(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
2021年12月2日(02.12.2021)

WIPO | PCT

(10) 国際公開番号
WO 2021/240678 A1

(51) 国際特許分類:
H04N 5/91 (2006.01)

(21) 国際出願番号： PCT/JP2020/020868

(22) 国際出願日： 2020年5月27日(27.05.2020)

(25) 国際出願の言語： 日本語

(26) 国際公開の言語： 日本語

(71) 出願人: 日本電気株式会社 (**NEC CORPORATION**) [JP/JP]; 〒1088001 東京都港区芝五丁目７番１号 Tokyo (JP).

(72) 発明者: 白石 壮馬 (**SHIRAISHI, Soma**); 〒1088001 東京都港区芝五丁目７番１号 日本電気株式会社内 Tokyo (JP). 菊池 克 (**KIKUCHI, Katsumi**); 〒1088001 東京都港区芝五丁目７番１号 日本電気株式会社内 Tokyo (JP). 鍋藤 悠 (**NABETO, Yu**); 〒1088001 東京都港区芝五丁目７番１号 日本電気株式会社内 Tokyo (JP). 渡辺 はるな (**WATANABE, Haruna**); 〒1088001 東京都港区芝五丁目７番１号 日本電気株式会社内 Tokyo (JP).

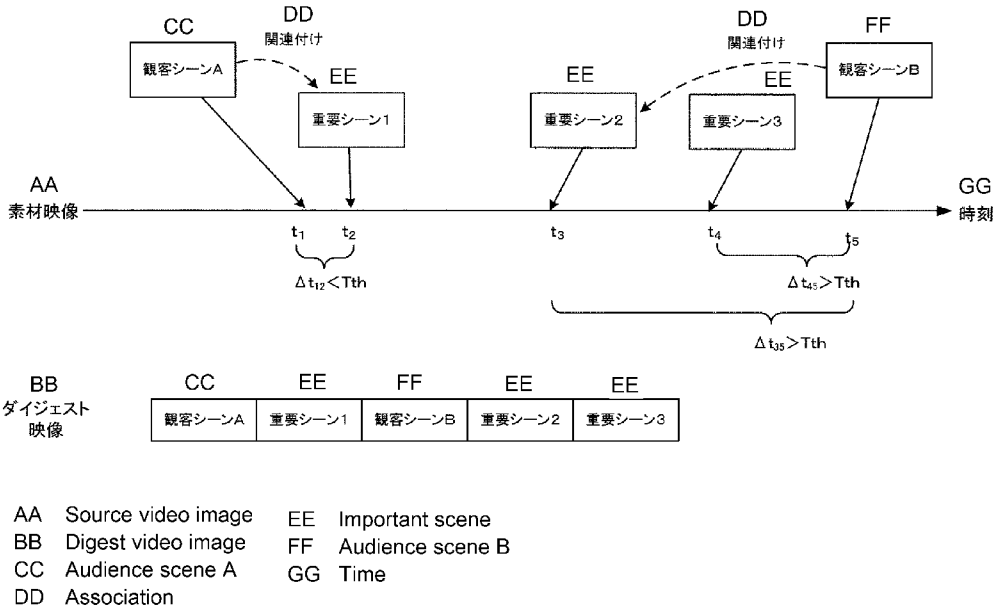
(74) 代理人: 中村 聡延, 外 (**NAKAMURA, Toshinobu et al.**); 〒1040031 東京都中央区京橋一丁目１６番１０号 オークビル京橋３階 東京セントラル特許事務所内 Tokyo (JP).

(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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH,

(54) **Title:** VIDEO IMAGE PROCESSING DEVICE, VIDEO IMAGE PROCESSING METHOD, AND RECORDING MEDIUM

(54) 発明の名称： 映像処理装置、映像処理方法、及び、記録媒体

[図6]



AA Source video image
BB Digest video image
CC Audience scene A
DD Association
EE Important scene
FF Audience scene B
GG Time

(57) **Abstract:** In this video image processing device, a video image acquisition means acquires a source video image. An audience scene extraction means extracts, from the source video image, an audience scene in which an audience appears. An important scene extraction means extracts an important scene from the source video image. An association means associates the audience scene with the important scene. A generation means generates a digest video image including the important scene and the audience scene which is associated with the important scene.

WO 2021/240678 A1

[続葉有]

KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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, ST, 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, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告(条約第21条(3))

(57) 要約:映像処理装置において、映像取得手段は、素材映像を取得する。観客シーン抽出手段は、素材映像から、観客を映した観客シーンを抽出する。重要シーン抽出手段は、素材映像から重要シーンを抽出する。関連付け手段は、観客シーンを重要シーンに関連付ける。生成手段は、重要シーンと、重要シーンに関連付けされた観客シーンとを含むダイジェスト映像を生成する。

明　細　書

発明の名称 ： 映像処理装置、映像処理方法、及び、記録媒体

技術分野

[0001] 本発明は、映像データの処理に関する。

背景技術

[0002] 動画像から映像ダイジェストを生成する技術が提案されている。特許文献１には、予め準備されたトレーニング動画像及びユーザが指定した重要シーン動画像から学習データファイルを作成し、当該学習データファイルに基づき、対象の動画像から重要シーンの検出を行うハイライト抽出装置が開示されている。

先行技術文献

特許文献

[0003] 特許文献1：特開２００８－０２２１０３号公報

発明の概要

発明が解決しようとする課題

[0004] スポーツの試合映像からダイジェスト映像を作成する場合、人間が編集したダイジェスト映像においては、選手がプレイする映像だけでなく、観客席にいる観客や観客が持つメッセージボードなどの映像が含められることが多い。しかし、このような観客のシーンは選手がプレイするシーンと比べて数が少ないため、機械学習により重要シーンとして学習することが困難であり、ダイジェスト映像に含めることが難しい。

[0005] 本発明の１つの目的は、スポーツ映像などにおける観客シーンをダイジェスト映像に含むダイジェスト映像を生成することが可能な映像処理装置を提供することにある。

課題を解決するための手段

[0006] 本発明の一つの観点では、映像処理装置は、

素材映像を取得する映像取得手段と、

前記素材映像から、観客を映した観客シーンを抽出する観客シーン抽出手段と、

前記素材映像から重要シーンを抽出する重要シーン抽出手段と、

前記観客シーンを前記重要シーンに関連付ける関連付け手段と、

前記重要シーンと、前記重要シーンに関連付けされた観客シーンとを含むダイジェスト映像を生成する生成手段と、を備える。

[0007] 本発明の他の観点では、映像処理方法は、

素材映像を取得し、

前記素材映像から、観客を映した観客シーンを抽出し、

前記素材映像から重要シーンを抽出し、

前記観客シーンを前記重要シーンに関連付け、

前記重要シーンと、前記重要シーンに関連付けされた観客シーンとを含むダイジェスト映像を生成する。

[0008] 本発明のさらに他の観点では、記録媒体は、

素材映像を取得し、

前記素材映像から、観客を映した観客シーンを抽出し、

前記素材映像から重要シーンを抽出し、

前記観客シーンを前記重要シーンに関連付け、

前記重要シーンと、前記重要シーンに関連付けされた観客シーンとを含むダイジェスト映像を生成する処理をコンピュータに実行させるプログラムを記録する。

発明の効果

[0009] 本発明によれば、スポーツ映像などにおける観客シーンをダイジェスト映像に含むダイジェスト映像を生成することが可能となる。

図面の簡単な説明

[0010] [図1]実施形態に係るダイジェスト生成装置の全体構成を示す。

[図2]ダイジェスト映像の例を示す。

[図3]ダイジェスト生成装置の訓練時及び推論時の構成を示す。

[図4]ダイジェスト生成装置のハードウェア構成を示すブロック図である。

[図5]観客席の映像の例を示す。

[図6]観客シーンをダイジェスト映像に含める方法を模式的に示す。

[図7]第１実施形態に係るダイジェスト生成装置の機能構成を示す。

[図8]ダイジェスト生成処理のフローチャートである。

[図9]観客シーン抽出処理のフローチャートである。

[図10]観客シーン抽出モデルの訓練装置の機能構成を示す。

[図11]訓練処理のフローチャートである。

[図12]第２実施形態に係る映像処理装置の機能構成を示すブロック図である。

発明を実施するための形態

[0011] 以下、図面を参照して、本発明の好適な実施形態について説明する。

＜基本構成＞

まず、実施形態に係るダイジェスト生成装置の基本構成について説明する。

［全体構成］

図１は、実施形態に係るダイジェスト生成装置１００の全体構成を示す。ダイジェスト生成装置１００は、素材映像データベース（以下、「データベース」を「ＤＢ」とも記す。）２に接続されている。素材映像ＤＢ２は、各種の素材映像、即ち、動画像を記憶している。素材映像は、例えば放送局から放送されるテレビ番組などの映像でもよく、インターネットなどで配信されている映像でもよい。なお、素材映像は、音声を含んでいてもよく、含んでいなくてもよい。

[0012] ダイジェスト生成装置１００は、素材映像ＤＢ２に保存されている素材映像の一部を用いるダイジェスト映像を生成し、出力する。ダイジェスト映像は、素材映像における重要シーンを時系列につなげた映像である。ダイジェスト生成装置１００は、機械学習により訓練済みのダイジェスト生成モデル（以下、単に「生成モデル」とも呼ぶ。）を用いてダイジェスト映像を生成

する。生成モデルとしては、例えば、ニューラルネットワークを用いたモデルを用いることができる。

[0013] 図２は、ダイジェスト映像の例を示す。図２の例では、ダイジェスト生成装置１００は、素材映像に含まれるシーンＡ～Ｄを重要シーンとして抽出し、これらを時系列につなげたダイジェスト映像を生成する。なお、素材映像から抽出された重要シーンは、その内容次第で、ダイジェスト映像中で繰り返し使用されてもよい。

[0014] ［機能構成］

図３（Ａ）は、ダイジェスト生成装置１００が使用する生成モデルを訓練するための構成を示すブロック図である。生成モデルの訓練には、予め用意された訓練データセットが使用される。訓練データセットは、訓練用素材映像と、訓練用素材映像に対する正解を示す正解データのペアである。正解データは、訓練用素材映像における重要シーンの位置に正解を示すタグ（以下、「正解タグ」と呼ぶ。）を付与したデータである。典型的には、正解データにおける正解タグの付与は、経験のある編集者などにより行われる。例えば野球中継の素材映像については、野球解説者などが試合中のハイライトシーンなどを選択し、正解タグを付与する。なお、編集者による正解タグの付与方法を機械学習などで学習し、正解タグを自動的に付与してもよい。

[0015] 訓練時には、訓練用素材映像が生成モデルＭに入力される。生成モデルＭは、素材映像から重要シーンを抽出する。具体的には、生成モデルＭは、素材映像を構成する１枚又は複数のフレームの集合から特徴量を抽出し、抽出した特徴量に基づいて素材映像に対する重要度（重要度スコア）を算出する。そして、生成モデルＭは、重要度が予め決められた閾値以上である部分を重要シーンとして出力する。訓練部４は、生成モデルＭの出力と、正解データとを用いて、生成モデルＭを最適化する。具体的には、訓練部４は、生成モデルＭが出力した重要シーンと、正解データに含まれる正解タグが示すシーンとを比較し、その誤差（損失）を小さくするように生成モデルＭのパラメータを更新する。こうして得られた訓練済みの生成モデルＭは、編集者が

正解タグを付与したシーンに近いシーンを素材映像から重要シーンとして抽出することができるようになる。

[0016] 図３（Ｂ）は、ダイジェスト生成装置１００による推論時の構成を示す。推論時には、ダイジェスト映像の生成の対象となる素材映像が訓練済みの生成モデルＭに入力される。生成モデルＭは、素材映像から重要度を算出し、重要度が予め決められた閾値以上である部分を重要シーンとして抽出し、ダイジェスト生成部５に出力する。ダイジェスト生成部５は、生成モデルＭが抽出した重要シーンをつなげてダイジェスト映像を生成し、出力する。こうして、ダイジェスト生成装置１００は、訓練済みの生成モデルＭを用いて、素材映像からダイジェスト映像を生成する。

[0017] ［ハードウェア構成］

図４は、ダイジェスト生成装置１００のハードウェア構成を示すブロック図である。図示のように、ダイジェスト生成装置１００は、インタフェース（ＩＦ）１１と、プロセッサ１２と、メモリ１３と、記録媒体１４と、データベース（ＤＢ）１５とを備える。

[0018] ＩＦ１１は、外部装置との間でデータの入出力を行う。具体的に、素材映像ＤＢ２に保存されている素材映像はＩＦ１１を介してダイジェスト生成装置１００に入力される。また、ダイジェスト生成装置１００により生成されたダイジェスト映像は、ＩＦ１１を通じて外部装置へ出力される。

[0019] プロセッサ１２は、ＣＰＵ（Ｃｅｎｔｒａｌ Ｐｒｏｃｅｓｓｉｎｇ Ｕｎｉｔ）などのコンピュータであり、予め用意されたプログラムを実行することにより、ダイジェスト生成装置１００の全体を制御する。具体的に、プロセッサ１２は、後述する訓練処理、及び、ダイジェスト生成処理を実行する。

[0020] メモリ１３は、ＲＯＭ（Ｒｅａｄ Ｏｎｌｙ Ｍｅｍｏｒｙ）、ＲＡＭ（Ｒａｎｄｏｍ Ａｃｃｅｓｓ Ｍｅｍｏｒｙ）などにより構成される。メモリ１３は、プロセッサ１２による各種の処理の実行中に作業メモリとしても使用される。

[0021] 記録媒体１４は、ディスク状記録媒体、半導体メモリなどの不揮発性で非一時的な記録媒体であり、ダイジェスト生成装置１００に対して着脱可能に構成される。記録媒体１４は、プロセッサ１２が実行する各種のプログラムを記録している。ダイジェスト生成装置１００が各種の処理を実行する際には、記録媒体１４に記録されているプログラムがメモリ１３にロードされ、プロセッサ１２により実行される。

[0022] データベース１５は、ＩＦ１１を通じて入力された素材映像、ダイジェスト生成装置１００が生成したダイジェスト映像などを一時的に記憶する。また、データベース１５は、ダイジェスト生成装置１００が使用する訓練済みの生成モデルの情報、及び、生成モデルの訓練に用いられる訓練データセットなどを記憶する。なお、ダイジェスト生成装置１００は、作成者が指示や入力を行うためのキーボード、マウスなどの入力部、及び、液晶ディスプレイなどの表示部を備えていてもよい。

[0023] ＜第１実施形態＞

次に、本発明の第１実施形態について説明する。

［原理説明］

第１実施形態では、ダイジェスト生成装置１００は、スポーツの試合映像などの素材映像からダイジェスト映像を生成する際に、観客席を映したシーン（以下、「観客シーン」と呼ぶ。）を抽出し、ダイジェスト映像に含める。この際、ダイジェスト生成装置１００は、素材映像から抽出した観客シーンを、素材映像から抽出された重要シーンに関連付けてダイジェスト映像に含める点に特徴を有する。図５（Ａ）は、観客席の映像の一例を示す。この映像は、多数の観客を含む観客席の映像である。

[0024] 図６は、観客シーンをダイジェスト映像に含める方法を模式的に示す。図６では、素材映像における時刻を横軸に示す。ダイジェスト生成装置１００は、素材映像から前処理により観客シーンを抽出しておく。図６の例では、素材映像から観客シーンＡ、Ｂが抽出されたとする。また、ダイジェスト生成装置１００は、素材映像から前述の方法で重要シーンを抽出する。図６の

例では、素材映像から重要シーン１～３が抽出されたとする。この場合、ダイジェスト生成装置１００は、観客シーンＡ、Ｂをいずれかの重要シーンに関連付ける処理を行う。そして、関連付けができた場合、ダイジェスト生成装置１００は、観客シーンを、時間軸上で関連付けされた重要シーンの前又は後に配置してダイジェスト映像を生成する。

[0025] 観客シーンを重要シーンに関連付けする方法としては、下記の方法が挙げられる。

（１）第１の方法

第１の方法は、素材映像における時刻に基づいて、観客シーンを重要シーンに関連付ける。具体的には、第１の方法は、観客シーンを、素材映像における時刻が最も近い重要シーンに関連付ける。なお、観客シーンと重要シーンの時間間隔（時刻の差）が所定の閾値以下である場合に限り、観客シーンを重要シーンに関連付けることとしてもよい。この場合、観客シーンと、それに最も近い重要シーンとの時間間隔が閾値より大きければ、その観客シーンはその重要シーンに関連付けされない。

[0026] なお、第１の方法で関連付けを行う場合、観客シーンの重要シーンに対する位置関係は、素材映像における両者の位置関係に従うことが好ましい。図６の例では、素材映像において観客シーンＡは重要シーン１の前にあるので、ダイジェスト映像の例に示すように、観客シーンＡを重要シーン１の前に配置する。逆に、素材映像において観客シーンが、関連付けの対象となる重要シーンより後にある場合には、その観客シーンを対象となる重要シーンの後に配置する。

[0027] （２）第２の方法

第２の方法は、観客シーンから色に関する情報を抽出し、これを用いて観客シーンを重要シーンに関連付けする。具体的には、ダイジェスト生成装置１００は、素材映像から抽出した観客シーンに含まれる人々の着用している服や帽子などの色、又は、それらの人々が所持している物体（例えば、メガホン、応援旗など）などの色を認識し、観客席の大部分を占める色に関する

情報を抽出する。

[0028] 通常、スポーツのチームには特定のチームカラーがあり、選手はそのチームカラーのユニフォームを着用している。また、そのチームのファンは、そのチームのユニフォームと同一又は類似したデザインのシャツ、帽子などを着用して試合を観戦することが多い。また、ファンは、チームカラーを採用したメガホン、応援旗などの応援グッズを持って応援することが多い。そこで、ダイジェスト生成装置１００は、観客シーンから色に関する情報を取得し、その色と同一又は類似するチームカラーを有するチームの重要シーンに関連付ける。例えば、素材映像がチームＡとチームＢの試合であり、チームＡのチームカラーが赤であり、チームＢのチームカラーが青であるとする。この場合、ダイジェスト生成装置１００は、観客席の大部分が赤で占められる観客シーンをチームＡに関する重要シーン（例えば、チームＡの得点シーンなど）に関連付け、観客席の大部分が青で占められる観客シーンをチームＢに関する重要シーンに関連付ける。

[0029] なお、あるチームについて、複数の観客シーンと複数の重要シーンが抽出された場合、各観客シーンを関連付ける重要シーンを選択する方法はいくつか考えらえる。例えば、各観客シーンを、そのチームの重要シーンのうち、時間的に最も近い重要シーンに関連付けてもよい。また、各観客シーンを、そのチームの複数の重要シーンのうち、ランダムに選択した重要シーンに関連付けてもよい。

[0030] （３）第３の方法

第３の方法は、観客シーンから文字列に関する情報を抽出し、これを用いて観客シーンを重要シーンに関連付けする。具体的には、ダイジェスト生成装置１００は、素材映像から抽出した観客シーンに含まれるメッセージボード、プラカード、応援旗などに書かれている応援メッセージなどの文字列を認識し、その観客シーンを、その文字列と関連性のある重要シーンに関連付けする。

[0031] 具体的に、観客シーンに映っているメッセージボードにチーム名、選手名

、選手の背番号などが書かれている場合、ダイジェスト生成装置１００は、その文字列が示すチームや、その文字列が示す選手が属するチームの重要シーンにその観客シーンを関連付けする。例えば、図５（Ｂ）に示すように、観客シーンに映っているメッセージボードに「Ｇｏ！ＧＩＡＮＴＳ！」のメッセージが書かれている場合、ダイジェスト生成装置１００は、この観客シーンをチーム「ＧＩＡＮＴＳ」の重要シーンに関連付けする。

[0032] 第３の方法でも、複数の観客シーンと複数の重要シーンが抽出された場合、ダイジェスト生成装置１００は、各観客シーンを、そのチームの重要シーンのうち、時間的に最も近い重要シーンに関連付けてもよいし、そのチームの複数の重要シーンのうち、ランダムに選択した重要シーンに関連付けてもよい。

[0033] 図６の例では、ダイジェスト生成装置１００は、第１の方法により観客シーンＡを重要シーン１に関連付け、重要シーン１の前に配置している。観客シーンＡについては、素材映像における観客シーンＡの時刻t_1と重要シーンの時刻t_2の時間間隔Δt_{12}が所定の閾値Ｔｔｈより小さいので、観客シーンＡが重要シーン１に関連付けされている。一方、観客シーンＢについては、重要シーン２の時間間隔Δt_{35}及び重要シーン３との時間間隔Δt_{45}がいずれも所定の閾値Ｔｔｈより大きいので、観客シーンＢは第１の方法によっては重要シーンに関連付けされない。但し、図６の例では、観客シーンＢは、第２の方法又は第３の方法のいずれかにより、重要シーン２に関連付けされている。

[0034] なお、上記の第１～第３の方法は、いずれか１つを使用してもよいし、２つ以上を組み合わせて使用してもよい。２つ以上を組み合わせて使用する場合の優先順位は任意に決定することができる。また、ダイジェスト生成装置１００は、素材映像から抽出した観客シーンの全てを重要シーンに関連付けし、ダイジェスト映像に含める必要はない。観客シーンが多い場合には、それらのうちのいくつかを選択して重要シーンに関連付けしてダイジェスト映像に含めてもよい。また、上記の第１～第３のうちの１つ又は複数の方法を

適用して関連付けの対象となった観客シーンのみを重要シーンに含め、関連付けの対象とならなかった観客シーンはダイジェスト映像に含めないこととしてもよい。

[0035] ［ダイジェスト生成装置］

（機能構成）

図７は、第１実施形態に係るダイジェスト生成装置１００の機能構成を示すブロック図である。ダイジェスト生成装置１００は、観客シーン抽出部２１と、観客シーンＤＢ２２と、重要シーン抽出部２３と、関連付け部２４と、ダイジェスト生成部２５とを備える。

[0036] 素材映像は、観客シーン抽出部２１及び重要シーン抽出部２３に入力される。観客シーン抽出部２１は、素材映像から観客シーンを抽出し、観客シーンＤＢ２２に保存する。観客シーンは、スポーツの試合映像などにおいて観客席を映した映像である。観客シーン抽出部２１は、例えばニューラルネットワークを用い、予め訓練済みのモデルを用いて観客シーンを抽出する。モデルの訓練方法については後述する。観客シーン抽出部２１は、ダイジェスト映像を生成するための前処理として素材映像から観客シーンを抽出し、観客シーンＤＢ２２に保存する。なお、観客シーン抽出部２１は、前述の第１の方法で用いる各観客シーンの時刻情報も付随情報として抽出し、観客シーンと対応付けて観客シーンＤＢ２２に保存する。また、観客シーン抽出部２１は、前述の第２の方法で用いる色に関する情報や文字列に関する情報も付随情報として抽出し、観客シーンと対応付けて観客シーンＤＢ２２に保存する。

[0037] 重要シーン抽出部２３は、図３を参照して説明した方法により、素材映像から重要シーンを抽出し、関連付け部２４に出力する。関連付け部２４は、観客シーンＤＢ２２に保存されている観客シーンを、重要シーン抽出部２３が抽出した重要シーンに関連付ける。具体的に、関連付け部２４は、前述の第１～第３の方法のいずれか又は複数の組み合わせを用いて観客シーンを重要シーンに関連付け、ダイジェスト生成部２５へ出力する。なお、関連付け

部２４は、観客シーンが関連付けされた重要シーンについては観客シーンと重要シーンのペアをダイジェスト生成部２５へ出力し、観客シーンが関連付けされなかった重要シーンについてはその重要シーンのみをダイジェスト生成部２５へ出力する。

[0038] ダイジェスト生成部２５は、関連付け部２４から入力された重要シーンを時系列につなげてダイジェスト映像を生成する。その際、ダイジェスト生成部２５は、各観客シーンを、関連付けされた重要シーンの前又は後に挿入する。なお、関連付け部２４は、各観客シーンを重要シーンの前後のいずれに配置するかを示す配置情報を生成し、観客シーン及び重要シーンとともにダイジェスト生成部２５に出力してもよい。この場合には、ダイジェスト生成部２５は入力された配置情報を参照して観客シーンの挿入位置を決定すればよい。こうして、ダイジェスト生成部２５は、観客シーンを含むダイジェスト映像を生成し、出力する。

[0039] （ダイジェスト映像生成処理）

図８は、ダイジェスト生成装置１００が実行するダイジェスト生成処理のフローチャートである。この処理は、図４に示すプロセッサ１２が予め用意されたプログラムを実行し、図７に示す各要素として動作することにより実現される。

[0040] まず、観客シーン抽出部２１は、前処理として観客シーン抽出処理を行う（ステップＳ１１）。図９は、観客シーン抽出処理のフローチャートである。まず、観客シーン抽出部２１は、素材映像を取得し（ステップＳ２１）、素材映像から観客シーンを検出する（ステップＳ２２）。そして、観客シーン抽出部２１は、観客シーンが検出されると（ステップＳ２３：Ｙｅｓ）、観客シーンＤＢ２２に保存する（ステップＳ２４）。次に、観客シーン抽出部２１は、ステップＳ２１～Ｓ２４の処理を素材映像の最後まで行ったか否かを判定し（ステップＳ２５：Ｎｏ）、最後まで行っていない場合は、ステップＳ２１～Ｓ２４を繰り返す。そして、観客シーン抽出部２１は、ステップＳ２１～Ｓ２４の処理を素材映像の最後まで行うと（ステップＳ２５：Ｙ

ｅｓ）、処理を終了する。これにより、素材映像に観客シーンが抽出される。また、観客シーンの付随情報として、各観客シーンの時刻、観客シーンに含まれる色や文字列に関する情報などが取得される。

[0041] 図８に戻り、重要シーン抽出部２３は素材映像から重要シーンを抽出する（ステップＳ１２）。次に、関連付け部２４は、抽出された重要シーンに対して、前述の第１～第３のいずれか又は複数を用いて、観客シーンＤＢ２２に保存されている観客シーンを関連付ける（ステップＳ１３）。関連付け部２４は、観客シーンが関連付けされた重要シーン、及び、観客シーンが関連付けされていない重要シーンをダイジェスト生成部２５へ出力する。そして、ダイジェスト生成部２５は、重要シーンを時系列につなげ、重要シーンの前又は後に対応する観客シーンを挿入してダイジェスト映像を生成する（ステップＳ１４）。こうして、ダイジェスト映像生成処理は終了する。

[0042] ［訓練装置］

次に、観客シーン抽出部２１が使用する観客シーン抽出モデルの訓練について説明する。図１０は、観客シーン抽出モデルＭｘを訓練する訓練装置の機能構成を示す。訓練装置２００は、観客シーン抽出モデルＭｘと、訓練部４ｘとを備える。また、観客シーン抽出モデルＭｘの訓練のための訓練データセットが用意される。訓練データセットは、訓練用素材映像と、正解データとを含む。正解データは、訓練用素材映像に含まれる観客シーンに正解を示す正解タグが付与されたデータである。

[0043] 観客シーン抽出モデルＭｘには、訓練用素材映像が入力される。観客シーン抽出モデルＭｘは、入力された訓練用素材映像から特徴量を抽出し、特徴量に基づいて観客シーンを抽出して訓練部４ｘに出力する。訓練部４ｘは、観客シーン抽出モデルＭｘが出力した観客シーンと、正解データとを用いて、観客シーン抽出モデルＭｘを最適化する。具体的には、訓練部４ｘは、観客シーン抽出モデルＭｘが抽出した観客シーンを、正解タグが付与されたシーンと比較して損失を算出し、損失が小さくなるように観客シーン抽出モデルＭｘのパラメータを更新する。こうして、訓練済みの観客シーン抽出モデ

ルＭｘが得られる。

[0044] （訓練処理）

図１１は、訓練装置２００による訓練処理のフローチャートである。この処理は、実際には図４に示すプロセッサ１２が予め用意されたプログラムを実行し、図１０に示す各要素として動作することにより実現される。まず、観客シーン抽出モデルＭｘは、訓練用素材映像から観客シーンを抽出する（ステップＳ３１）。次に、訓練部４ｘは、観客シーン抽出モデルＭｘから出力された観客シーンと、正解データとを用いて各モデルを最適化する（ステップＳ３２）。

[0045] 次に、訓練装置２００は、訓練終了条件が具備されたか否かを判定する（ステップＳ３３）。訓練終了条件とは、例えば予め用意された訓練データセットを全て使用したこと、訓練部４ｘが算出する損失の値が所定範囲内に収束したことなどである。こうして訓練終了条件が具備されるまで、観客シーン抽出モデルＭｘの訓練が行われ、訓練終了条件が具備されると、訓練処理は終了する。

[0046] ＜第２実施形態＞

次に、本発明の第２実施形態について説明する。図１２は、第２実施形態に係る映像処理装置の機能構成を示すブロック図である。図示のように、映像処理装置７０は、映像取得手段７１と、観客シーン抽出手段７２と、重要シーン抽出手段７３と、関連付け手段７４と、生成手段７５と、を備える。映像取得手段７１は、素材映像を取得する。観客シーン抽出手段７２は、素材映像から、観客を映した観客シーンを抽出する。重要シーン抽出手段７３は、素材映像から重要シーンを抽出する。関連付け手段７４は、観客シーンを重要シーンに関連付ける。生成手段７５は、重要シーンと、重要シーンに関連付けされた観客シーンとを含むダイジェスト映像を生成する。

[0047] 上記の実施形態の一部又は全部は、以下の付記のようにも記載されうるが、以下には限られない。

[0048] （付記１）

素材映像を取得する映像取得手段と、

前記素材映像から、観客を映した観客シーンを抽出する観客シーン抽出手段と、

前記素材映像から重要シーンを抽出する重要シーン抽出手段と、

前記観客シーンを前記重要シーンに関連付ける関連付け手段と、

前記重要シーンと、前記重要シーンに関連付けされた観客シーンとを含むダイジェスト映像を生成する生成手段と、

を備える映像処理装置。

[0049] （付記２）

前記生成手段は、前記重要シーンを時系列に配置して前記ダイジェスト映像を生成し、

前記生成手段は、前記重要シーンに関連付けされた観客シーンを、当該重要シーンの前又は後に配置して前記ダイジェスト映像を生成する付記１に記載の映像処理装置。

[0050] （付記３）

前記関連付け手段は、前記重要シーンの前後の所定時間以内の位置に存在する観客シーンを、当該重要シーンに関連付けする付記１又は２に記載の映像処理装置。

[0051] （付記４）

前記観客シーン抽出手段は、前記観客シーンに含まれる色に関する情報を抽出し、

前記関連付け手段は、前記色に関する情報に基づいて、当該観客シーンを前記重要シーンに関連付けする付記１乃至３のいずれか一項に記載の映像処理装置。

[0052] （付記５）

前記素材映像は、スポーツの映像であり、

前記観客シーン抽出手段は、前記観客シーンに含まれる人々の服装又は人々が所持する物体の色を抽出し、

前記関連付け手段は、前記観客シーンを、当該観客シーンから抽出された色をチームカラーとするチームを映した重要シーンに関連付けする付記１乃至３のいずれか一項に記載の映像処理装置。

[0053] （付記６）

前記観客シーン抽出手段は、前記観客シーンに含まれる文字列を抽出し、

前記関連付け手段は、前記文字列に基づいて、前記観客シーンを前記重要シーンに関連付けする付記１乃至５のいずれか一項に記載の映像処理装置。

[0054] （付記７）

前記素材映像は、スポーツの映像であり、

前記観客シーン抽出手段は、前記観客シーンに含まれるメッセージボード、もしくは、前記観客シーンに含まれる人物が着用又は所持している物体が示す文字列を抽出し、

前記関連付け手段は、前記観客シーンを、当該観客シーンから抽出された文字列が示すチーム、又は、当該文字列が示す選手が属するチームを映した重要シーンに関連付けする付記１乃至５のいずれか一項に記載の映像処理装置。

[0055] （付記８）

前記観客シーン抽出手段は、予め用意された訓練用素材映像と、当該訓練用素材映像中の観客シーンを示す正解データとを含む訓練データセットを用いて訓練されたモデルを用いて、前記観客シーンを抽出する付記１乃至７のいずれか一項に記載の映像処理装置。

[0056] （付記９）

素材映像を取得し、

前記素材映像から、観客を映した観客シーンを抽出し、

前記素材映像から重要シーンを抽出し、

前記観客シーンを前記重要シーンに関連付け、

前記重要シーンと、前記重要シーンに関連付けされた観客シーンとを含むダイジェスト映像を生成する映像処理方法。

[0057] （付記１０）

素材映像を取得し、

前記素材映像から、観客を映した観客シーンを抽出し、

前記素材映像から重要シーンを抽出し、

前記観客シーンを前記重要シーンに関連付け、

前記重要シーンと、前記重要シーンに関連付けされた観客シーンとを含むダイジェスト映像を生成する処理をコンピュータに実行させるプログラムを記録した記録媒体。

[0058] 以上、実施形態及び実施例を参照して本発明を説明したが、本発明は上記実施形態及び実施例に限定されるものではない。本発明の構成や詳細には、本発明のスコープ内で当業者が理解し得る様々な変更をすることができる。

符号の説明

[0059] ２　素材映像ＤＢ

３、３ｘ　正解データ

４、４ｘ　訓練部

５、２５　ダイジェスト生成部

１２　プロセッサ

２１　観客シーン抽出部

２２　観客シーンＤＢ

２３　重要シーン抽出部

２４　関連付け部

１００　ダイジェスト生成装置

２００　訓練装置

請求の範囲

[請求項1] 素材映像を取得する映像取得手段と、

前記素材映像から、観客を映した観客シーンを抽出する観客シーン抽出手段と、

前記素材映像から重要シーンを抽出する重要シーン抽出手段と、

前記観客シーンを前記重要シーンに関連付ける関連付け手段と、

前記重要シーンと、前記重要シーンに関連付けされた観客シーンとを含むダイジェスト映像を生成する生成手段と、

を備える映像処理装置。

[請求項2] 前記生成手段は、前記重要シーンを時系列に配置して前記ダイジェスト映像を生成し、

前記生成手段は、前記重要シーンに関連付けされた観客シーンを、当該重要シーンの前又は後に配置して前記ダイジェスト映像を生成する請求項１に記載の映像処理装置。

[請求項3] 前記関連付け手段は、前記重要シーンの前後の所定時間以内の位置に存在する観客シーンを、当該重要シーンに関連付けする請求項１又は２に記載の映像処理装置。

[請求項4] 前記観客シーン抽出手段は、前記観客シーンに含まれる色に関する情報を抽出し、

前記関連付け手段は、前記色に関する情報に基づいて、当該観客シーンを前記重要シーンに関連付けする請求項１乃至３のいずれか一項に記載の映像処理装置。

[請求項5] 前記素材映像は、スポーツの映像であり、

前記観客シーン抽出手段は、前記観客シーンに含まれる人々の服装又は人々が所持する物体の色を抽出し、

前記関連付け手段は、前記観客シーンを、当該観客シーンから抽出された色をチームカラーとするチームを映した重要シーンに関連付けする請求項１乃至３のいずれか一項に記載の映像処理装置。

[請求項6] 前記観客シーン抽出手段は、前記観客シーンに含まれる文字列を抽出し、

前記関連付け手段は、前記文字列に基づいて、前記観客シーンを前記重要シーンに関連付けする請求項１乃至５のいずれか一項に記載の映像処理装置。

[請求項7] 前記素材映像は、スポーツの映像であり、

前記観客シーン抽出手段は、前記観客シーンに含まれるメッセージボード、もしくは、前記観客シーンに含まれる人物が着用又は所持している物体が示す文字列を抽出し、

前記関連付け手段は、前記観客シーンを、当該観客シーンから抽出された文字列が示すチーム、又は、当該文字列が示す選手が属するチームを映した重要シーンに関連付けする請求項１乃至５のいずれか一項に記載の映像処理装置。

[請求項8] 前記観客シーン抽出手段は、予め用意された訓練用素材映像と、当該訓練用素材映像中の観客シーンを示す正解データとを含む訓練データセットを用いて訓練されたモデルを用いて、前記観客シーンを抽出する請求項１乃至７のいずれか一項に記載の映像処理装置。

[請求項9] 素材映像を取得し、

前記素材映像から、観客を映した観客シーンを抽出し、

前記素材映像から重要シーンを抽出し、

前記観客シーンを前記重要シーンに関連付け、

前記重要シーンと、前記重要シーンに関連付けされた観客シーンとを含むダイジェスト映像を生成する映像処理方法。

[請求項10] 素材映像を取得し、

前記素材映像から、観客を映した観客シーンを抽出し、

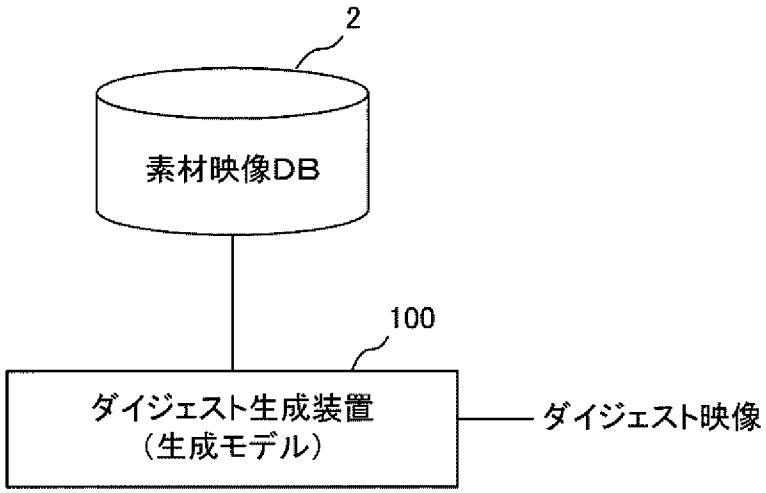
前記素材映像から重要シーンを抽出し、

前記観客シーンを前記重要シーンに関連付け、

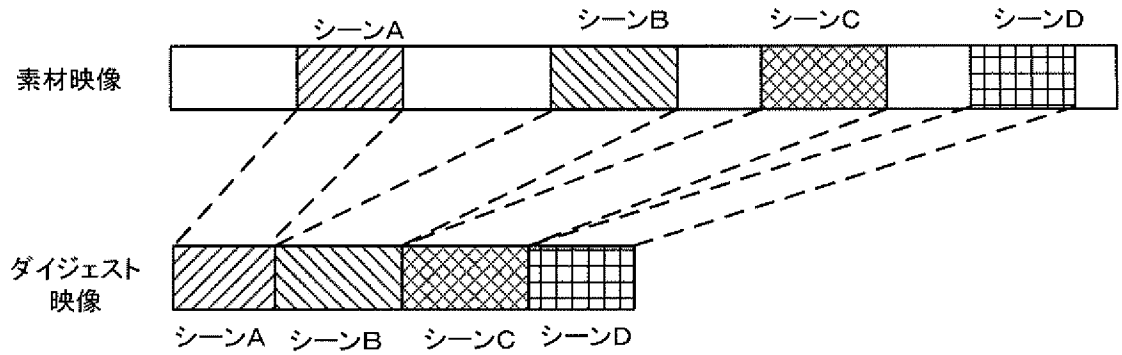
前記重要シーンと、前記重要シーンに関連付けされた観客シーンと

を含むダイジェスト映像を生成する処理をコンピュータに実行させるプログラムを記録した記録媒体。

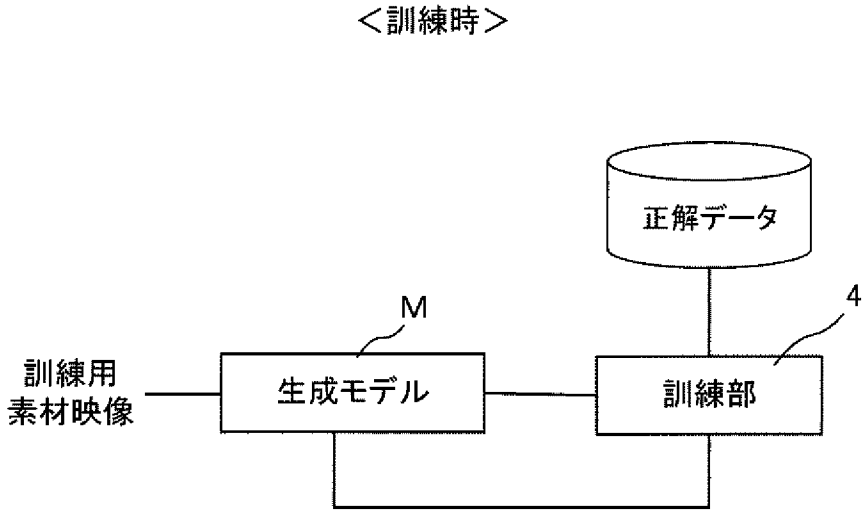
[図1]


2
素材映像DB
100
ダイジェスト生成装置
（生成モデル）
ダイジェスト映像

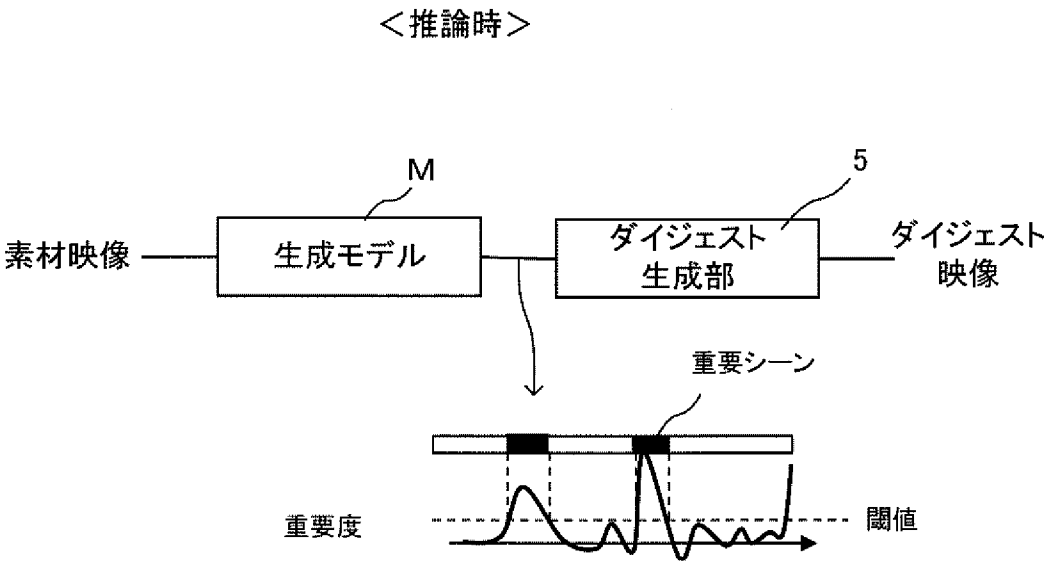
[図2]


シーンA
シーンB
シーンC
シーンD
素材映像
ダイジェスト
映像
シーンA
シーンB
シーンC
シーンD

[図3]

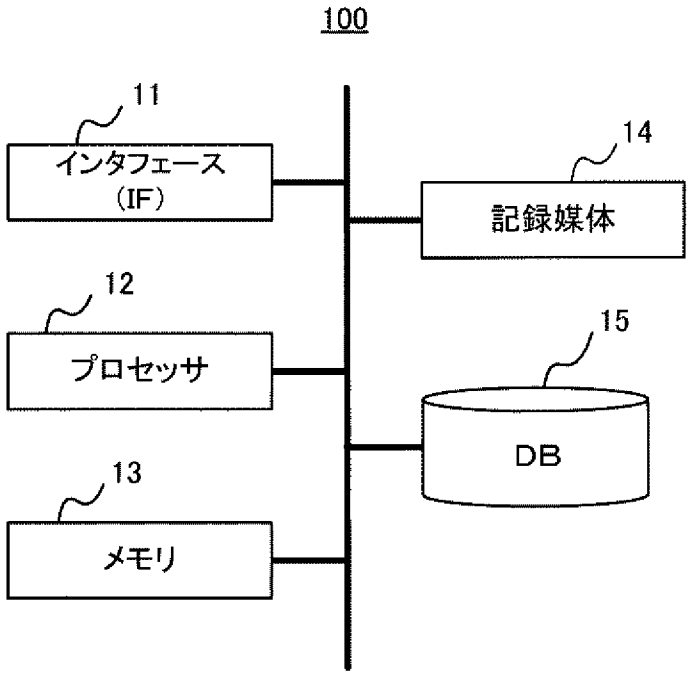

＜訓練時＞
正解データ
M
4
訓練用
素材映像
生成モデル
訓練部

(A)


＜推論時＞
M
5
素材映像
生成モデル
ダイジェスト
生成部
ダイジェスト
映像
重要シーン
重要度
閾値

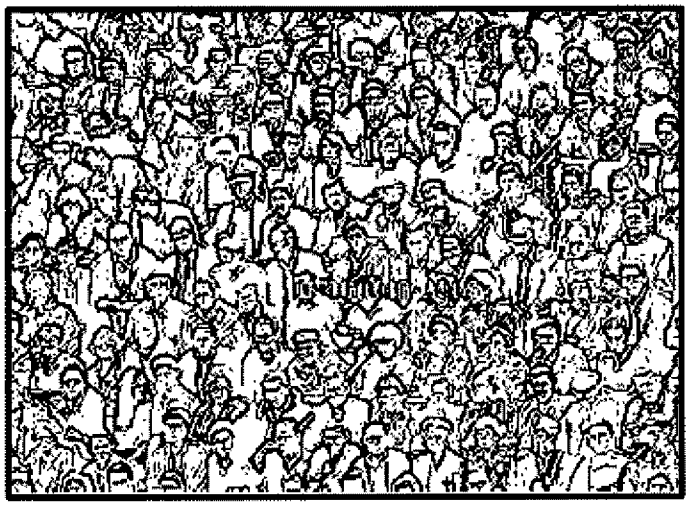
(B)

[図4]



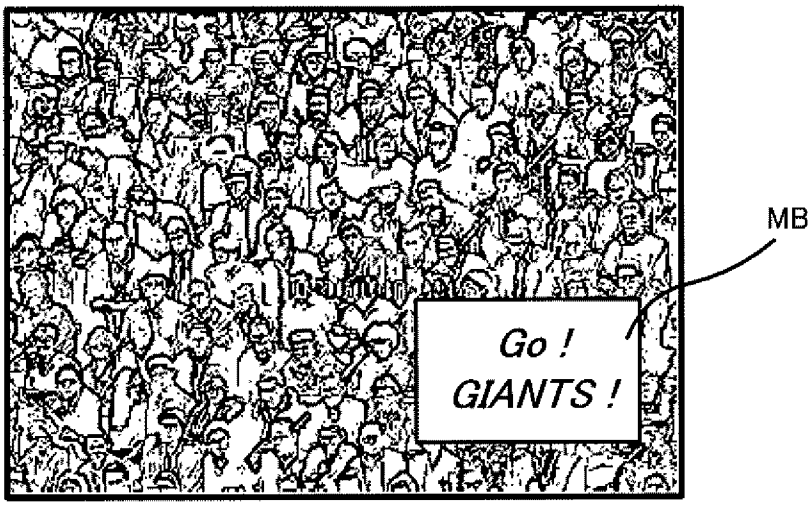
[図5]

（観客席映像）



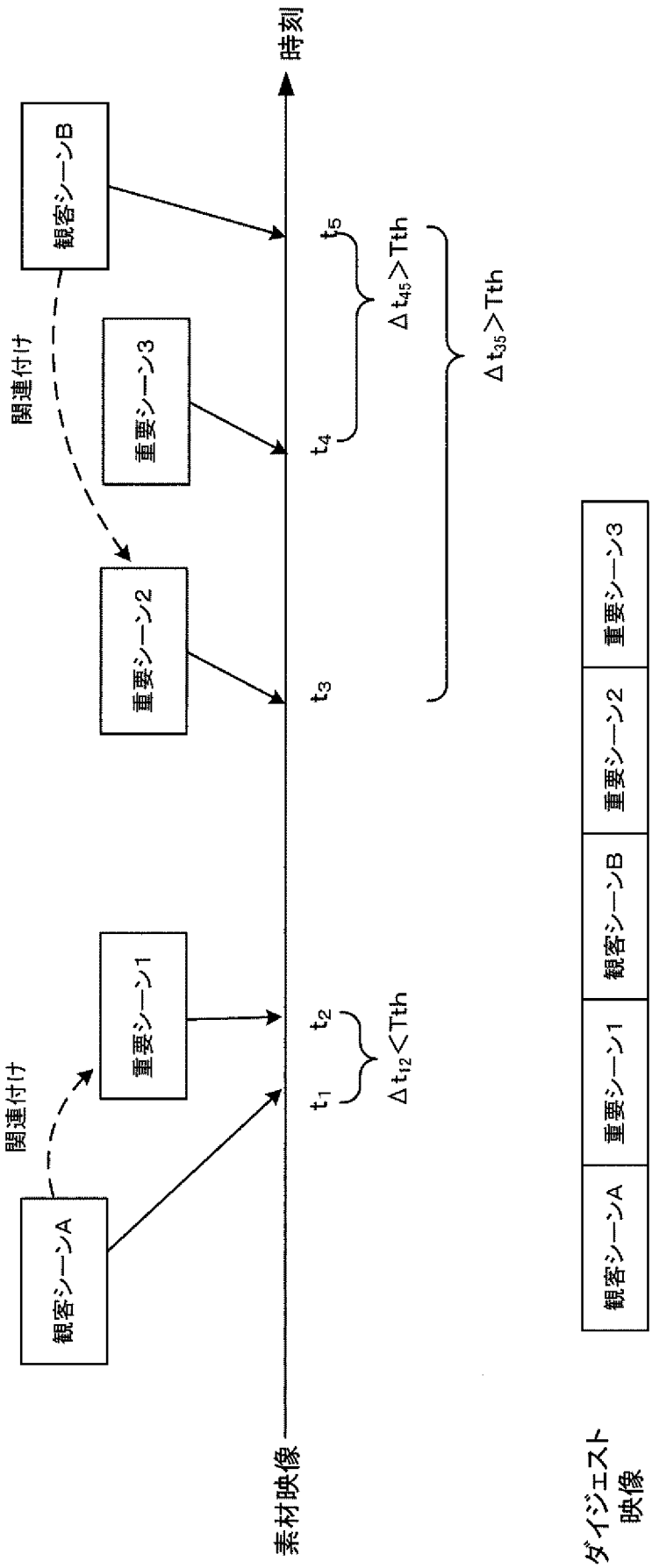
（A）

（観客席映像）

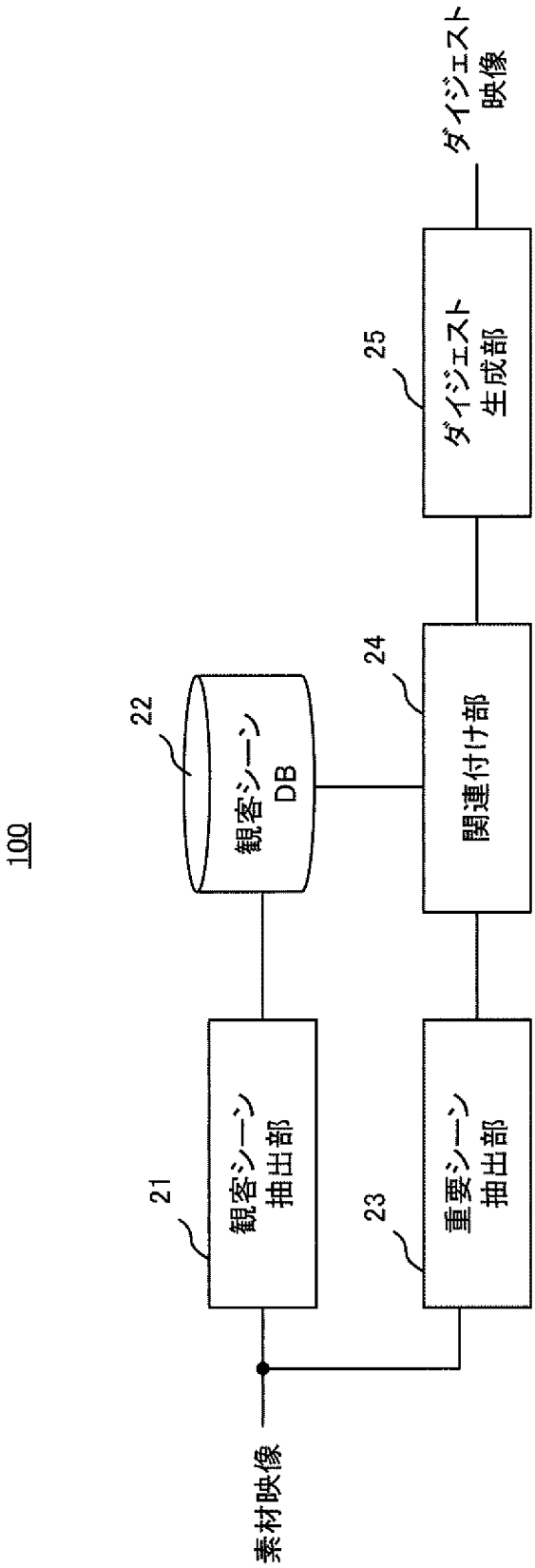


（B）

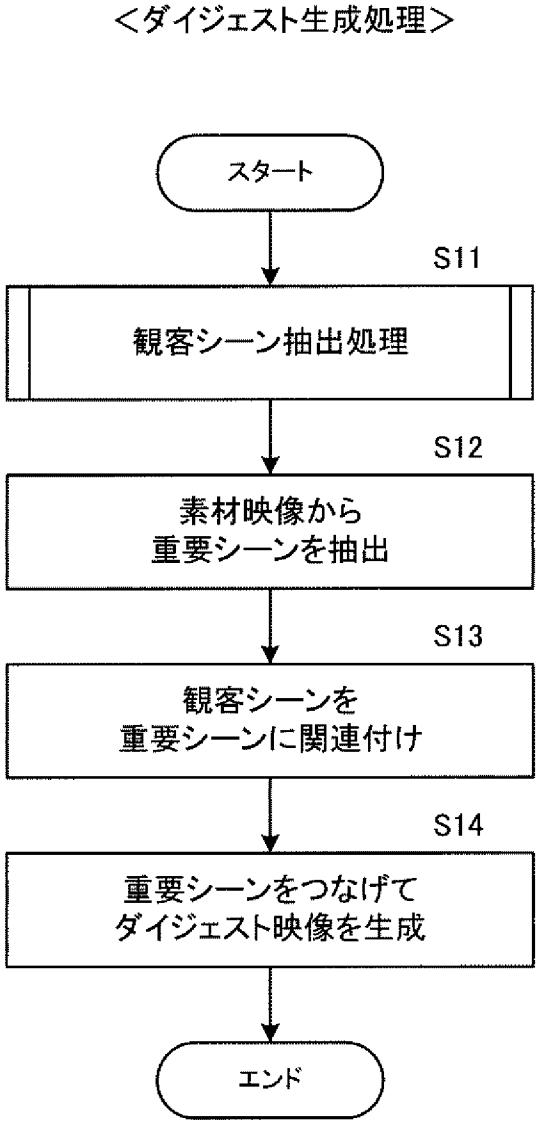
[図6]


時刻
観客シーンB
重要シーン3
重要シーン2
重要シーン1
観客シーンA
関連付け
関連付け
t1
t2
t3
t4
t5
Δt12<Tth
Δt45>Tth
Δt35>Tth
素材映像
観客シーンA
重要シーン1
観客シーンB
重要シーン2
重要シーン3
ダイジェスト映像

[図7]

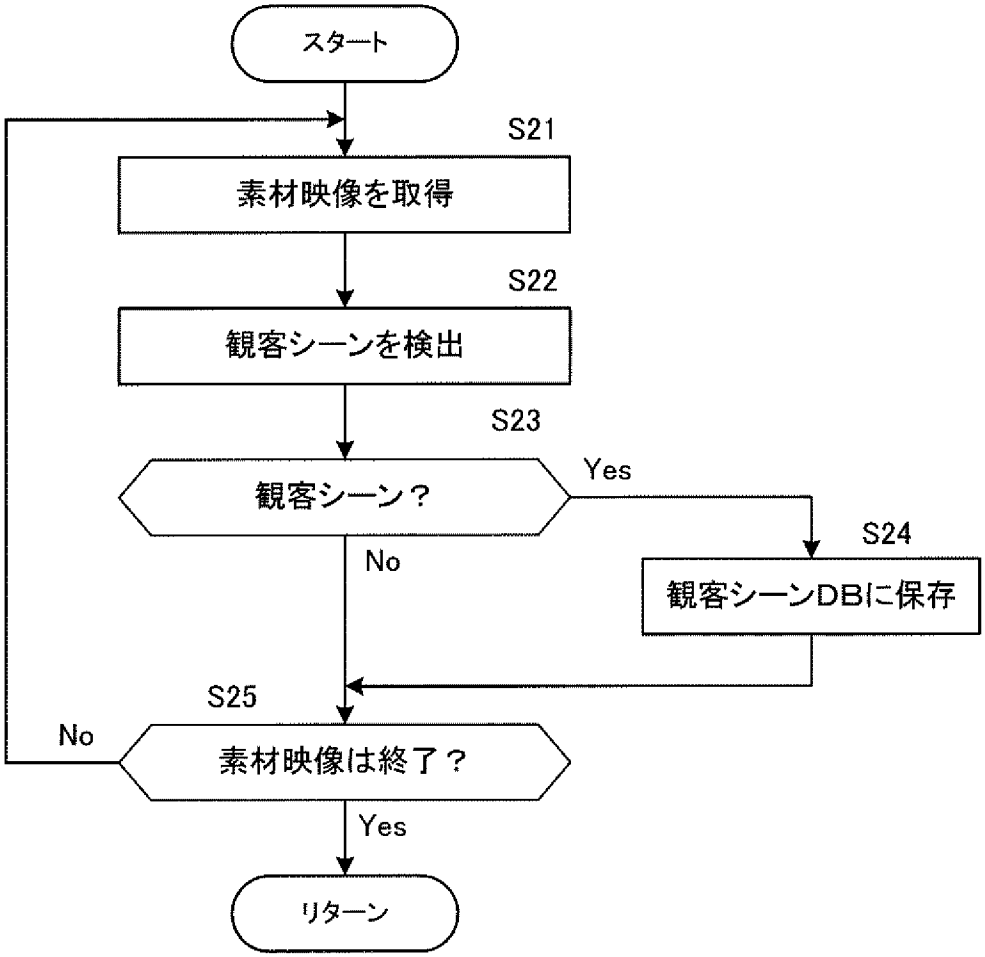

100
素材映像
観客シーン
抽出部
21
重要シーン
抽出部
23
観客シーン
DB
22
関連付け部
24
ダイジェスト
生成部
25
ダイジェスト
映像

[図8]

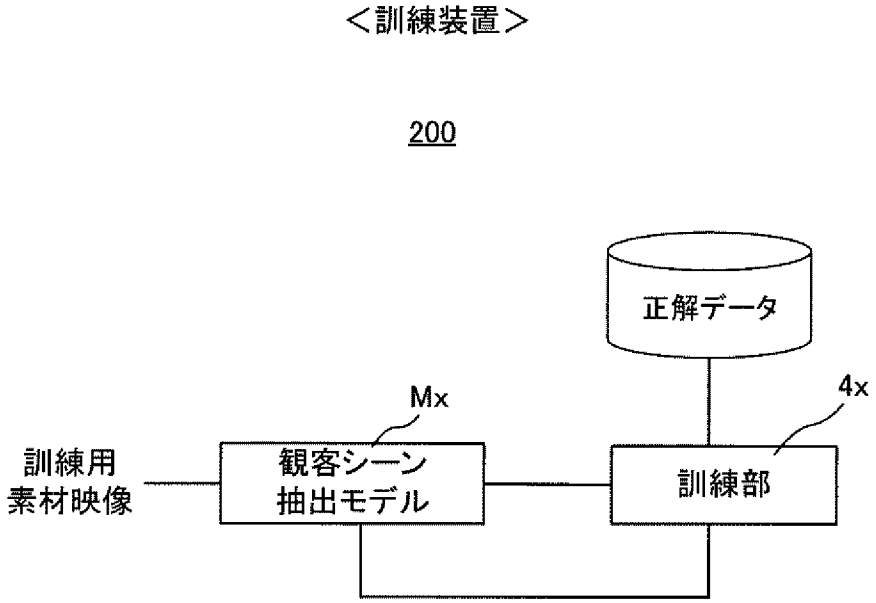

<ダイジェスト生成処理>
スタート
S11
観客シーン抽出処理
S12
素材映像から
重要シーンを抽出
S13
観客シーンを
重要シーンに関連付け
S14
重要シーンをつなげて
ダイジェスト映像を生成
エンド

[図9]

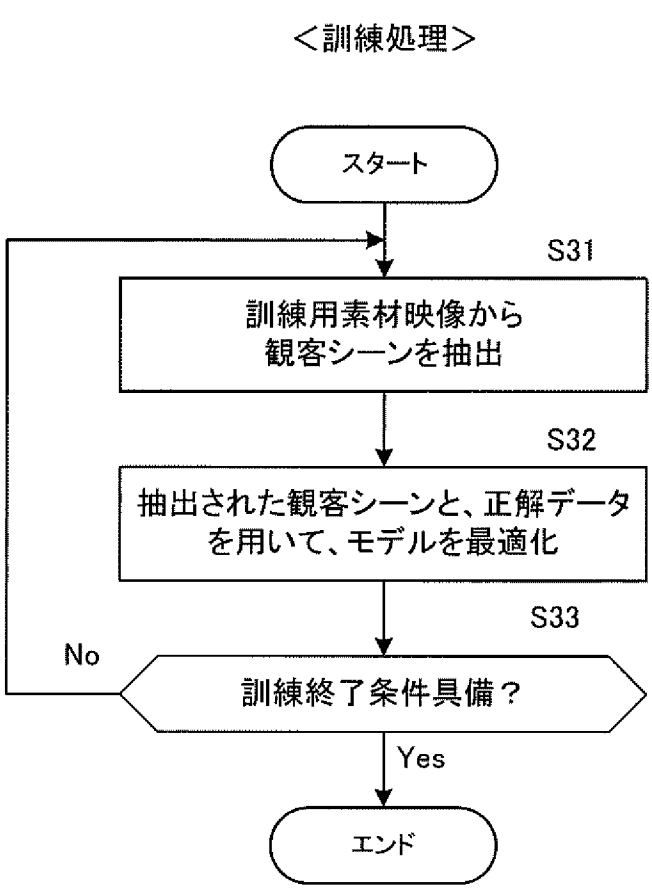
＜観客シーン抽出処理＞



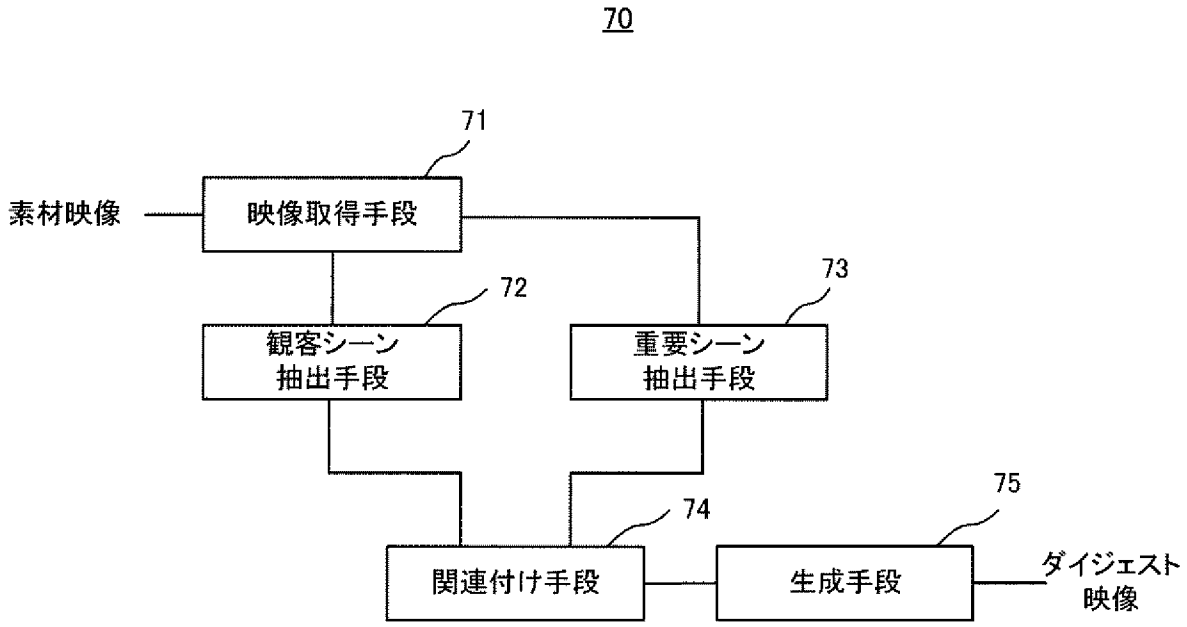
[図10]


＜訓練装置＞
200
正解データ
Mx
4x
訓練用
素材映像
観客シーン
抽出モデル
訓練部

[図11]


＜訓練処理＞
スタート
S31
訓練用素材映像から
観客シーンを抽出
S32
抽出された観客シーンと、正解データ
を用いて、モデルを最適化
S33
No
訓練終了条件具備？
Yes
エンド

[図12]


70
71
素材映像
映像取得手段
72
観客シーン
抽出手段
73
重要シーン
抽出手段
74
関連付け手段
75
生成手段
ダイジェスト
映像

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2020/020868

A. CLASSIFICATION OF SUBJECT MATTER
Int. Cl. H04N5/91(2006.01)i
FI: H04N5/91

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)
Int. Cl. H04N5/91

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Published examined utility model applications of Japan 1922-1996
Published unexamined utility model applications of Japan 1971-2020
Registered utility model specifications of Japan 1996-2020
Published registered utility model applications of Japan 1994-2020

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.
X Y A	WO 2006/008923 A2 (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.) 26 January 2006, claim 1, paragraphs [0015]-[0019], [0022]-[0032], [0053], [0054], [0075], [0076], fig. 1-5, 18, claim 1, paragraphs [0015]-[0019], [0022]-[0032], [0053], [0054], [0075], [0076], fig. 1-5, 18, claim 1, paragraphs [0015]-[0019], [0022]-[0032], [0053], [0054], [0075], [0076], fig. 1-5, 18	1-3, 9-10 8 4-7
Y A	JP 2014-229092 A (NIKON CORP.) 08 December 2014, paragraphs [0002], [0019], [0036]-[0046], fig. 1, paragraphs [0002], [0019], [0036]-[0046], fig. 1	8 1-7, 9-10

☒ 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	Date of mailing of the international search report
20.08.2020	01.09.2020
Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2020/020868

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2011-504702 A (KONINKLIJKE PHILIPS ELECTRONICS N.V.) 10 February 2011, paragraphs [0037]-[0039], [0042]-[0044], [0048], fig. 3	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2020/020868

Patent Documents referred to in the Report	Publication Date	Patent Family	Publication Date
WO 2006/008923 A2	26.01.2006	US 2008/0193099 A1 claim 1, paragraphs [0019]-[0021], [0051]-[0073], [0108]-[0112], [0147], [0148], fig. 1-5, 18 EP 1765011 A2 CN 1977530 A	
JP 2014-229092 A	08.12.2014	(Family: none)	
JP 2011-504702 A	10.02.2011	US 2010/0289959 A1 paragraphs [0050]-[0052], [0055]-[0058], [0061], fig. 3 WO 2009/066213 A1 EP 2227758 A1 CN 101868795 A KR 10-2010-0097173 A	

国際調査報告

国際出願番号
PCT/JP2020/020868

A. 発明の属する分野の分類（国際特許分類（ＩＰＣ））

H04N 5/91(2006.01)i
FI: H04N5/91

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））

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	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０２０年
日本国実用新案登録公報	１９９６－２０２０年
日本国登録実用新案公報	１９９４－２０２０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
X	WO 2006/008923 A2 (松下電器産業株式会社) 26.01.2006 (2006-01-26) 請求項1, 段落[0015]-[0019], [0022]-[0032], [0053]-[0054], [0075]-[0076], 図1-5, 18	1-3, 9-10
Y	請求項1, 段落[0015]-[0019], [0022]-[0032], [0053]-[0054], [0075]-[0076], 図1-5, 18	8
A	請求項1, 段落[0015]-[0019], [0022]-[0032], [0053]-[0054], [0075]-[0076], 図1-5, 18	4-7
Y	JP 2014-229092 A (株式会社ニコン) 08.12.2014 (2014-12-08) 段落[0002], [0019], [0036]-[0046], 図1	8
A	段落[0002], [0019], [0036]-[0046], 図1	1-7, 9-10
A	JP 2011-504702 A (コーニンクレッカ フィリップス エレクトロニクス エヌ ヴィ) 10.02.2011 (2011-02-10) 段落[0037]-[0039], [0042]-[0044], [0048], 図3	1-10

☐ 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抵触するものではなく、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	国際調査報告の発送日
20.08.2020	01.09.2020

名称及びあて先

日本国特許庁(ISA/JP)
〒100-8915
日本国
東京都千代田区霞が関三丁目４番３号

権限のある職員（特許庁審査官）

川中　龍太 5C 5892

電話番号 03-3581-1101 内線 3541

国際調査報告
パテントファミリーに関する情報

国際出願番号
PCT/JP2020/020868

引用文献	公表日	パテントファミリー文献	公表日
WO 2006/008923 A2	26.01.2006	US 2008/0193099 A1 claim1, paragraphs[0019]-[0021], [0051]-[0073], [0108]-[0112], [0147]-[0148], Figs.1-5, 18 EP 1765011 A2 CN 1977530 A	
JP 2014-229092 A	08.12.2014	（ファミリーなし）	
JP 2011-504702 A	10.02.2011	US 2010/0289959 A1 paragraphs[0050]-[0052], [0055]-[0058], [0061], Fig.3 WO 2009/066213 A1 EP 2227758 A1 CN 101868795 A KR 10-2010-0097173 A	