	.	.
.	.	.
X	顔検出(人物A)	100
Y	顔検出(人物B)	80
Z	顔検出	50

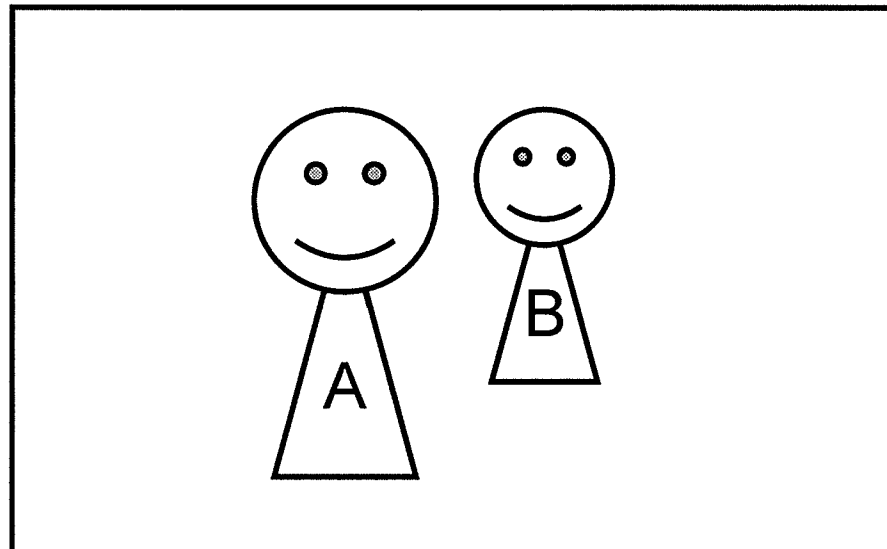
[図7A]



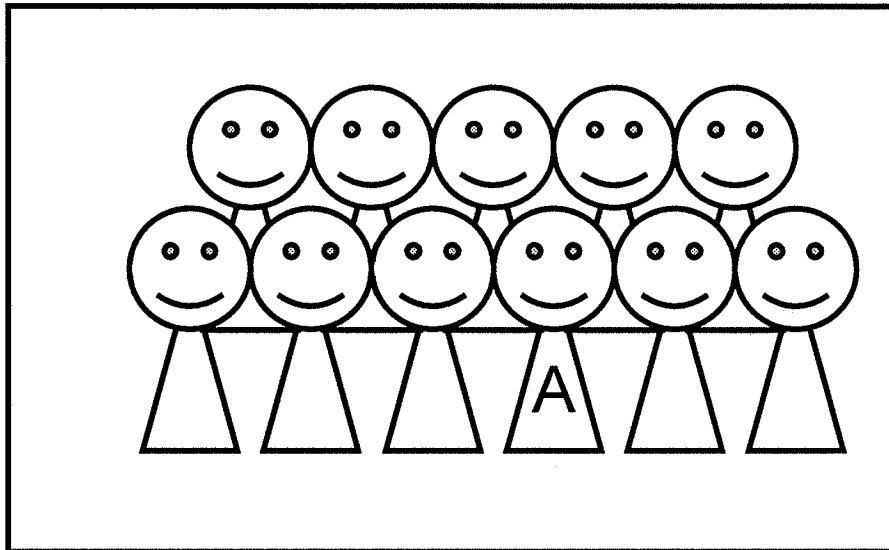
[図7B]



[図7C]



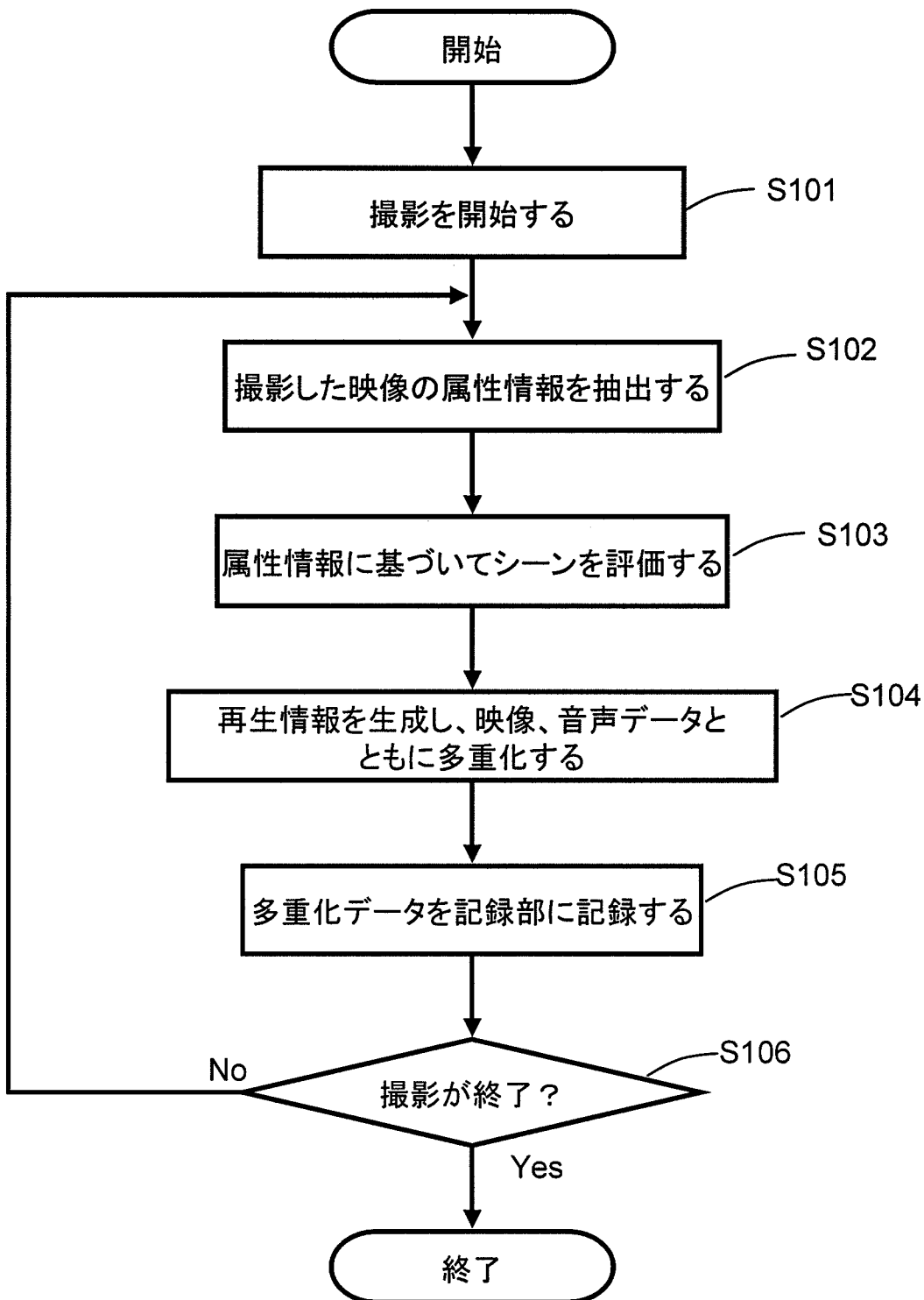
[図7D]



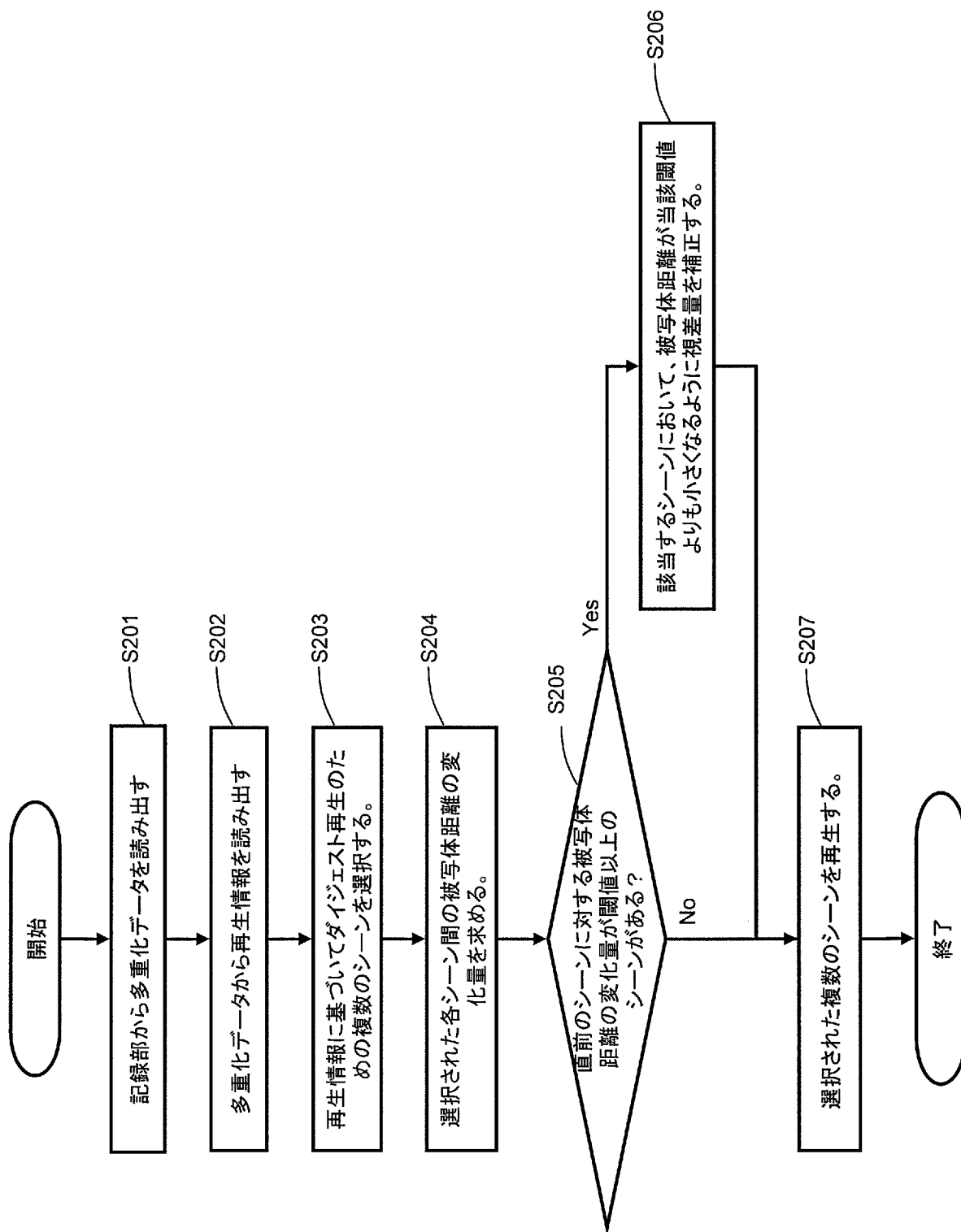
[図8]

シーン番号	開始時間	終了時間	代表時間
1	0:00:00	0:01:00	0:00:30
2	0:02:00	0:03:00	0:02:30
5	0:20:00	0:21:00	0:20:40

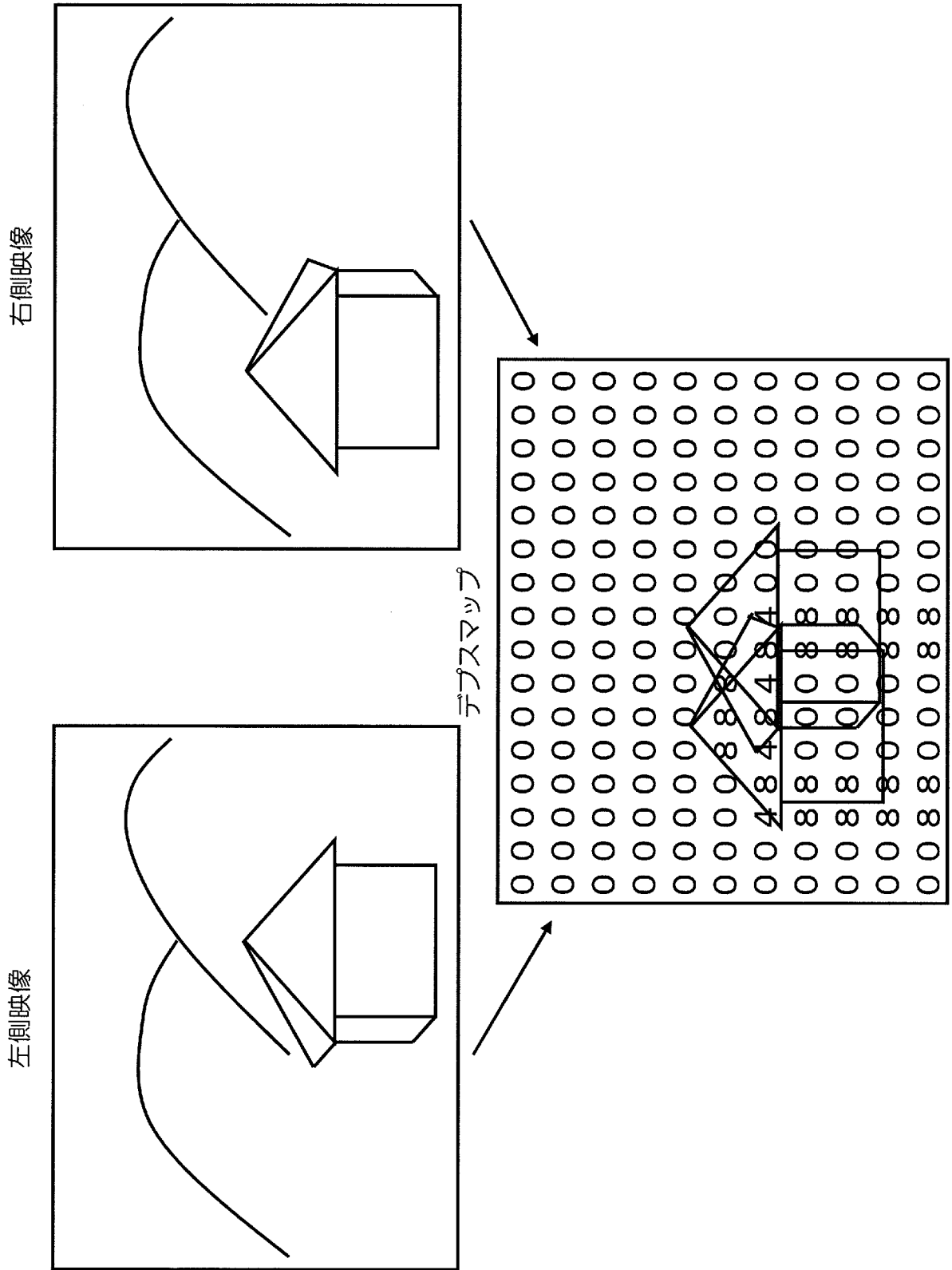
[図9]



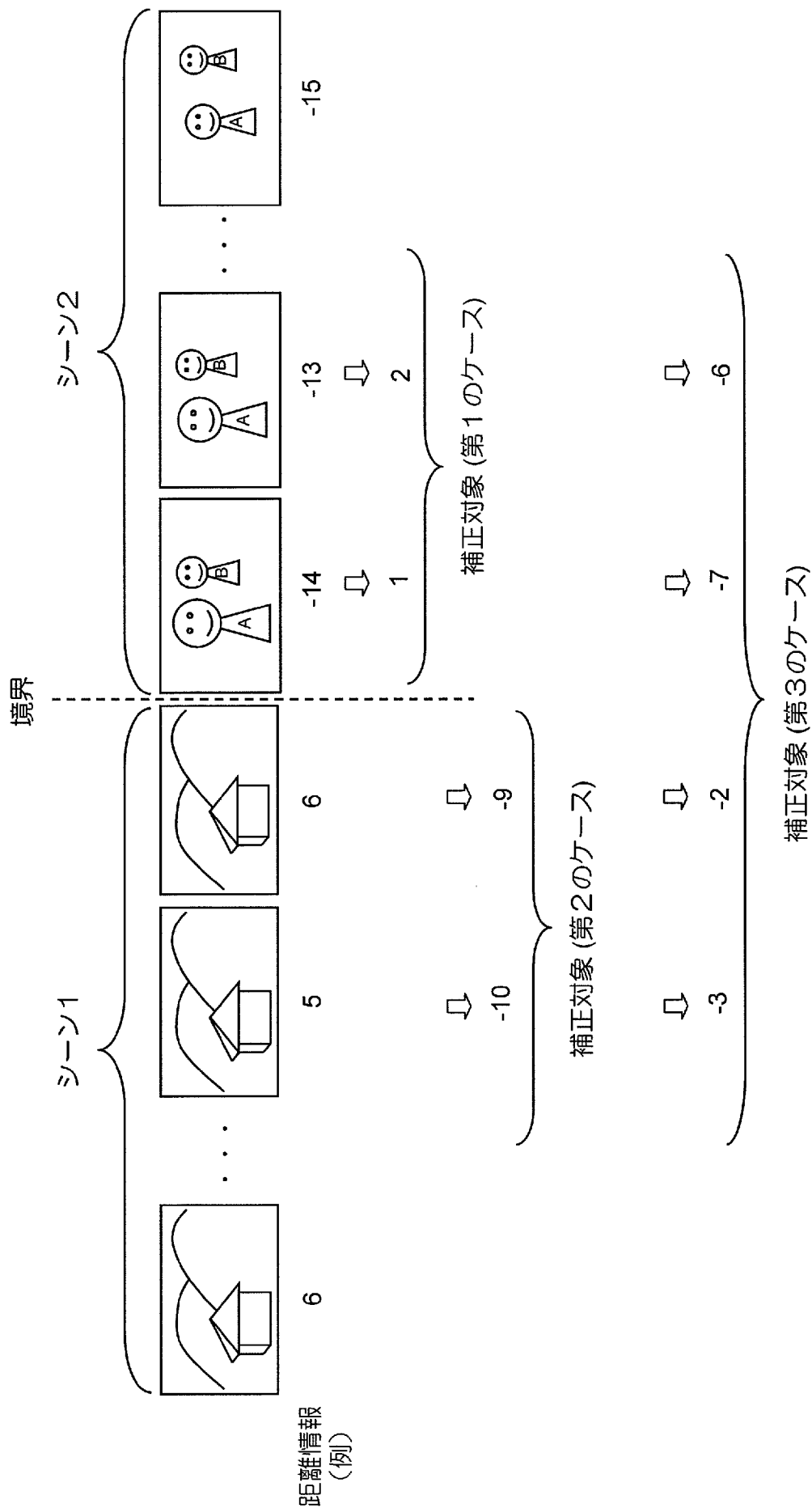
[図10A]



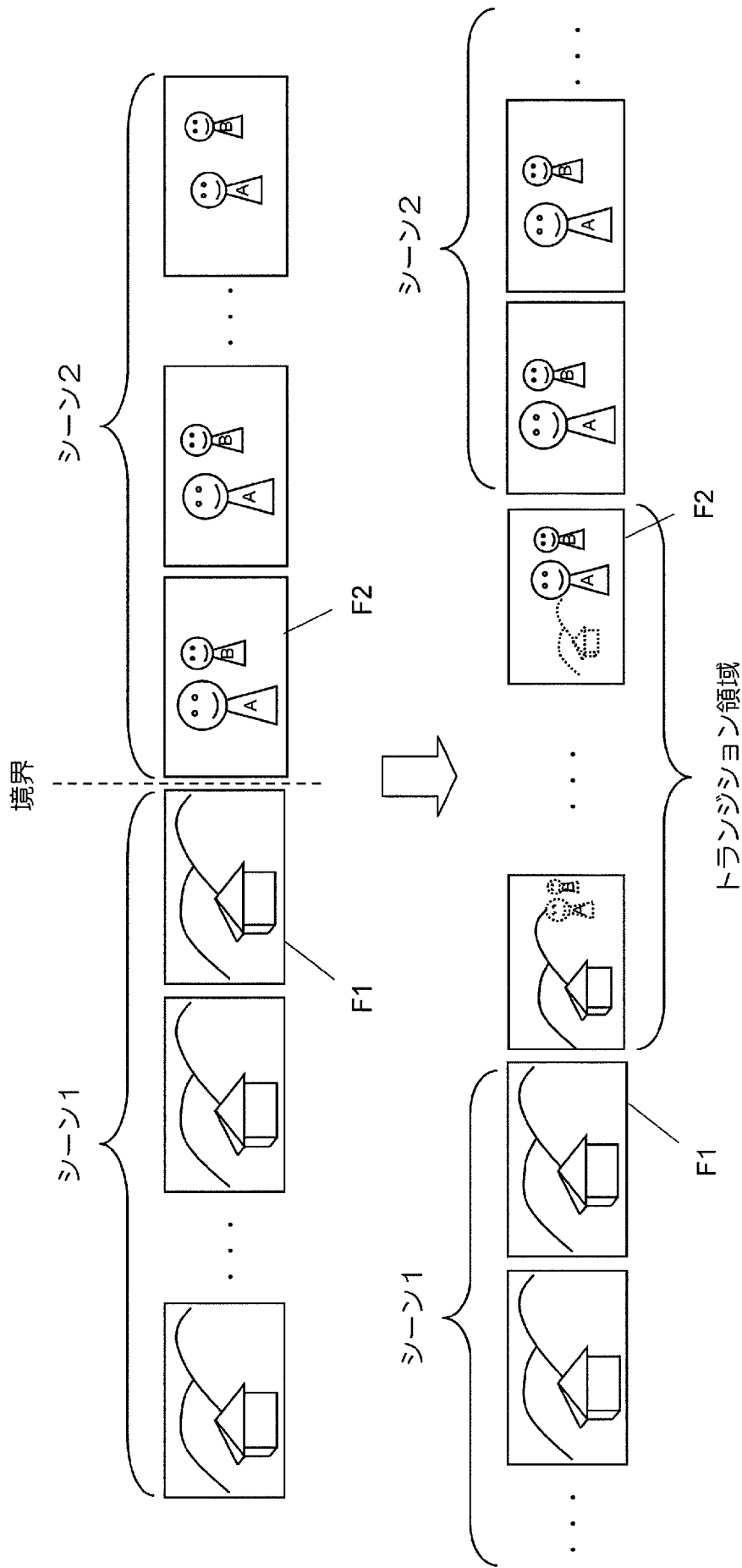
[図10B]



[図10D]



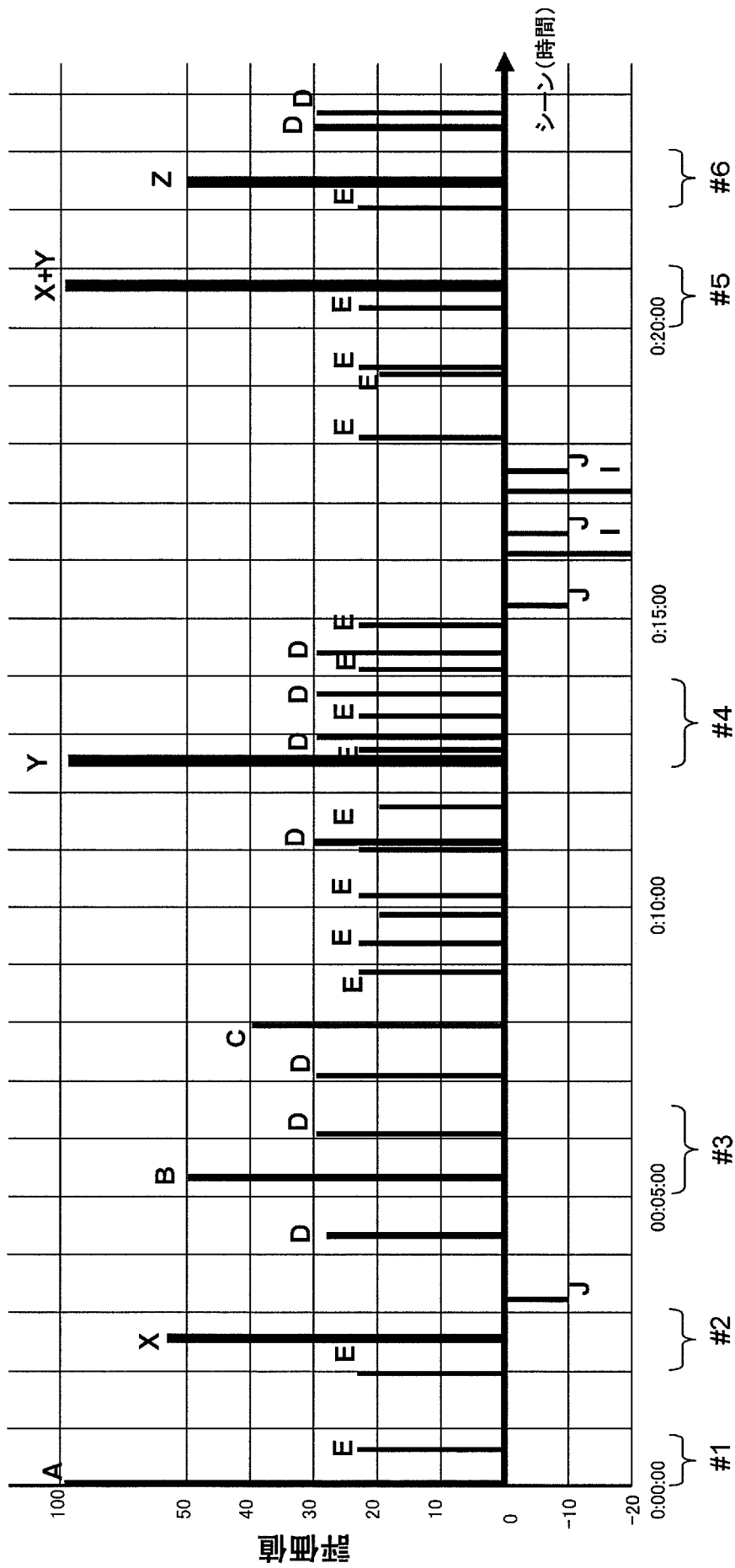
[図10E]



[図11]

識別	属性情報の内容	評価値
A	クリップイン	100
B	特定音声の抽出	50
C	(カメラワーク後に)静止して撮影	40
D	ズームアップ	30
E	パン、ティルト	25
F	クリップアウト	90
G	ズームダウン	30
H	特定色の抽出	50
I	像ぶれ	-20
J	地面(垂直下方向)撮影	-10
K	被写体距離(近)	70
L	被写体距離(中)	30
.	.	.
.	.	.
.	.	.
X	顔検出(人物A)	60
Y	顔検出(人物B)	90
Z	顔検出	50

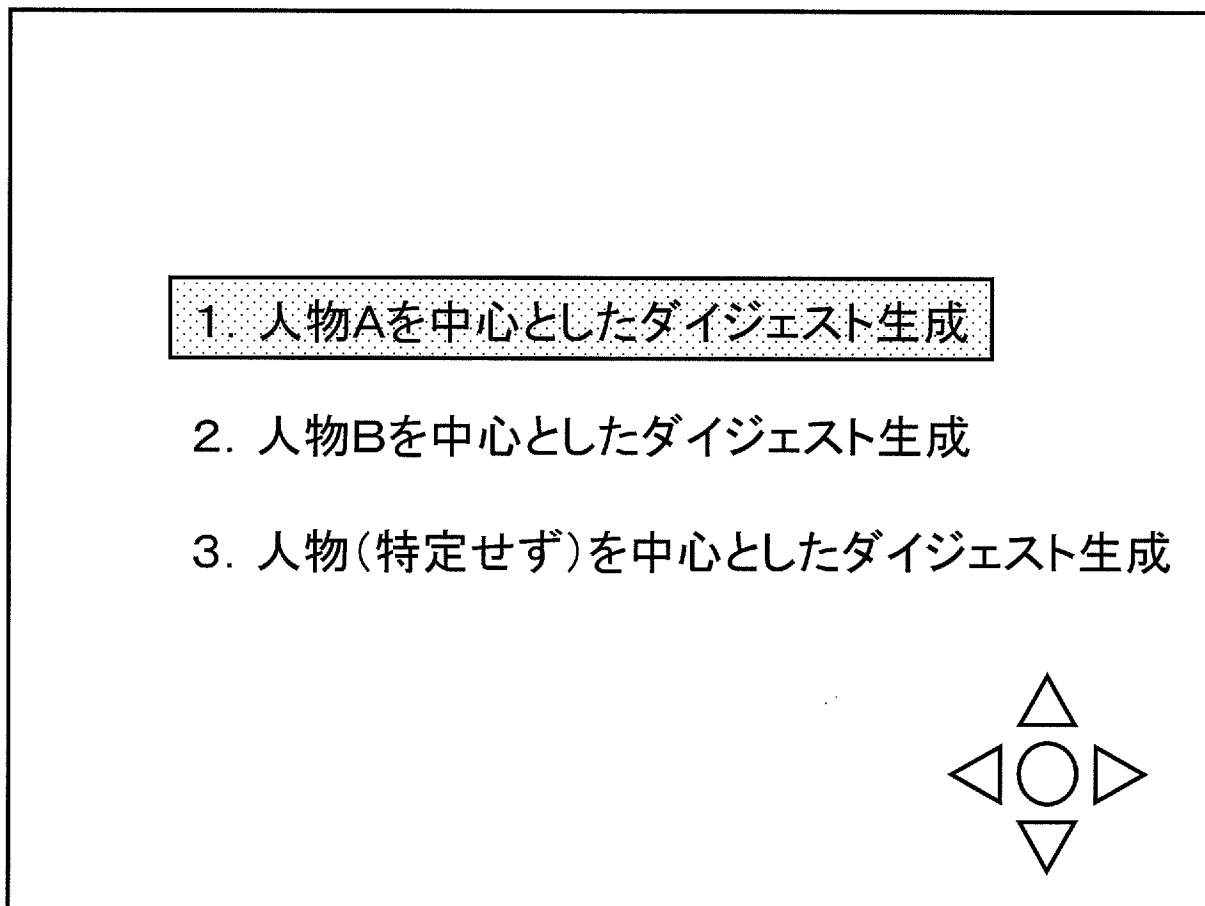
[図12]



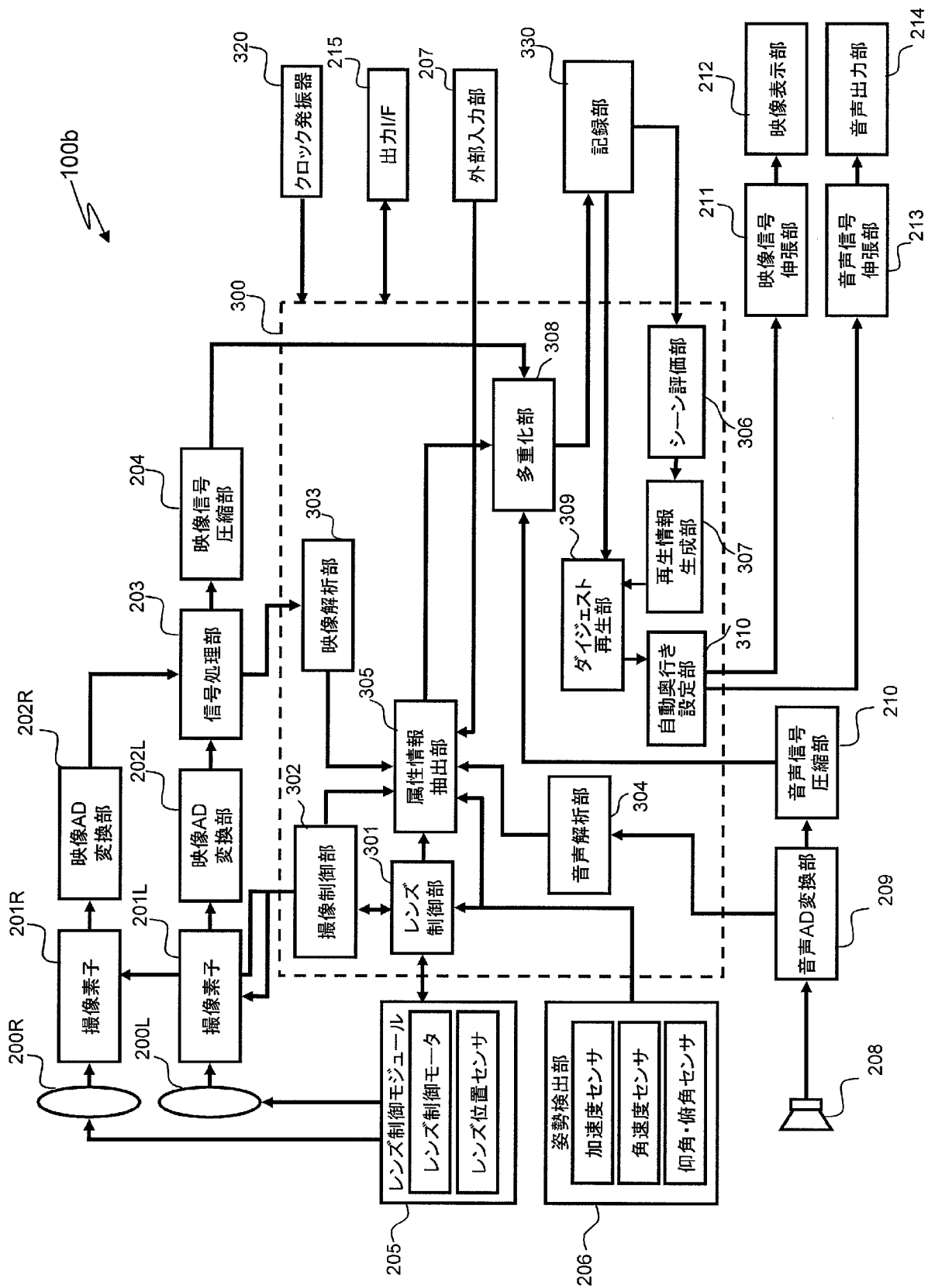
[図13]

シーン番号	開始時間	終了時間	代表時間
1	0:00:00	0:01:00	0:00:30
4	0:12:30	0:14:00	0:12:35
5	0:20:00	0:21:00	0:20:40

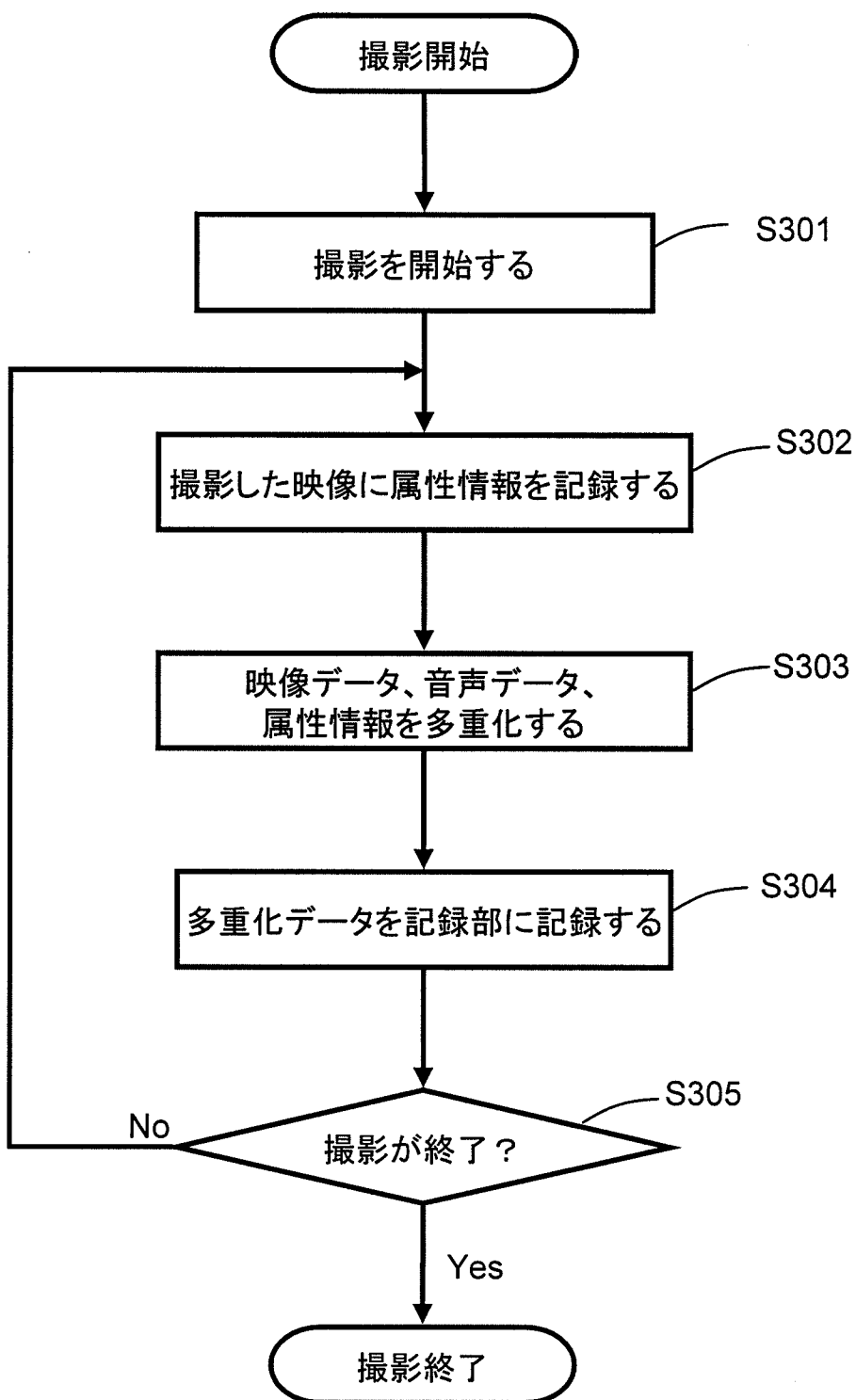
[図14]



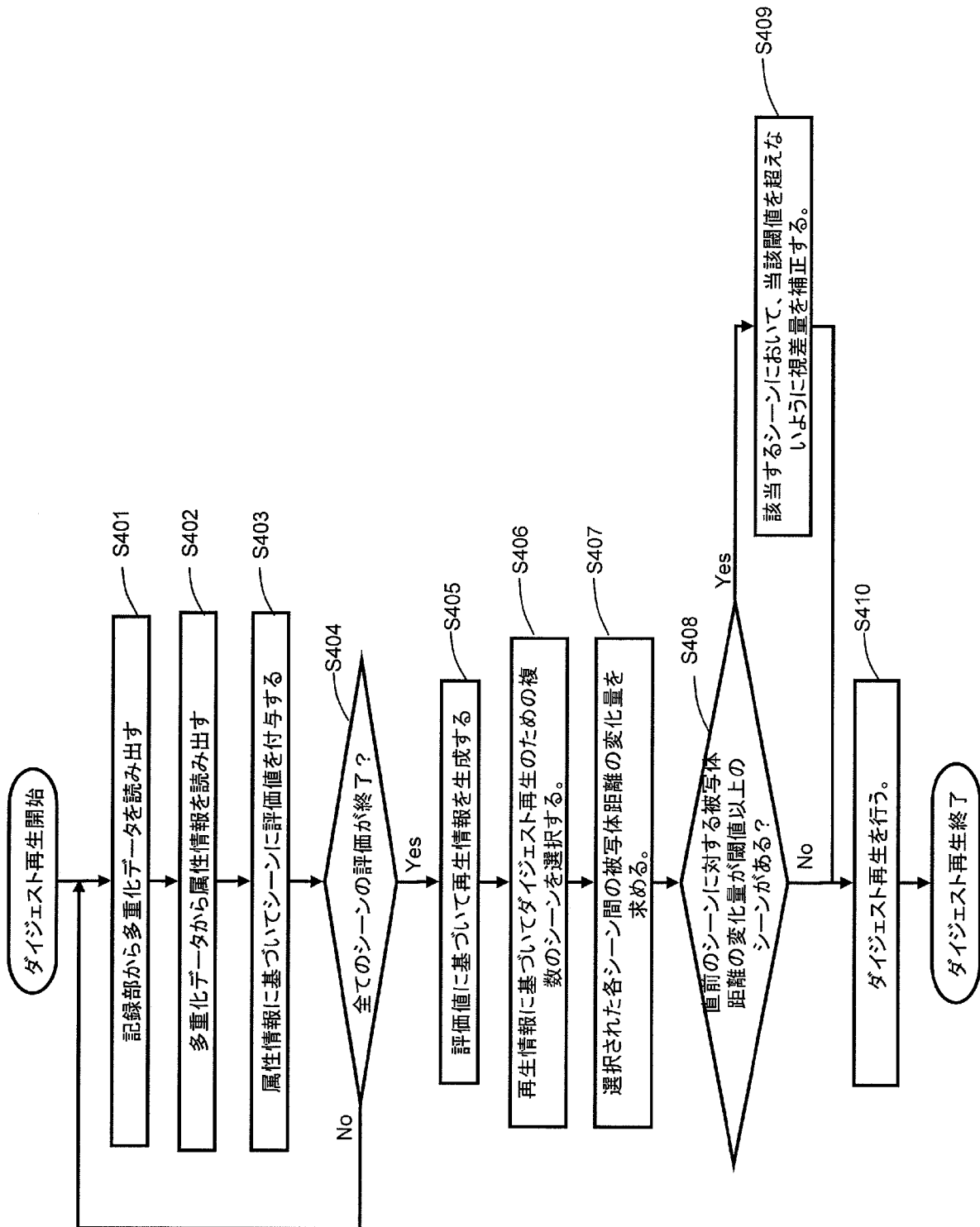
[図15]



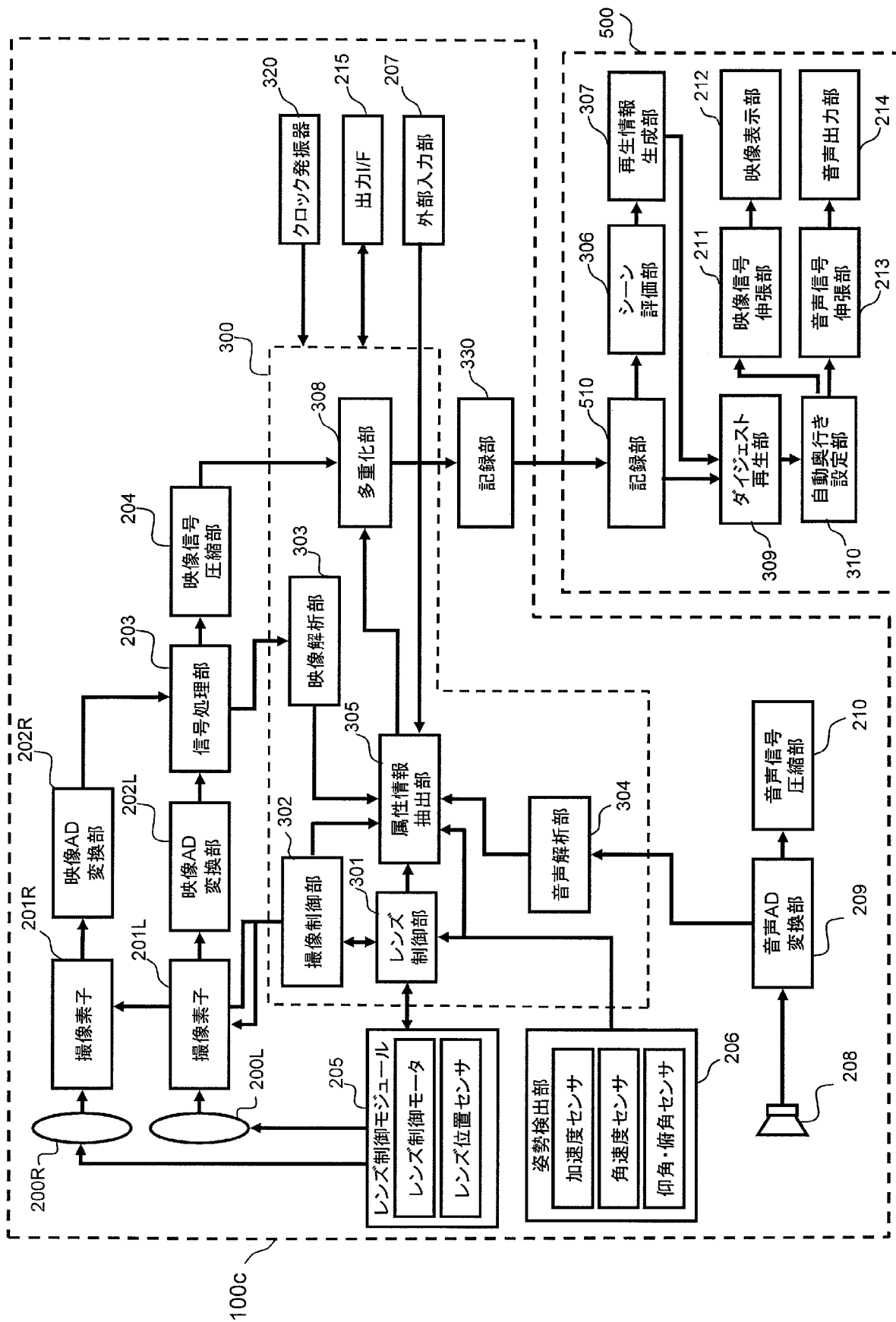
[図16]



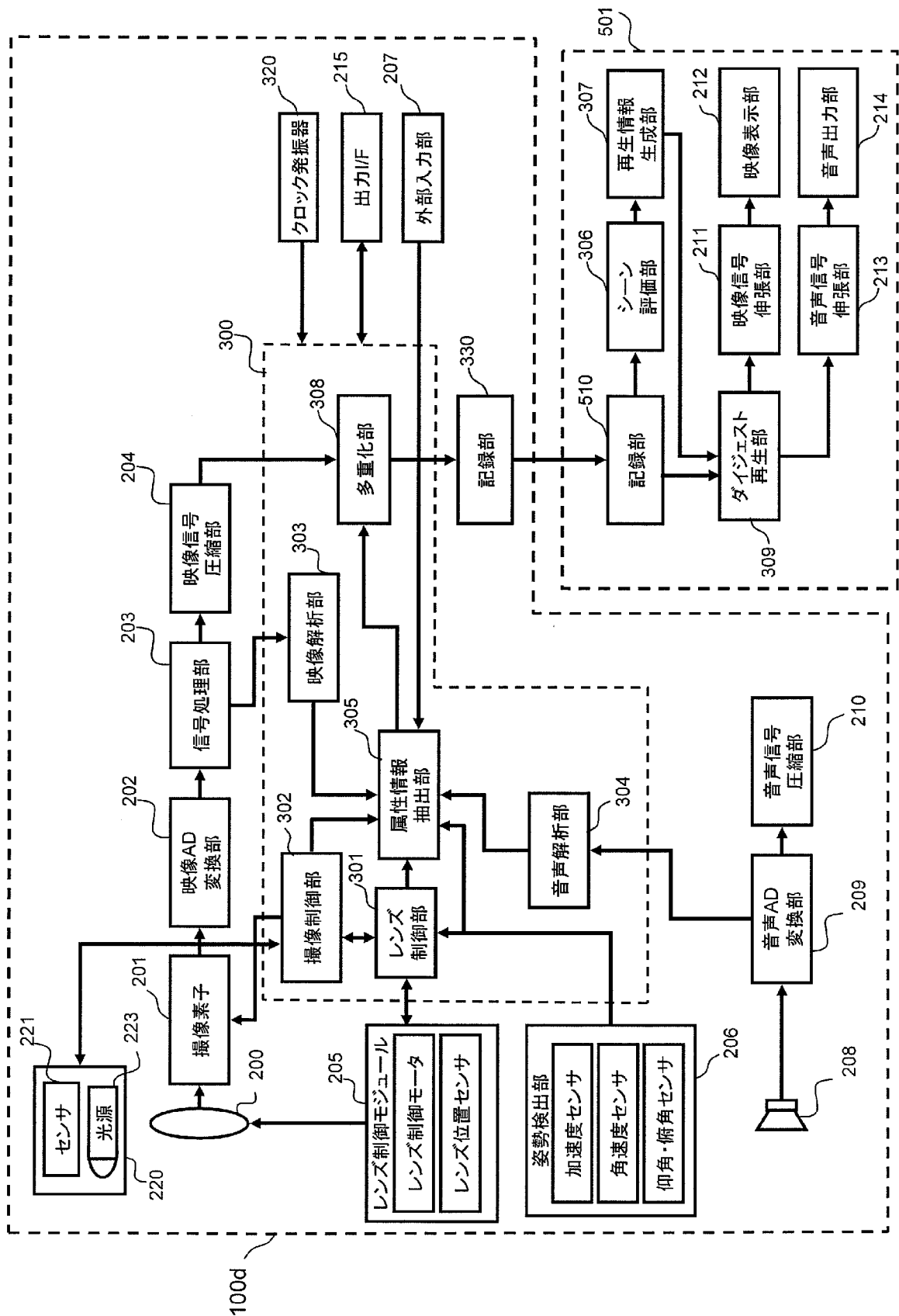
[図17]



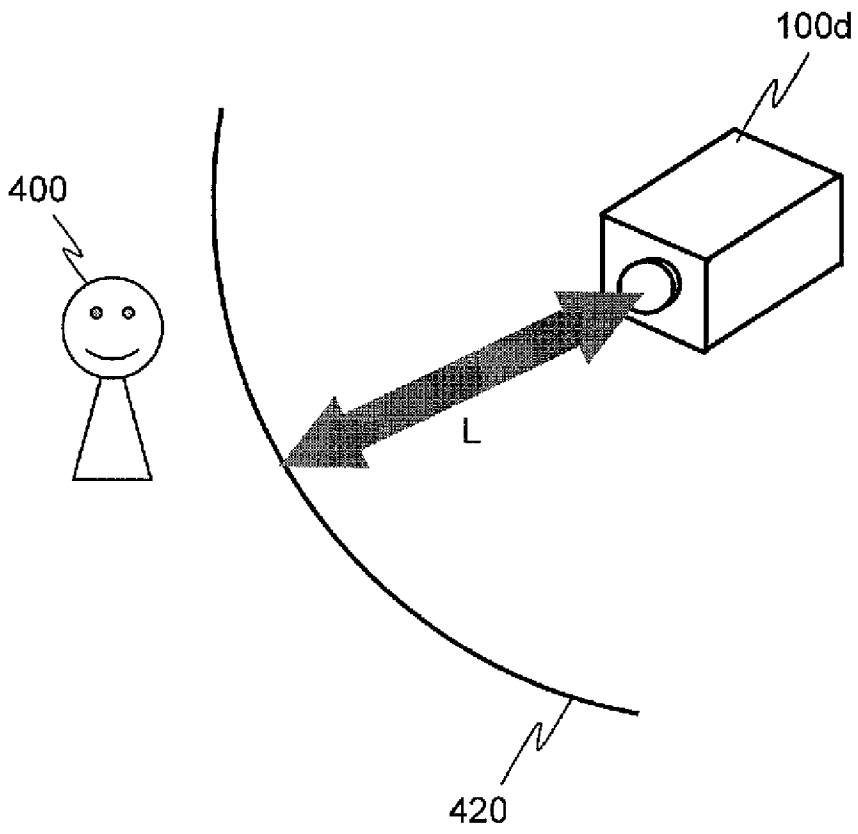
[図18]



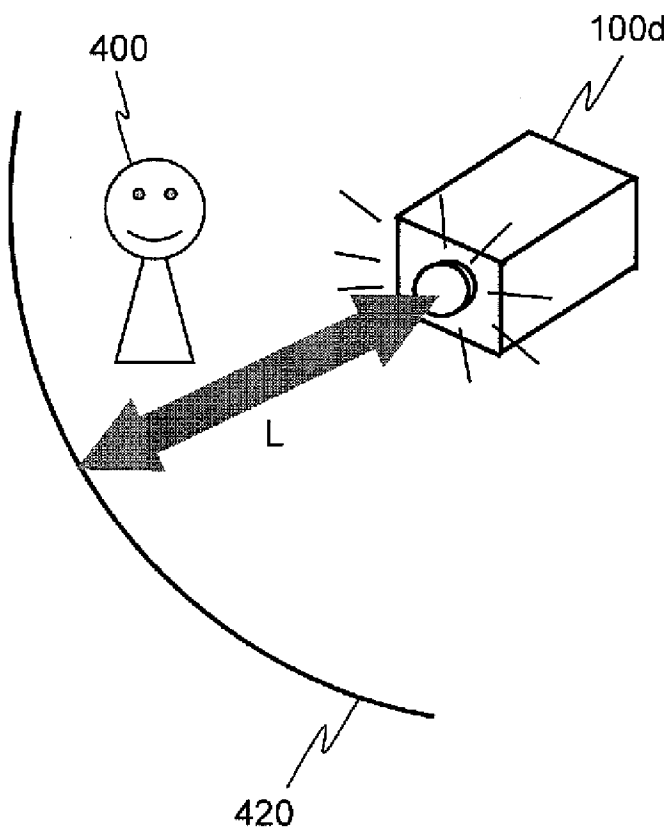
[図19]



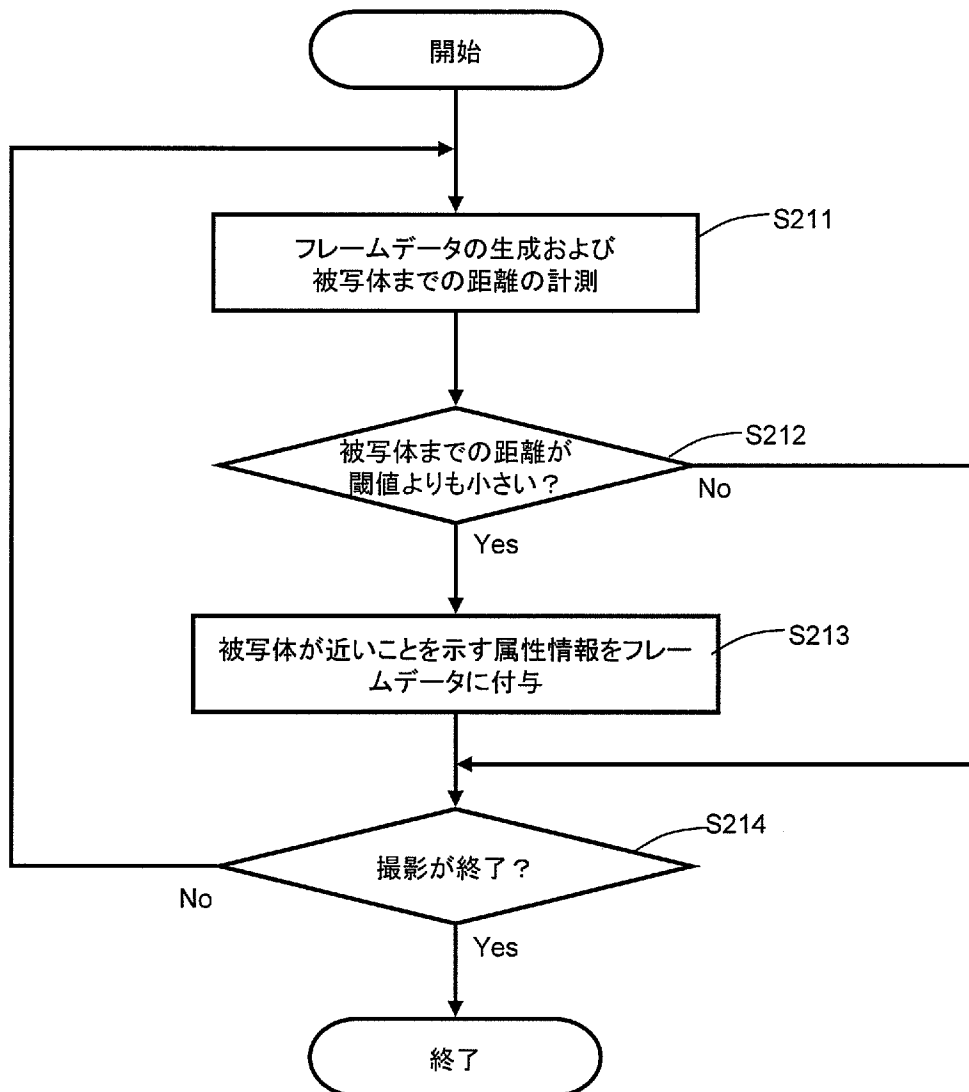
[図20A]



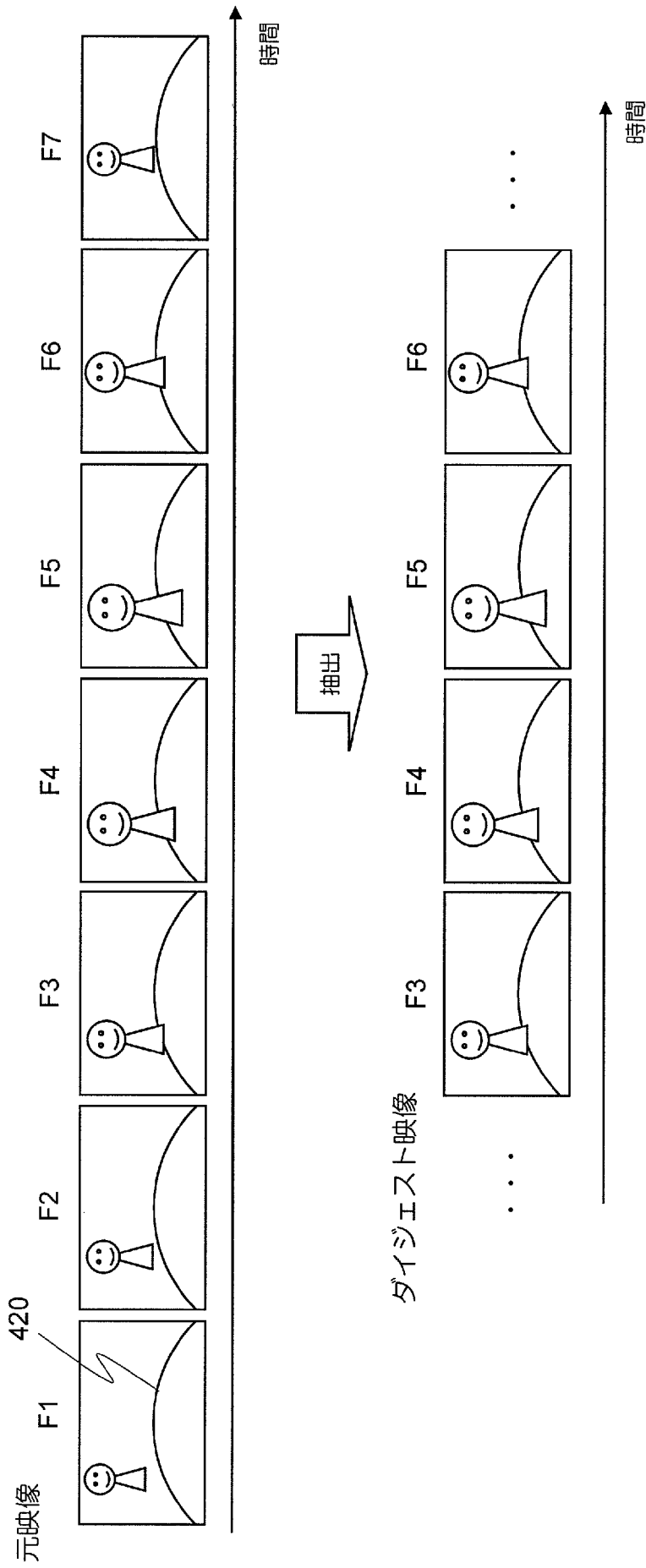
[図20B]



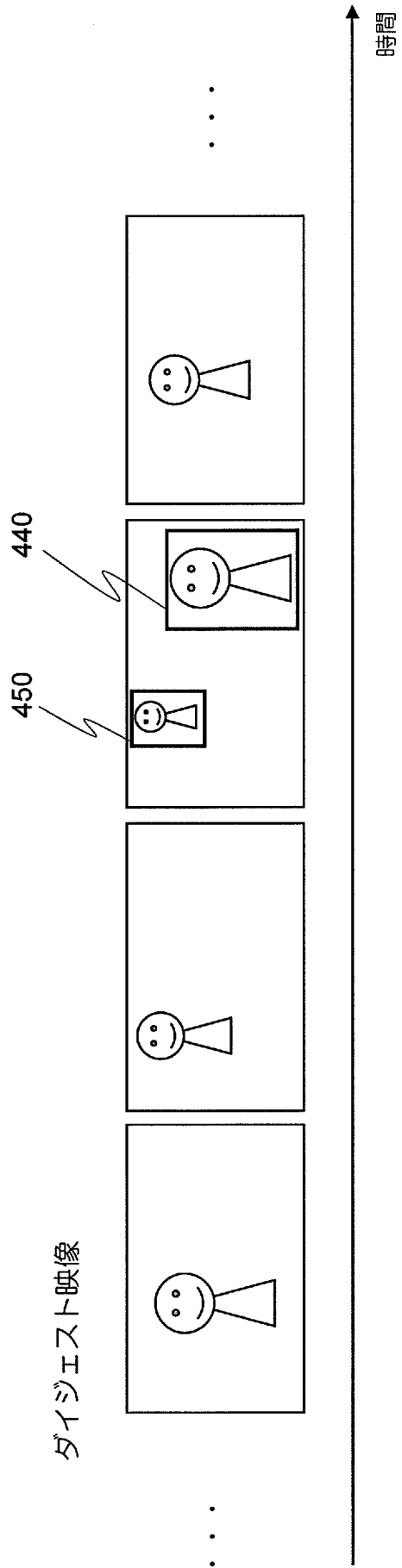
[図21]



[図22]



[図23]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/001069

A. CLASSIFICATION OF SUBJECT MATTER

H04N13/04(2006.01)i, H04N5/225(2006.01)i, H04N5/232(2006.01)i, H04N5/91(2006.01)i, H04N5/92(2006.01)i, H04N5/93(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)

H04N13/04, H04N5/225, H04N5/232, H04N5/91, H04N5/92, H04N5/93

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 2012-54862 A (Sony Corp.), 15 March 2012 (15.03.2012), paragraphs [0011], [0129], [0096], [0097]; fig. 13, 16 & WO 2012/029615 A1	1-15
Y	JP 2009-239388 A (Fujifilm Corp.), 15 October 2009 (15.10.2009), paragraphs [0007], [0008], [0042], [0070], [0092]; fig. 7 to 9, 17 to 19 & US 2009/0244268 A1 & EP 2106150 A2 & KR 10-2009-0102705 A	1-15

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
02 April, 2013 (02.04.13)

Date of mailing of the international search report
16 April, 2013 (16.04.13)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/001069

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2011-15256 A (Sony Corp.), 20 January 2011 (20.01.2011), paragraphs [0030], [0035], [0036], [0059], [0061], [0140]; fig. 2, 7, 16 to 18 & US 2011/0001800 A1 & CN 101945212 A	1-15

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. H04N13/04(2006.01)i, H04N5/225(2006.01)i, H04N5/232(2006.01)i, H04N5/91(2006.01)i, H04N5/92(2006.01)i, H04N5/93(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 (IPC))

Int.Cl. H04N13/04, H04N5/225, H04N5/232, H04N5/91, H04N5/92, H04N5/93

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	1922-1996年
日本国公開実用新案公報	1971-2013年
日本国実用新案登録公報	1996-2013年
日本国登録実用新案公報	1994-2013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 2012-54862 A (ソニー株式会社) 2012.03.15, 【0011】、【0129】、【0096】、【0097】、図13、 図16 & WO 2012/029615 A1	1-15
Y	JP 2009-239388 A (富士フイルム株式会社) 2009.10.15, 【0007】、【0008】、【0042】、【0070】、 【0092】、図7-図9、図17-図19 & US 2009/0244268 A1 & EP 2106150 A2 & KR 10-2009-0102705 A	1-15

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

02.04.2013

国際調査報告の発送日

16.04.2013

国際調査機関の名称及びあて先

日本国特許庁 (ISA/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)

益戸 宏

電話番号 03-3581-1101 内線 3581

5 P

9380

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	JP 2011-15256 A (ソニー株式会社) 2011.01.20, 【0030】 , 【0035】 , 【0036】 , 【0059】 , 【0061】 , 【0140】 , 図2, 図7, 図16 - 図18 & US 2011/0001800 A1 & CN 101945212 A	1 - 15