(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 2 月 13 日(13.02.2014)

WIPO | PCT

(10) 国際公開番号
WO 2014/024475 A1

(51) 国際特許分類:
H04N 21/2668 (2011.01) *H04N 21/2187* (2011.01)
H04N 5/265 (2006.01)

(21) 国際出願番号: PCT/JP2013/004742

(22) 国際出願日: 2013 年 8 月 6 日(06.08.2013)

(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 2012-178623 2012 年 8 月 10 日(10.08.2012) 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: 佐々木　泰治 (SASAKI, Taiji). 矢羽田　洋 (YAHATA, Hiroshi).

(74) 代理人: 新居　広守 (NII, Hiromori); 〒5320011 大阪府大阪市淀川区西中島５丁目３番１０号タナカ・イトーピア新大阪ビル６階新居国際特許事務所内 Osaka (JP).

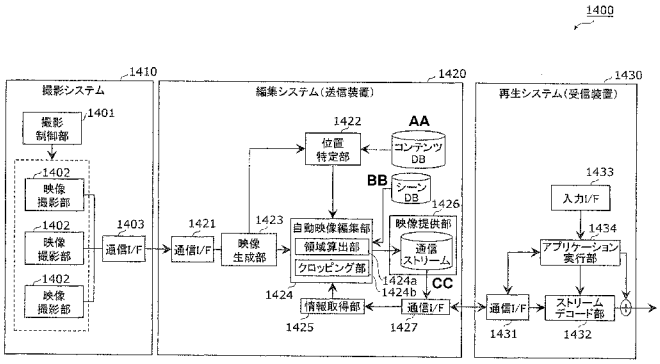
(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, 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, 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, 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:
— 国際調査報告（条約第 21 条(3)）

(54) **Title**: VIDEO PROVISION METHOD, TRANSMISSION DEVICE, AND RECEPTION DEVICE

(54) 発明の名称 : 映像提供方法、送信装置および受信装置



1401 Imaging controller
1402 Video imaging unit
1403, 1421, 1427, 1431 Communication interface
1410 Imaging system
1420 Editing system (transmission device)
1422 Position identification unit
1423 Video generation unit
1424 Automatic video editing unit
1424a Area calculation unit
1424b Cropping unit
1425 Information acquisition unit
1426 Video provision unit
1430 Playback system (reception device)
1432 Stream decoder
1433 Input interface
1434 Application execution unit
AA Content database
BB Scene database
CC Communication stream

(57) **Abstract**: Provided is a video provision method which includes: a video acquisition step (S2201) for acquiring a first main video in which a first imaging space is imaged, and a second main video in which a second imaging space including a space other than the first space is imaged; a video generation step (S2202) in which the plurality of main videos acquired in the video acquisition step are synthesized to generate a wide angle video; an information acquisition step (S2203) in which user preference information is acquired via a network; an area calculation step (S2205) in which a cropping area in the wide angle video is calculated on the basis of the user preference information acquired in the information acquisition step, said cropping area being smaller than the wide-angle-video area; a cropping step in which the cropping area is used to crop the wide angle video; and a video provision step (S2206) in which a cropped video generated by being subjected to cropping is provided to a user.

(57) 要約: 映像提供方法であって、第一撮影空間が撮影された第一主映像と、第一空間以外の空間を含む第二撮影空間が撮影された第二主映像と、を取得する映像取得ステップ（Ｓ２２０１）と、映像取得ステップにおいて取得された複数の主映像を合成することにより広角映像を生成する映像生成ステップ（Ｓ２２０２）と、ネットワークを経由して、ユーザの嗜好情報を取得する情報取得ステップ（Ｓ２２０３）と、情報取得ステップにおいて取得されたユーザの嗜好情報に基づいて、広角映像のうちの領域であって、広角映像の領域よりも小さいクロッピング領域を算出する領域算出ステップ（Ｓ２２０５）と、広角映像をクロッピング領域でクロッピングするクロッピングステップと、クロッピングされることにより生成されたクロッピング映像をユーザに提供する映像提供ステップ（Ｓ２２０６）と、を含む。

WO 2014/024475 A1

明　細　書

発明の名称 ： 映像提供方法、送信装置および受信装置

技術分野

[0001] 本発明は、映像コンテンツの作成および伝送し再生するための、映像提供方法、送信装置および受信装置に関する。

背景技術

[0002] 従来、映像コンテンツの配信は、放送局が制作した映像コンテンツが、放送波を通じて一般家庭に配信されることにより行われてきた。放送のデジタル化が進み、一般家庭のテレビにおいて、高画質なハイビジョンの映像を楽しむことが可能となっており、バラエティやドラマ、スポーツ等様々なジャンルの映像コンテンツの制作および配信がなされている。

[0003] また、一方でブロードバンド環境が普及することで、映像コンテンツの配信をインターネット経由で行うサービスが広く普及し始めている。例えば、特許文献１では個人の映像コンテンツをアップロードして共有可能なサーバについての開示が有り、ユーザはインターネット経由で個人がアップロードされている映像コンテンツを選択して視聴することができる。また、特許文献２では、カメラで撮影した映像をそのままインターネットにアップロード可能なシステムについての開示が有り、ユーザは配信されているライブ映像をインターネット経由で再生して楽しむことができる。

先行技術文献

特許文献

[0004] 特許文献1：特開２００８－２８９７０号公報

特許文献2：特開２００５－２７７９４９号公報

発明の概要

発明が解決しようとする課題

[0005] しかしながら、放送局による映像コンテンツの配信・視聴システムおよびインターネットによる動画配信サービスでは、ユーザの嗜好が反映された映

像を視聴することはできない。

[0006] そこで、本発明の目的は、かかる問題に鑑みてなされたものであって、ユーザの嗜好が反映された映像を提供することができる映像提供方法などを提供することにある。

課題を解決するための手段

[0007] 上記目的を達成するために、本発明の一態様に係る映像提供方法は、コンピュータが映像をユーザに提供するための映像提供方法であって、（ｉ）撮影空間のうちの一部の第一撮影空間が撮影された第一主映像と、（ｉｉ）前記撮影空間のうちの一部の空間であって、前記第一空間以外の空間を含む第二撮影空間が撮影された第二主映像と、を取得する映像取得ステップと、前記映像取得ステップにおいて取得された前記第一主映像および前記第二主映像を合成することにより広角映像を生成する映像生成ステップと、ネットワークを経由して、前記ユーザの嗜好情報を取得する情報取得ステップと、前記情報取得ステップにおいて取得された前記ユーザの嗜好情報に基づいて、前記広角映像のうちの一部の領域であって、前記広角映像の領域よりも小さいクロッピング領域を算出する領域算出ステップと、前記映像生成ステップにおいて生成された前記広角映像を、前記領域算出ステップにおいて算出された前記クロッピング領域でクロッピングするクロッピングステップと、前記クロッピングステップにおいてクロッピングされることにより生成されたクロッピング映像を前記ユーザに提供する映像提供ステップと、を含む。

発明の効果

[0008] 本発明の映像提供方法、送信装置および受信装置は、ユーザの意図を反映させた映像コンテンツの自動生成が可能となるため、ユーザは映像コンテンツを個人の好きなように好きな見方で楽しむことが可能となる。

図面の簡単な説明

[0009] [図1]図１は、放送波による映像コンテンツの配信・視聴システムの構成を示す図である。

[図2]図２は、実施の形態１における再生装置の使用形態を説明するための図

である。

[図3]図３は、トランスポートストリーム形式のデジタルストリームの構成を示す図である。

[図4]図４は、ビデオストリームの構造を説明するための図である。

[図5]図５は、ビデオストリームのアクセスユニットの内部構成を説明するための図である。

[図6]図６は、クロッピング領域情報とスケーリング情報を説明するための図である。

[図7]図７は、クロッピング領域情報とスケーリング情報の具体的な指定方法を説明するための図である。

[図8]図８は、ＰＥＳパケットの構成を説明するための図である。

[図9]図９は、トランスポートストリームを構成するＴＳパケットのデータ構造を示す図である。

[図10]図１０は、ＰＭＴのデータ構造を説明するための図である。

[図11]図１１は、ビデオストリームの参照関係を説明するための図である。

[図12]図１２は、ソースパケットの構造を説明するための図である。

[図13]図１３は、ＴＳストリームからＴＴＳストリームへの変換を説明するための図である。

[図14]図１４は、個人嗜好を反映した映像コンテンツの配信・視聴システムを説明するための図である。

[図15]図１５は、複数の映像撮影部によるイベントの広角撮影方法を説明するための図である。

[図16]図１６は、広角映像の生成方法を説明するための図である。

[図17]図１７は、映像の位置情報からコートの位置情報に変換する方法を説明するための図である。

[図18]図１８は、ユーザの嗜好情報に基づいた映像コンテンツの生成方法を説明するための図である。

[図19]図１９は、広角映像からの映像切り出し（クロッピング）方法を説明

するための図である。

[図20]図２０は、広角映像からの映像クロッピング方法の変形例を説明するための図である。

[図21]図２１は、音声データの生成方法を説明するための図である。

[図22]図２２は、編集システムにより行われる映像提供処理の流れを示すフローチャートである。

[図23]図２３は、映像撮影部から編集システムへの伝送の変形例を説明するための図である。

[図24]図２４は、撮影制御部の変形例を説明するための図である。

[図25]図２５は、被写体位置情報の補正例を説明するための図である。

[図26]図２６は、違和感のない映像を生成するための広角映像からのクロッピング方法を説明するための図である。

[図27]図２７は、クロッピング領域のサイズ変形の方法を説明するための図である。

[図28]図２８は、ユーザの嗜好情報として複数の対象が設定される場合のクロッピング方法を説明するための図である。

[図29]図２９は、実施の形態２に係る個人嗜好を反映した映像コンテンツの自動生成・視聴システムを説明するための図である。

[図30]図３０は、スポット映像撮影部の配置例を説明するための図である。

[図31]図３１は、自動映像選択編集部による編集例を説明するための図である。

[図32]図３２は、シーン情報を利用した自動映像選択編集部による編集例を説明するための図である。

[図33]図３３は、シーン区切りの方法を説明するための図である。

[図34]図３４は、シーン区切りのアルゴリズムのフローチャートを説明するための図である。

[図35]図３５は、リプレイ映像に利用する映像シーンの時間関係を説明するための図である。

[図36]図３６は、自動映像選択編集部による選手の動きベクトルを使った映像選択の例を説明するための図である。

[図37]図３７は、許容遅延量が設定される場合のシーン区切りの例を説明するための図である。

[図38]図３８は、シーン区切りを攻守交替によって行う方法を説明するための図である。

[図39]図３９は、本実施の形態１および２を適用する応用例１を説明するための図である。

[図40]図４０は、本実施の形態を適用する応用例２を説明するための図である。

[図41]図４１は、本実施の形態の視聴システムの変形例を説明するための図である。

[図42]図４２は、ユーザ嗜好データを反映させるユーザインターフェースを説明するための図である。

[図43]図４３は、複数のテレビで広角映像を表示する構成１を説明するための図である。

[図44]図４４は、複数のテレビで広角映像を表示する構成２を説明するための図である。

[図45]図４５は、ユーザ嗜好を反映したハイライト再生の実現方法を説明するための図である。

[図46]図４６は、複数カメラを使った人物認識方法を説明するための図である。

[図47]図４７は、年代ごとに顔認証データベースを持つ構成を説明するための図である。

[図48]図４８は、電子コミックを配信する方法を説明するための図である。

発明を実施するための形態

[0010] 以下、本発明の実施の形態について、図面を参照しながら説明する。

[0011] （本発明の基礎となった知見）

本発明者らは、「背景技術」の欄において記載した、配信・視聴システムに関し、以下の問題が生じることを見出した。

[0012] 図１を用いて背景技術として放送波による映像コンテンツの配信・視聴システムを説明する。配信・視聴システム１０は、図１に示すように、映像コンテンツを制作し送信する放送局のシステムである放送システム１００と、放送波から映像コンテンツを受信する再生装置１１０とから構成される。

[0013] 放送システム１００は、放送映像撮影部１０１、放送映像編集部１０２、および放送ストリーム作成部１０３から構成される。

[0014] 放送映像撮影部１０１は、主に放送局のビデオカメラを指し、映像を撮影し、かつ、音声を集音する（以下、単に「映像を撮影する」という。）。つまり、映像は、一般に複数のカメラマンが様々な角度から放送映像撮影部１０１を利用することにより、撮影される。例えば、サッカーのコンテンツを作成する場合には、サッカーのピッチの俯瞰映像、選手をズームアップした映像、ゴール裏などからの別視点映像など様々な視点からの映像を撮影するために、様々な位置でカメラマンが放送映像撮影部１０１を使用して撮影を行う。

[0015] 放送映像編集部１０２は、放送映像撮影部１０１によって撮影されることにより記録された映像および音声を、編集する。具体的には、複数の放送映像撮影部１０１によって撮影された映像の中で、放送するシーンの選択、撮影した映像にスコア情報や字幕情報などのグラフィックスを重ねあわせる画像処理などが放送映像編集部１０２により行われる。複数の放送映像撮影部１０１によって撮影された映像の中から放送するシーンの映像の選択は、シーンの選択を専門に行うディレクタにより行われる。ディレクタは、撮影されたコンテンツの状況に応じて判断を行い、適宜利用するシーンを選択する。例えば、サッカーの例では、ディレクタは、試合状況を見ながら選手およびボールがよく写るカメラの映像を選択する。

[0016] 放送ストリーム作成部１０３は、放送映像編集部１０２によって編集された映像および音声のコンテンツを、放送波に流すためのフォーマットである

放送ストリーム１０４に変換する。放送ストリーム作成部１０３は、例えば、映像であればＭＰＥＧ－２やＭＰＥＧ－４　ＡＶＣなどのビデオコーデックで符号化してビデオストリームを生成し、音声であればＡＣ３やＡＡＣなどのオーディオコーデックで符号化してオーディオストリームを生成し、それらを１本のＭＰＥＧ－２　ＴＳなどのシステムストリームに多重化を行う。

[0017] 再生装置１１０は、チューナ１１１および放送ストリームデコード部１１２から構成される。

[0018] チューナ１１１は、システムストリームを受信し、受信した信号を復調する機能を有している。

[0019] 放送ストリームデコード部１１２は、システムストリームをデコードする。放送ストリームデコード部１１２は、システムストリーム内の圧縮符号化されたビデオストリームをデコードすることにより非圧縮のイメージ映像を生成して、ビデオプレーンに出力しテレビ等に出力する。放送ストリームデコード部１１２はまた、システムストリーム内に圧縮符号化されたオーディオストリームをデコードし、非圧縮のＬＰＣＭ（Ｌｉｎｅａｒ　Ｐｕｌｓｅ　Ｃｏｄｅ　Ｍｏｄｕｌａｔｉｏｎ）状態のオーディオフレームを生成して、テレビなどのスピーカに出力する。

[0020] 以上が従来から広く普及する放送波による映像コンテンツの配信・視聴システム１０の構成である。

[0021] しかしながら、図１に示す放送局による映像コンテンツの配信・視聴システム１０では、ユーザは放送局によって作られた映像コンテンツを視聴することができるが、ユーザの意図を反映した編集がなされた映像コンテンツを楽しむことができない。つまり、映像コンテンツの内容は、放送映像撮影部１０１を使って撮影するカメラマン、および、放送映像編集部１０２を使って複数のシーンから映像を選択するディレクタの意思によって決まり、ユーザの好みが反映されることはない。

[0022] 例えば、サッカーの中継映像の場合、ユーザがお気に入りの選手がクロー

ズアップされた映像を視聴したいと考えても、放送局による編集は、大衆に広く受け入れられるように行われているため、特定の選手が偏って撮影されている映像が放送されることは少ない。例えば、アイドルのコンサートの中継映像の場合、お気に入りのアイドルがクローズアップされた映像を視聴したいとユーザが考えていても、放送局による編集は、大衆に広く受け入れられるように行われているため、特定のアイドルが偏って撮影されている映像が放送されることは少ない。

[0023] また、特許文献１および特許文献２に示したようなインターネットによる動画配信サービスにおいてもこの課題は解決されない。ユーザは複数ある映像コンテンツから自分の好みのコンテンツを選択してオンデマンドで再生することは可能であるが、視聴する映像コンテンツを動的に自分の嗜好を反映させた映像コンテンツに変えて視聴することはできない。

[0024] このような問題を解決するために、本発明の一態様に係る映像提供方法は、コンピュータが映像をユーザに提供するための映像提供方法であって、（ｉ）撮影空間のうちの一部の第一撮影空間が撮影された第一主映像と、（ｉｉ）前記撮影空間のうちの一部の空間であって、前記第一空間以外の空間を含む第二撮影空間が撮影された第二主映像と、を取得する映像取得ステップと、前記映像取得ステップにおいて取得された前記第一主映像および前記第二主映像を合成することにより広角映像を生成する映像生成ステップと、ネットワークを経由して、前記ユーザの嗜好情報を取得する情報取得ステップと、前記情報取得ステップにおいて取得された前記ユーザの嗜好情報に基づいて、前記広角映像のうちの一部の領域であって、前記広角映像の領域よりも小さいクロッピング領域を算出する領域算出ステップと、前記映像生成ステップにおいて生成された前記広角映像を、前記領域算出ステップにおいて算出された前記クロッピング領域でクロッピングするクロッピングステップと、前記クロッピングステップにおいてクロッピングされることにより生成されたクロッピング映像を前記ユーザに提供する映像提供ステップと、を含む。

[0025] これによれば、ユーザの嗜好情報に基づいて複数の映像を合成した広角映像のうちのクロッピング領域を特定し、特定したクロッピング領域でクロッピングすることにより生成したクロッピング映像をユーザに提供するため、ユーザの嗜好情報に応じた映像をユーザに対して提供することができる。

[0026] また、例えば、前記ユーザの嗜好情報は、ユーザが視聴したい対象である視聴対象を示し、前記映像提供方法は、さらに、前記ユーザの嗜好情報に基づいて前記広角映像に対して画像認識を行うことで、前記広角映像における前記視聴対象の位置を特定する位置特定ステップを含み、前記領域算出ステップでは、前記広角映像のうち、前記位置特定ステップにおいて特定された前記視聴対象の位置を用いて、前記視聴対象が含まれる領域を、前記クロッピング領域として算出してもよい。

[0027] これによれば、ユーザの嗜好情報に基づいて特定されるユーザが視聴したい対象である視聴対象について、広角映像に対して画像認識を行うことで、広角映像のうちの視聴対象が写り込んでいる領域をクロッピング領域として特定できるため、ユーザが視聴したい対象が写り込んでいる映像をユーザに対して提供することができる。

[0028] また、例えば、前記領域算出ステップでは、前記広角映像のうち、前記視聴対象の位置を、前記広角映像をクロッピングするための予め定められたサイズのクロッピング枠における所定の基準位置に一致させた場合に前記クロッピング枠で特定される領域を前記クロッピング領域として算出してもよい。

[0029] これによれば、クロッピングするためのクロッピング枠の基準位置に視聴対象の位置が一致するようにクロッピング領域を特定するため、確実に視聴対象を含む映像をクロッピング映像とすることができる。

[0030] また、例えば、前記領域算出ステップでは、前記広角映像のうち、処理対象のフレームよりも所定時間以前のフレームでの前記視聴対象の位置が、前記クロッピング枠の前記所定の基準位置に一致させた場合に前記クロッピング枠で特定される領域を前記クロッピング領域として算出してもよい。

[0031] 視聴対象の移動に対してクロッピング領域の移動を遅らせることができるため、人間がカメラを操作したかのような映像を生成できる。このため、ユーザに違和感を与えない自然な印象の映像を提供できる。

[0032] また、例えば、前記映像取得ステップでは、さらに、前記第一主映像および前記第二主映像と同じタイミングで、前記撮影空間のうちの少なくとも一部の空間が前記第一主映像及び前記第二主映像とは異なる角度で撮影された副映像を取得し、前記映像提供方法は、さらに、前記クロッピングステップでクロッピングされた前記クロッピング映像と、前記映像取得ステップで取得された前記副映像とのそれぞれを、所定のアルゴリズムに基づいて複数のシーンに分割するシーン分割ステップと、前記情報取得ステップにおいて取得された前記ユーザの嗜好情報に基づいて、前記複数のシーンのそれぞれについて、前記クロッピング映像および前記副映像のいずれかを選択する映像選択ステップと、を含み、前記映像提供ステップでは、前記映像選択ステップにおいて選択された前記クロッピング映像および前記副映像のいずれかをユーザに提供してもよい。

[0033] これによれば、複数の映像を複数のシーンに分割し、複数のシーンのそれぞれについて、ユーザの嗜好情報に応じて最適な映像を選択できるため、よりユーザに適した映像を提供することができる。

[0034] また、例えば、前記シーン分割ステップでは、前記クロッピング映像および前記副映像のそれぞれを前記複数のシーンに分割するときに、前記所定のアルゴリズムとは別に、所定時間毎に分割してもよい。

[0035] これによれば、所定のアルゴリズムとは別に所定時間毎に分割するため、所定のアルゴリズムによるシーンの分割で一つのシーンが長くなってしまっても、当該一つのシーンをさらに所定時間毎に分割できる。これにより、映像提供方法に係る処理単位を小さくできるため、複数の映像に対してほぼリアルタイムに処理を行うことができる。

[0036] また、例えば、前記所定のアルゴリズムは、前記撮影空間内で行われているイベントの種類毎に異なってもよい。

[0037] これによれば、イベントの種類毎に所定のアルゴリズムは異なるため、イベントの種類に適したシーンの分割を行うことができる。

[0038] また、例えば、前記シーン分割ステップでは、前記撮影空間内で行われているイベントの種類がスポーツである場合には、当該イベントの状態が「ゲーム中」であるか、「非ゲーム中」であるかを前記所定のアルゴリズムで判定することにより、判定結果が前記「ゲーム中」および前記「非ゲーム中」の一方から他方へ切り替わったタイミングで、前記クロッピング映像と前記副映像とのそれぞれを、複数のシーンに分割してもよい。

[0039] イベントの種類がスポーツである場合には、イベントの状態が「ゲーム中」であるか「非ゲーム中」であるかに応じてシーンを分割するため、適切にシーンを分割できる。

[0040] また、例えば、前記映像選択ステップでは、前記撮影空間内で行われているイベントの種類がスポーツである場合、前記「ゲーム中」から前記「非ゲーム中」に切り替わったときに、当該「非ゲーム中」のシーンを選択する代わりに直前の「ゲーム中」のシーンの映像の中から選択してもよい。

[0041] 「ゲーム中」から「非ゲーム中」に切り替わったときに、直前の「ゲーム中」の映像を選択するため、ユーザにリプレイ映像を提供することができる。

[0042] また、例えば、前記シーン分割ステップでは、前記撮影空間内で行われているイベントの種類がコンサートである場合には、当該イベントの状態が「演奏中」であるか、「非演奏中」であるかを前記所定のアルゴリズムで判定することにより、前記クロッピング映像と前記副映像とのそれぞれを、複数のシーンに分割してもよい。

[0043] イベントの種類がコンサートである場合には、イベントの状態が「演奏中」であるか「非演奏中」であるかに応じてシーンを分割するため、適切にシーンを分割できる。

[0044] また、例えば、前記シーン分割ステップでは、前記撮影空間内で行われているイベントの種類が討論会である場合には、当該討論会に参加している複

数の参加者のうちで当該討論会における話者の交代を前記所定のアルゴリズムで判定することにより、前記クロッピング映像と前記副映像とのそれぞれを、複数のシーンに分割してもよい。

[0045] イベントの種類が討論会である場合には、話者が交代したタイミングでシーンを分割するため、適切にシーンを分割できる。

[0046] また、例えば、前記映像提供方法は、さらに、前記シーン分割ステップにおいて分割された前記複数のシーンのそれぞれを、前記情報取得ステップにおいて取得された前記ユーザの嗜好情報と、所定の評価指標とに基づいて評価する評価ステップを含み、前記映像選択ステップでは、前記評価ステップにおいて評価された結果に基づいて、前記複数のシーンのそれぞれについて、前記クロッピング映像および前記副映像のいずれかを選択してもよい。

[0047] 複数のシーンのそれぞれについて評価した結果に応じて、提供する映像を選択するため、よりユーザの嗜好に適した映像を提供できる。

[0048] また、例えば、前記所定の評価指標は、映像を撮影した複数のカメラのうちで、前記視聴対象が画角に含まれており、かつ、前記視聴対象への距離が近いカメラによって撮影された映像のシーンほど高く評価する指標を含んでもよい。

[0049] 視聴対象に最も近いカメラによって撮影された映像のシーンほど高く評価するため、ユーザの嗜好に適した映像を選択できる。

[0050] また、例えば、前記所定の評価指標は、映像を撮影した複数のカメラのうちで、前記視聴対象が画角に含まれており、かつ、前記視聴対象との間にあるオブジェクトの数が少ないカメラによって撮影された映像のシーンほど高く評価する指標を含んでもよい。

[0051] 視聴対象までの間にオブジェクトが少ない映像のシーンほど高く評価するため、ユーザの嗜好に適した映像を選択できる。

[0052] また、例えば、前記所定の評価指標は、映像を撮影した複数のカメラのうちで、前記視聴対象が画角に含まれており、かつ、当該映像に映り込んでいる前記視聴対象の面積が大きいカメラによって撮影された映像のシーンほど

高く評価する指標を含んでもよい。

[0053] 視聴対象が最も大きく写り込んでいる映像のシーンほど高く評価するため、ユーザの嗜好に適した映像を選択できる。

[0054] また、例えば、前記所定の評価指標は、映像を撮影した複数のカメラのうちで、前記視聴対象が画角に含まれており、かつ、前記視聴対象への距離が近いカメラによって撮影された映像のシーンほど高く評価する第一指標と、映像を撮影した複数のカメラのうちで、前記視聴対象が画角に含まれており、かつ、前記視聴対象との間にあるオブジェクトの数が少ないカメラによって撮影された映像のシーンほど高く評価する第二指標と、映像を撮影した複数のカメラのうちで、前記視聴対象が画角に含まれており、かつ、当該映像に映り込んでいる前記視聴対象の面積が大きいカメラによって撮影された映像のシーンほど高く評価する第三指標と、のうちの２以上の指標を含み、前記評価ステップでは、前記複数のシーンのそれぞれについて、当該シーンに対する前記２以上の指標により評価された複数の結果について、前記２以上に指標に関連付けられている予め定められた重み付けによって重み付加算された加算値に基づいて評価してもよい。

[0055] 複数の評価指標を組み合わせたもので、映像のシーンを評価できるため、多角的にユーザの嗜好に適した映像を選択できる。

[0056] また、例えば、前記情報取得ステップでは、前記コンピュータに前記ネットワークを介して接続される情報端末に対して、前記ユーザにより入力された前記ユーザの嗜好情報を、前記ネットワークを経由して取得してもよい。

[0057] ユーザは手もとの情報端末を操作すれば、嗜好情報が反映された映像を取得できるため、容易にユーザの嗜好に適した映像を閲覧することができる。

[0058] なお、これらの全般的または具体的な態様は、システム、送信装置、受信装置、集積回路、コンピュータプログラムまたはコンピュータ読み取り可能なＣＤ－ＲＯＭなどの記録媒体で実現されてもよく、システム、送信装置、受信装置、集積回路、コンピュータプログラムまたは記録媒体の任意な組み合わせで実現されてもよい。

[0059] 以下、本発明の一態様に係る映像提供方法、送信装置および受信装置について、図面を参照しながら具体的に説明する。

[0060] なお、以下で説明する実施の形態は、いずれも本発明の一具体例を示すものである。以下の実施の形態で示される数値、形状、材料、構成要素、構成要素の配置位置及び接続形態、ステップ、ステップの順序などは、一例であり、本発明を限定する主旨ではない。また、以下の実施の形態における構成要素のうち、最上位概念を示す独立請求項に記載されていない構成要素については、任意の構成要素として説明される。

[0061] （実施の形態１）

本実施の形態に係る映像コンテンツの作成および伝送し再生するための、映像提供方法、送信装置および受信装置について説明する。

[0062] まず、本実施の形態における受信装置の使用形態について図２を用いながら説明する。

[0063] ユーザは受信装置を利用して、通信Ｉ／Ｆ等で受信した映像コンテンツを再生する。ここでは、受信装置の例として図２に示すようにデジタルテレビ２０２を挙げて説明する。

[0064] 図２の（ａ）に示すように、デジタルテレビ２０２は、ユーザインターフェースとしてリモコン２０１が付属しており、ユーザはリモコン２０１に対して入力を行うことにより、デジタルテレビ２０２への操作を行う。デジタルテレビ２０２は、ユーザの嗜好を反映するためのメニュー画面を表示する。図２の例では、デジタルテレビ２０２は、サッカーに対して、何を中心にフォーカスする映像が好みなのかをユーザに選択させる画面を表示する。ユーザが、例えば「ボール」を中心に見たい場合には、メニュー画面の「ボール」ボタンを選択すれば、図２の（ｂ）のように、ボールを中心にフォーカスされた映像が表示され、ユーザはボールを中心とした映像を視聴できる。ユーザが、例えば「選手Ａ」を中心に見たい場合には、「選手Ａ」ボタンを選択すれば、図２の（ｃ）のように、選手Ａを中心にフォーカスされた映像が表示され、ユーザは選手Ａを中心とした映像を視聴できる。このように、

本実施の形態における再生装置の使用形態においては、ユーザの好みに応じた映像コンテンツを視聴することが可能となる。

[0065] 以上が、受信装置の使用形態についての説明である。

[0066] 次に、デジタルテレビの放送波や通信等で伝送される一般的なストリームの構造について説明する。

[0067] デジタルテレビの放送波等での伝送では、ＭＰＥＧ－２トランスポートストリーム形式のデジタルストリームが使われている。ＭＰＥＧ－２トランスポートストリームとは、ビデオやオーディオなど様々なストリームを多重化して伝送するための規格である。ＩＳＯ／ＩＥＣ１３８１８－１およびＩＴＵ－Ｔ勧告Ｈ２２２．０において標準化されている。

[0068] 図３は、ＭＰＥＧ－２トランスポートストリーム形式のデジタルストリームの構成を示す図である。本図に示すようにトランスポートストリームは、ビデオストリーム、オーディオストリーム、字幕ストリームなどを多重化することで得られる。ビデオストリームは番組の主映像を、オーディオストリームは番組の主音声部分や副音声を、字幕ストリームは番組の字幕情報をそれぞれ格納している。ビデオストリームは、ＭＰＥＧ－２、ＭＰＥＧ－４ＡＶＣなどの方式を使って符号化記録される。オーディオストリームは、ドルビーＡＣ－３、ＭＰＥＧ－２　ＡＡＣ、ＭＰＥＧ－４　ＡＡＣ、ＨＥ－ＡＡＣなどの方式で圧縮・符号化記録されている。

[0069] ビデオストリームの構成について説明する。ＭＰＥＧ－２、ＭＰＥＧ－４ＡＶＣ、ＳＭＰＴＥ　ＶＣ－１などの動画圧縮符号化においては、動画像の空間方向および時間方向の冗長性を利用してデータ量の圧縮を行う。時間方向の冗長性を利用する方法として、ピクチャ間予測符号化が用いられる。ピクチャ間予測符号化では、あるピクチャを符号化する際に、表示時間順で前方または後方にあるピクチャを参照ピクチャとする。そして、その参照ピクチャからの動き量を検出し、動き補償を行ったピクチャと符号化対照のピクチャとの差分値に対して空間方向の冗長度を取り除くことによりデータ量の圧縮を行う。図１１に一般的なビデオストリームのピクチャの参照構造を

示す。矢印は参照して圧縮されていることを示す。

[0070] ここでは、参照ピクチャを持たずに符号化対象ピクチャのみを用いてピクチャ内予測符号化を行うピクチャをＩピクチャと呼ぶ。ピクチャとは、フレームおよびフィールドの両者を包含する１つの符号化の単位である。また、既に処理済の１枚のピクチャを参照してピクチャ間予測符号化するピクチャをＰピクチャと呼び、既に処理済みの２枚のピクチャを同時に参照してピクチャ間予測符号化するピクチャをＢピクチャと呼び、Ｂピクチャの中で他のピクチャから参照されるピクチャをＢｒピクチャと呼ぶ。また、フレーム構造の場合のフレーム、および、フィールド構造のフィールドを、ここではビデオアクセスユニットと呼ぶ。

[0071] また、ビデオストリームは、図４に示すような階層構造を有している。ビデオストリームは、複数のＧＯＰ（Ｇｒｏｕｐ　ｏｆ　Ｐｉｃｔｕｒｅｓ）から構成されており、これを符合化処理の基本単位とすることで動画像の編集やランダムアクセスが可能となっている。ＧＯＰは１つ以上のビデオアクセスユニットにより構成されている。ビデオアクセスユニットは、ピクチャの符合化データを格納する単位であり、フレーム構造の場合には１フレームが格納され、フィールド構造の場合には１フィールドのデータが格納される。各ビデオアクセスユニットは、ＡＵ識別コード、シーケンスヘッダ、ピクチャヘッダ、補足データ、圧縮ピクチャデータ、パディングデータ、シーケンス終端コード、ストリーム終端コードなどから構成される。各データはＭＰＥＧ－４　ＡＶＣの場合には、ＮＡＬユニットと呼ばれる単位で格納される。

[0072] ＡＵ識別コードはアクセスユニットの先頭を示す開始符号である。シーケンスヘッダは、複数ビデオアクセスユニットから構成される再生シーケンスでの共通の情報を格納したヘッダであり、解像度、フレームレート、アスペクト比、ビットレートなどの情報が格納される。ピクチャヘッダはピクチャ全体の符合化の方式などの情報を格納したヘッダである。補足データは圧縮データの復号に必須ではない付加情報であり、例えば、映像と同期してＴＶ

に表示するクローズドキャプションの文字情報やＧＯＰ構造情報などが格納される。圧縮ピクチャデータには、圧縮符号化されたピクチャのデータが格納される。パディングデータは、形式を整えるための意味のないデータが格納される。例えば、決められたビットレートを保つためのスタッフィングデータとして用いる。シーケンス終端コードは、再生シーケンスの終端を示すデータである。ストリーム終端コードは、ビットストリームの終端を示すデータである。

[0073] ＡＵ識別コード、シーケンスヘッダ、ピクチャヘッダ、補足データ、圧縮ピクチャデータ、パディングデータ、シーケンス終端コード、およびストリーム終端コードの中身の構成は、ビデオの符合化方式によって異なる。

[0074] 例えば、ＭＰＥＧ－４ ＡＶＣの場合であれば、ＡＵ識別コードは、ＡＵデリミタ（Ａｃｃｅｓｓ Ｕｎｉｔ Ｄｅｌｉｍｉｔｅｒ）に、シーケンスヘッダはＳＰＳ（Ｓｅｑｕｅｎｃｅ Ｐａｒａｍｔｅｒ Ｓｅｔ）に、ピクチャヘッダはＰＰＳ（Ｐｉｃｔｕｒｅ Ｐａｒａｍｅｔｅｒ Ｓｅｔ）に、圧縮ピクチャデータは複数個のスライスに、補足データはＳＥＩ（Ｓｕｐｐｌｅｍｅｎｔａｌ Ｅｎｈａｎｃｅｍｅｎｔ Ｉｎｆｏｒｍａｔｉｏｎ）に、パディングデータはＦｉｌｌｅｒＤａｔａに、シーケンス終端コードはＥｎｄ ｏｆ Ｓｅｑｕｅｎｃｅに、ストリーム終端コードはＥｎｄ ｏｆ Ｓｔｒｅａｍにそれぞれ対応する。

[0075] また、例えば、ＭＰＥＧ－２の場合であれば、シーケンスヘッダはｓｅｑｕｅｎｃｅ＿Ｈｅａｄｅｒ、ｓｅｑｕｅｎｃｅ＿ｅｘｔｅｎｓｉｏｎ、ｇｒｏｕｐ＿ｏｆ＿ｐｉｃｔｕｒｅ＿ｈｅａｄｅｒに、ピクチャヘッダはｐｉｃｔｕｒｅ＿ｈｅａｄｅｒ、ｐｉｃｔｕｒｅ＿ｃｏｄｉｎｇ＿ｅｘｔｅｎｓｉｏｎに、圧縮ピクチャデータは複数個のスライスに、補足データはｕｓｅｒ＿ｄａｔａに、シーケンス終端コードはｓｅｑｕｅｎｃｅ＿ｅｎｄ＿ｃｏｄｅにそれぞれ対応する。なお、この場合にＡＵ識別コードは存在しないが、それぞれのヘッダのスタートコードを使えば、アクセスユニットの切れ目を判断できる。

[0076] 各データは常に必要ではなく、例えば、シーケンスヘッダはＧＯＰ先頭のビデオアクセスユニットでのみ必要で、それ以外のビデオアクセスユニットにはなくてもよい、としてもよい。また、符号化方式によっては、ピクチャヘッダは符号順で前のビデオアクセスユニットのものを参照して、自身のビデオアクセスユニット内にピクチャヘッダがなくてもよい。

[0077] また、図５に示すようにＧＯＰ先頭のビデオアクセスユニットは、圧縮ピクチャデータとしてＩピクチャのデータが格納され、ＡＵ識別コード、シーケンスヘッダ、ピクチャヘッダ、および圧縮ピクチャデータが必ず格納され、補足データ、パディングデータ、シーケンス終端コード、およびストリーム終端コードが格納されている。ＧＯＰ先頭以外のビデオアクセスユニットは、ＡＵ識別コードおよび圧縮ピクチャデータが必ず格納され、補足データ、パディングデータ、シーケンス終端コード、およびストリーム終端コードが格納されている。

[0078] 次に、クロッピング領域情報およびスケーリング情報について図６および図７を参照しながら説明する。

[0079] ビデオ符号化方式によっては、符号化されたフレームの領域と、実際に表示に使う領域を変更することができる。図６に示すように、符号化されたフレーム領域の中から実際に表示する領域を、「クロッピング領域」として指定することができる。例えば、ＭＰＥＧ－４　ＡＶＣの場合には、ＳＰＳに格納されるｆｒａｍｅ＿ｃｒｏｐｐｉｎｇ情報を使って指定できる。ｆｒａｍｅ＿ｃｒｏｐｐｉｎｇ情報は、図７の（ａ）のように、クロッピンング領域の上線／下線／左線／右線と、符号化されたフレーム領域の上線／下線／左線／右線との差分を、上下左右のクロップ量として指定する。より具体的には、クロッピング領域を指定する場合には、ｆｒａｍｅ＿ｃｒｏｐｐｉｎｇ＿ｆｌａｇを１に設定し、ｆｒａｍｅ＿ｃｒｏｐ＿ｔｏｐ＿ｏｆｆｓｅｔ／ｆｒａｍｅ＿ｃｒｏｐ＿ｂｏｔｔｏｍ＿ｏｆｆｓｅｔ／ｆｒａｍｅ＿ｃｒｏｐ＿ｌｅｆｔ＿ｏｆｆｓｅｔ／ｆｒａｍｅ＿ｃｒｏｐ＿ｒｉｇｈｔ＿ｏｆｆｓｅｔに上／下／左／右のクロップ量を指定する。ＭＰＥＧ

ー２の場合には、図７の（ｂ）のように、クロッピング領域の縦横のサイズ（ｓｅｑｕｅｎｃｅ＿ｄｉｓｐｌａｙ＿ｅｘｔｅｎｓｉｏｎのｄｉｓｐｌａｙ＿ｈｏｒｉｚｏｎｔａｌ＿ｓｉｚｅ，ｄｉｓｐｌａｙ＿ｖｅｒｔｉｃａｌ＿ｓｉｚｅ）と、符号化されたフレーム領域の中心とクロッピング領域の中心との差分情報（ｐｉｃｔｕｒｅ＿ｄｉｓｐｌａｙ＿ｅｘｔｅｎｓｉｏｎのｆｒａｍｅ＿ｃｅｎｔｒｅ＿ｈｏｒｉｚｏｎｔａｌ＿ｏｆｆｓｅｔ，ｆｒａｍｅ＿ｃｅｎｔｒｅ＿ｖｅｒｔｉｃａｌ＿ｏｆｆｓｅｔ）を使ってクロッピング領域を指定できる。

[0080] また、ビデオ符号化方式によっては、クロッピング領域を実際にテレビなどに表示する際のスケーリング方法を示すスケーリング情報が存在する。これは、例えばアスペクト比として設定される。再生装置はそのアスペクト比の情報を使って、クロッピング領域をアップコンバートして表示を行う。例えば、ＭＰＥＧ－４　ＡＶＣの場合には、スケーリング情報として、ＳＰＳにアスペクト比の情報（ａｓｐｅｃｔ＿ｒａｔｉｏ＿ｉｄｃ）が格納される。ＭＰＥＧ－４　ＡＶＣの場合、１４４０ｘ１０８０のクロッピング領域を、１９２０ｘ１０８０に拡大して表示するためには、アスペクト比は４：３を指定する。この場合水平方向に４／３倍にアップコンバート（１４４０ｘ４／３＝１９２０）され、１９２０ｘ１０８０に拡大されて表示される。ＭＰＥＧ－２の場合にも同様にｓｅｑｕｅｎｃｅ＿ｈｅａｄｅｒにアスペクト比の情報（ａｓｐｅｃｔ＿ｒａｔｉｏ＿ｉｎｆｏｒｍａｔｉｏｎ）が格納されている。

[0081] 以上がビデオストリームの構成の説明である。

[0082] トランスポートストリームに含まれる各ストリームはＰＩＤと呼ばれるストリーム識別ＩＤによって識別される。このＰＩＤのパケットを抽出することで復号装置は、対象のストリームを抽出することができる。ＰＩＤとストリームの対応は以降で説明するＰＭＴパケットのディスクリプタに格納される。

[0083] 図３は、トランスポートストリームがどのように多重化されるかを模式的

に示している。まず、複数のビデオフレームからなるビデオストリーム５０１、複数のオーディオフレームからなるオーディオストリーム５０４を、それぞれＰＥＳパケット列５０２および５０５に変換し、ＴＳパケット５０３および５０６に変換する。同じく字幕ストリーム５０７のデータをそれぞれＰＥＳパケット列５０８に変換し、更にＴＳパケット５０９に変換する。ＭＰＥＧ－２トランスポートストリーム５１３はこれらのＴＳパケットを１本のストリームに多重化することで構成される。

[0084] 図８は、ＰＥＳパケット列に、ビデオストリームがどのように格納されるかを更に詳しく示している。本図における第１段目はビデオストリームのビデオフレーム列を示す。第２段目は、ＰＥＳパケット列を示す。本図の矢印ｙｙ１，ｙｙ２，ｙｙ３，ｙｙ４に示すように、ビデオストリームにおける複数のＶｉｄｅｏ　Ｐｒｅｓｅｎｔａｔｉｏｎ　ＵｎｉｔであるＩピクチャ、Ｂピクチャ、Ｐピクチャは、ピクチャ毎に分割され、ＰＥＳパケットのペイロードに格納される。各ＰＥＳパケットはＰＥＳヘッダを持ち、ＰＥＳヘッダには、ピクチャの表示時刻であるＰＴＳ（Ｐｒｅｓｅｎｔａｔｉｏｎ　Ｔｉｍｅ－Ｓｔａｍｐ）やピクチャの復号時刻であるＤＴＳ（Ｄｅｃｏｄｉｎｇ　Ｔｉｍｅ－Ｓｔａｍｐ）が格納される。

[0085] 図９は、トランスポートストリームを構成するＴＳパケットのデータ構造を示す図である。ＴＳパケットは、４ＢｙｔｅのＴＳヘッダと、アダプテーションフィールドとＴＳペイロードから構成される１８８Ｂｙｔｅ固定長のパケットである。ＴＳヘッダは、ｔｒａｎｓｐｏｒｔ＿ｐｒｉｏｒｉｔｙ、ＰＩＤ、ａｄａｐｔａｉｏｎ＿ｆｉｅｌｄ＿ｃｏｎｔｒｏｌなどから構成される。ＰＩＤは前述したとおりトランスポートストリームに多重化されているストリームを識別するためのＩＤである。ｔｒａｎｓｐｏｒｔ＿ｐｒｉｏｒｉｔｙは、同一ＰＩＤのＴＳパケットの中のパケットの種別を識別するための情報である。ａｄａｐｔａｔｉｏｎ＿ｆｉｅｌｄ＿ｃｏｎｔｒｏｌは、アダプテーションフィールドとＴＳペイロードの構成を制御するための情報である。アダプテーションフィールドとＴＳペイロードはどちらかだけが存

在する場合と両方が存在する場合があり、ａｄａｐｔａｔｉｏｎ＿ｆｉｅｌｄ＿ｃｏｎｔｒｏｌはその有無を示す。ａｄａｐｔａｔｉｏｎ＿ｆｉｅｌｄ＿ｃｏｎｔｒｏｌが１の場合は、ＴＳペイロードのみが存在し、ａｄａｐｔａｔｉｏｎ＿ｆｉｅｌｄ＿ｃｏｎｔｒｏｌが２の場合は、アダプテーションフィールドのみが存在し、ａｄａｐｔａｔｉｏｎ＿ｆｉｅｌｄ＿ｃｏｎｔｒｏｌが３の場合は、ＴＳペイロードとアダプテーションフィールドの両方が存在することを示す。

[0086] アダプテーションフィールドは、ＰＣＲなどの情報の格納や、ＴＳパケットを１８８バイト固定長にするためのスタッフィングするデータの格納領域である。ＴＳペイロードにはＰＥＳパケットが分割されて格納される。

[0087] トランスポートストリームに含まれるＴＳパケットには、映像・音声・字幕などの各ストリーム以外にもＰＡＴ（Ｐｒｏｇｒａｍ Ａｓｓｏｃｉａｔｉｏｎ Ｔａｂｌｅ）、ＰＭＴ（Ｐｒｏｇｒａｍ Ｍａｐ Ｔａｂｌｅ）、ＰＣＲ（Ｐｒｏｇｒａｍ Ｃｌｏｃｋ Ｒｅｆｅｒｅｎｃｅ）などがある。これらのパケットはＰＳＩ（Ｐｒｏｇｒａｍ Ｓｐｅｃｉｆｉｃ Ｉｎｆｏｒｍａｔｉｏｎ）と呼ばれる。ＰＡＴはトランスポートストリーム中に利用されるＰＭＴのＰＩＤが何であるかを示し、ＰＡＴ自身のＰＩＤは０で登録される。ＰＭＴは、トランスポートストリーム中に含まれる映像・音声・字幕などの各ストリームのＰＩＤと各ＰＩＤに対応するストリームの属性情報を持ち、またトランスポートストリームに関する各種ディスクリプタを持つ。ディスクリプタにはＡＶストリームのコピーを許可・不許可を指示するコピーコントロール情報などがある。ＰＣＲは、ＴＳパケットのデコーダへの到着時刻とＰＴＳ・ＤＴＳの時間軸であるＳＴＣ（Ｓｙｓｔｅｍ Ｔｉｍｅ Ｃｌｏｃｋ）の同期を取るために、そのＰＣＲパケットがデコーダに転送されるタイミングに対応するＳＴＣ時間の情報を持つ。

[0088] 図１０はＰＭＴのデータ構造を詳しく説明するための図である。ＰＭＴの先頭には、そのＰＭＴに含まれるデータの長さなどを記したＰＭＴヘッダが配置される。その後ろには、トランスポートストリームに関するディスクリ

プタが複数配置される。前述したコピーコントロール情報などが、ディスクリプタとして記載される。ディスクリプタの後には、トランスポートストリームに含まれる各ストリームに関するストリーム情報が複数配置される。ストリーム情報は、ストリームの圧縮コーデックなどを識別するためストリームタイプ、ストリームのＰＩＤ、ストリームの属性情報（フレームレート、アスペクト比など）が記載されたストリームディスクリプタから構成される。

[0089] ここで、図９に示すトランスポートストリームは、ＴＳパケットが並ぶストリームであり、一般的に放送波に使われるストリームはこの形式である。図９に示すトランスポートストリームは以降ではＴＳストリームと呼ぶことにする。一方で、図１２に示すトランスポートストリームは、１８８バイトのＴＳパケットの先頭に４Ｂｙｔｅのタイムスタンプを付与したソースパケットが並ぶストリームであり、一般的に通信で伝送されるストリームはこの形式である。図１２に示すトランスポートストリームは、以降ではＴＴＳストリームと呼ぶことにする。ＴＳパケットに付与される先頭のタイムスタンプは、以降ではＡＴＳ（Ａｒｒｉｖａｌ＿ｔｉｍｅ＿ｓｔａｍｐ）と呼び、ＡＴＳは付与されるＴＳパケットの、ストリームのデコーダへの転送開始時刻を示す。ＴＴＳストリームは図１２に示すようにソースパケットが並ぶこととなり、ＴＴＳストリームの先頭からインクリメントする番号はＳＰＮ（ソースパケットナンバー）と呼ぶ。

[0090] 通常の放送波では複数チャンネル文のＴＳが多重化されたフルＴＳで送出される。フルＴＳは１８８バイトの固定長ＴＳパケット列で構成されるＴＳストリームである。一方で、ＢＤ－ＲＥやＨＤＤ等の蓄積メディアに放送番組を記録する場合には、フルＴＳの中から、必要なチャンネルのデータのみを抜き出しパーシャルＴＳとして記録する。パーシャルＴＳは、ＴＴＳストリームである。ここでＴＳストリームをＴＴＳストリームに変換する場合に、フルＴＳから不要となったＴＳパケットを単純につめて記録してしまえば、ＴＳパケット間の時間間隔情報がなくなる。このため、デコーダへの入力

タイミングが、送出時に想定していたタイミングとずれることになり、デコーダが正しく再生できなくなってしまう。そこで、フルＴＳから不要となったＴＳパケットの時間間隔情報を残すために、ＡＴＳを付与する。このように構成し、デコーダへのデータの入力タイミングを、ＡＴＳで制御することによって、デコーダを破綻させることなく、再生することが可能となる。

[0091] ＴＳストリームからＴＴＳストリームへの変換は、図１３の仕組みを用いて行われる。図１３は、ＴＳストリームをＴＴＳストリームに変換する方式を示しており、その方式は、ＴＳパケットフィルタリング、ＡＴＳ付与器、ＡＴＣカウンタ、および高周波発信器から構成される。

[0092] 水晶振動子は、水晶（石英）の圧電効果を利用して高い周波数精度の発振を起こす装置であり、ここでは２７Ｍｈｚのクロックを発振するとする。

[0093] ＡＴＣカウンタは、水晶振動子のクロックにしたがって、ＡＴＣ時間軸を刻むカウンタである。ＡＴＣカウンタは、データバッファから入力されるＴＳパケットのＡＴＳで初期化し、２７Ｍｈｚの周波数で値をインクリメントする。

[0094] ＴＳパケットフィルタリングは、ＥＩＴの番組情報、およびＰＭＴパケットのプログラム内のストリーム構成情報を利用して、ユーザが選択する番組を構成するＴＳパケットのみをフィルタリングして、ＡＴＳ付与器に入力する。

[0095] ＡＴＳ付与器は、ＴＳパケットフィルタリングを経由して入力される１８８バイトのＴＳパケットに対して、ＡＴＣカウンタのＡＴＣ値を参照して、ＴＳパケットの先頭にＡＴＳ値を付与し、１９２バイトのＴＳパケットを生成する。ＡＴＳのフィールドは４バイトなので、０ｘ０から０ｘＦＦＦＦＦＦＦＦの値をとり、ＡＴＣ値が０ｘＦＦＦＦＦＦＦＦ以上の値になった場合には、再びＷｒａｐ－ａｒｏｕｎｄして０に戻る。なお、Ｂｌｕ－ｒａｙ（登録商標）の場合には、ＴＳパケットの先頭４Ｂｙｔｅの先頭２ｂｉｔはコピー制御情報に利用されるため、ＡＴＳ値は３０ｂｉｔであり、３０ｂｉｔでＷｒａｐ－ａｒｏｕｎｄする。

[0096] 以上がデジタルテレビの放送波や通信等で伝送される一般的なストリームの構造の説明である。

[0097] （個人嗜好を反映した映像コンテンツの自動生成・視聴システム）

次に本実施の形態に係る、個人嗜好を反映した映像コンテンツの自動生成・視聴システム（以下、「配信・視聴システム」という。）について、図面を参照しながら説明を行う。

[0098] 図１４は、配信・視聴システムの全体像を示す。配信・視聴システム１４００は、撮影システム１４１０、編集システム１４２０、再生システム１４３０から構成される。

[0099] （撮影システム）

撮影システム１４１０は、撮影制御部１４０１、複数の映像撮影部１４０２、通信Ｉ／Ｆ１４０３から構成される。撮影システム１４１０は、撮影制御部１４０１により制御された複数の映像撮影部１４０２を用いて、イベントを撮影し、撮影した映像を圧縮符号化し、圧縮符号化した映像を通信Ｉ／Ｆ１４０３を通じて編集システム１４２０に伝送する。

[0100] 映像撮影部１４０２は、主にビデオカメラを指し、撮影制御部１４０１の制御を元に、映像（音声を含む）を撮影し、圧縮符号化した映像データを通信Ｉ／Ｆ１４０３に伝送する。ここで、映像撮影部は１つまたは複数存在しており、図１５の（ａ）に示すようにイベント全体が広角で入るように配置される。図１５の（ａ）では、サッカーの試合の撮影例を示しており、コート全体が写るように、複数の映像撮影部である第一カメラ１５０１、第二カメラ１５０２、および第三カメラ１５０３が広角に配置される。つまり、第一カメラ１５０１はコート左側が写る向きで設置され、第二カメラ１５０２はコート真ん中が写る向きで設置され、第三カメラ１５０３はコート右側が写る向きで設置されている。図１５の（ｂ）は、各カメラ１５０１～１５０３によって撮影される映像を模式的に示している。第一主映像１５１１は第一カメラ１５０１によって撮影された映像、第二主映像１５１２は第二カメラ１５０２によって撮影された映像、第三主映像１５１３は第三カメラ１５

０３によって撮影された映像である。要するに、第一カメラ１５０１で撮影された映像は、撮影空間のうちの一部の第一撮影空間が撮影された第一主映像１５１１である。また、第二カメラ１５０２で撮影された映像は、撮影空間のうちの一部の空間であって、第一空間以外の空間を含む第二撮影空間が撮影された第二主映像１５１２である。第三カメラ１５０３で撮影された映像は、撮影空間のうちの第一空間および第二空間以外の空間を含む第三撮影空間が撮影された第三主映像１５１３である。このように映像撮影部１４０２は、１つまたは複数が、イベント全体が写るように向きや位置が固定されて配置される。なお、映像撮影部１４０２は、３台のカメラ１５０１～１５０３から構成されるが、複数台のカメラから構成されていればよく、少なくとも２台のカメラ１５０１、１５０２により構成されていればよい。

[0101] 撮影制御部１４０１は、複数の映像撮影部１４０２に対して、同期した撮影開始、撮影停止などの制御を行う。図１５の（ａ）では、撮影制御部１４０１は、タブレット型端末１５０４とである。タブレット型端末１５０４は、複数の映像撮影部１４０２である第一カメラ１５０１、第二カメラ１５０２および第三カメラ１５０３と無線や有線等で通信可能な通信部を有しており、タブレット型端末１５０４上で実行されるアプリケーションによって、第一カメラ１５０１、第二カメラ１５０２および第三カメラ１５０３の動作を制御できる。タブレット型端末１５０４は、具体的には、撮影開始、撮影停止などの指示を第一カメラ１５０１、第二カメラ１５０２および第三カメラ１５０３に対して行うことができる。また、タブレット型端末１５０４は、通信部を通じて無線や有線等で、同期信号を第一カメラ１５０１、第二カメラ１５０２および第三カメラ１５０３に送る。この同期信号が、第一カメラ１５０１、第二カメラ１５０２および第三カメラ１５０３で撮影して生成するストリームに埋め込まれることによって、後段の処理において、この同期信号を利用すれば、複数のストリーム間の同期を取ることが可能となる。つまり、あるストリームのフレームの時間と同じ別のストリームのフレームがどこであるかの判断が容易となる。同期信号は、例えば、ＮＴＰサーバに

よる信号情報であってもよい。なお、撮影制御部１４０１の機能は、映像撮影部１４０２のどれか１つが有していてもよい。なお、これらタブレット型端末１５０４上で表示される、第一カメラ１５０１、第二カメラ１５０２および第三カメラ１５０３を制御するためのＧＵＩは、ＨＴＭＬ５やＪａｖａ（登録商標）等のアプリケーションによって実現されていてもよい。

[0102] 通信Ｉ／Ｆ１４０３は、インターネットと接続するためのＩ／Ｆを示しており、例えばルータ等を示す。図１５の（ａ）に示すように、各カメラ１５０１～１５０３で撮影された映像ストリームは、通信Ｉ／Ｆ１４０３であるルータ等を通じて、インターネット上にある編集システム１４２０に伝送される。なお、通信Ｉ／Ｆ１４０３は、ネットワーク上に存在する編集システムに伝送するためのＩ／Ｆであればよく、例えば、携帯電話網（３ＧやＬＴＥ等）に接続するものでもよい。なお、映像撮影部１４０２の撮影映像を端末内部のローカルストレージ（メモリやＨＤＤ）に格納しておき、撮影後に、そのデータを、パソコン等の情報端末を使って、編集システムにアップロードしてもよい。

[0103] （編集システム）

編集システム１４２０は、位置特定部１４２２、映像生成部１４２３、自動映像編集部１４２４、情報取得部１４２５、映像提供部１４２６、および通信Ｉ／Ｆ１４２１、１４２７から構成される。編集システム１４２０は、撮影システム１４１０によって撮影されたイベントの映像ストリームから広角映像を生成すると共に、画像認識を行うことにより被写体の位置情報を特定し、その位置情報とユーザの嗜好情報とから、ユーザ最適な映像ストリームを生成する。なお、編集システム１４２０は、コンピュータにより構成され、ユーザの嗜好情報に基づいて編集された映像を提供する、送信装置として機能する。

[0104] 通信Ｉ／Ｆ１４２１は、映像取得部として機能し、第一カメラ１５０１により撮影された第一主映像、第二カメラ１５０２により撮影された第二主映像、および、第三カメラ１５０３により撮影された第三主映像を取得する。

[0105] 映像生成部１４２３は、撮影システム１４１０によって撮影された複数の映像ストリームから、広角映像（パノラマ映像）を生成する。つまり、映像生成部１４２３は、複数の映像ストリームである第一主映像１５１１、第二主映像１５１２および第三主映像１５１３を合成することにより広角映像を生成する。

[0106] 図１６は、広角映像の具体的な生成方法を模式的に示す図である。図１６の（ａ）は、撮影システム１４１０によって撮影された複数の映像であり、図１５の例で示した第一主映像１５１１、第二主映像１５１２、および第三主映像１５１３である。図１６の（ａ）に示すように、第一主映像１５１１および第二主映像１５１２は、互いに同じ空間を撮影している領域であるオーバラップ領域を含み、また、第二主映像１５１２および第三主映像１５１３は、オーバラップ領域を含む。映像生成部１４２３は、各映像に含まれるオーバラップ領域を重ね合わせて、図１６の（ｃ）に示すような１枚の広角映像を生成する。

[0107] 具体的には、映像生成部１４２３は、次の処理を行う。

[0108] まず、映像生成部１４２３は、（１）各映像に含まれるオーバラップ領域に対して、画像特徴点を抽出し、映像間の画像特徴点のマッチングを行う。ここで、画像特徴点の抽出には、たとえば、ＳＩＦＴやＳＵＲＦといったアルゴリズムが使用される。図１６の（ｂ）の第一主映像１５１１および第二主映像１５１２の例では、丸で囲んだ箇所が特徴点であり、当該特徴点の第一主映像１５１１および第二主映像１５１２間のマッチングを矢印で示している。

[0109] 次に、映像生成部１４２３は、（２）各映像１５１１～１５１３間の画像特徴点が一致するように画像を変形する。図１６の（ｂ）の例でいえば、画像特徴点から第一主映像１５１１は第二主映像１５１２よりも拡大された映像であることが分かるため、第一主映像１５１１を縮小もしくは第二主映像１５１２を拡大することで、第一主映像１５１１と第二主映像１５１２との接続をシームレスにすることが可能となる。画像特徴点からの画像変形を行

うためには、特徴点から形状変形用にホモグラフィ行列等の行列を生成し、画像に対して行列演算を行うことで変形できる。

[0110] 次に、映像生成部１４２３は、（３）変形された映像を１枚の広角映像に合成する。合成する場合には、各映像１５１１～１５１３に含まれるオーバラップ領域部分をブレンディングしてもよいし、どちらかのオーバラップ領域を削除するように構成してもよい。このような複数の映像から広角映像を生成する手段は、一般的に「スティッチング」と呼ばれ、広角映像の生成手段として広く利用されてきており、ＯｐｅｎＣＶ等の様々なソフトウェアで実装されている。なお、（１）のステップにおいて、特徴点マッチングではなく、複数のカメラ１５０１～１５０３のそれぞれの位置、向き情報や画角パラメータ等を用いることで、画像の歪みを特定し、特定した画像の歪みを利用して各映像１５１１～１５１３を合成してもよい。

[0111] なお、複数の映像１５１１～１５１３を用いた広角映像を生成するときには、映像生成部１４２３は、複数の映像１５１１～１５１３のうちの同一のタイミングで撮影された３つのフレームに対して上記の画像合成を行う。つまり、映像生成部１４２３は、第一主映像１５１１、第二主映像１５１２および第三主映像１５１３のそれぞれに埋め込まれた、第一主映像１５１１、第二主映像１５１２および第三主映像１５１３を同期するための同期信号に基づいて、同期を行いつつ、同一のタイミングで撮影された第一主映像１５１１、第二主映像１５１２、および第三主映像１５１３の各フレームに対して画像合成を行う。

[0112] 位置特定部１４２２は、映像生成部１４２３によって生成された広角映像に対して、コンテンツデータベースを参照しながら画像認識処理を行うことで、被写体の位置情報を解析し、かつ、特定する。なお、ここで「コンテンツデータベース」は、例えば、ボールの形状、グラウンドの形状、選手の名前、ポジション、背番号、顔写真といった情報を格納している。例えば、ボールの位置情報は、映像生成部１４２３により生成された広角映像に対して、ボールの形状や色とのパターンマッチングを行うことにより特定される。

また、例えば、選手の位置情報は、広角映像に対して、選手の顔やユニフォーム、背番号、体型などのパターンマッチングを行うことにより特定される。つまり、ユーザが視聴したい対象である視聴対象が分かれば、位置特定部１４２２は、当該視聴対象に基づいてコンテンツデータベースを参照しながら広角映像に対して画像認識を行うことで、広角映像における視聴対象の位置を特定する。

[0113] また、特定した選手およびボールの動きをトラッキングすれば、選手およびボールの位置情報を特定できる。ここで、選手、ボールなどの物体のトラッキング処理は、背景差分を行い、動きのある物体のみを抽出して、画像の動きを計測することで実現できる。画像処理による物体の追跡処理としては、オプティカルフローなどが有名であり、ＯｐｅｎＣＶ等の様々なソフトウェアで実装されている。また、選手が重なるなどしてトラッキングが外れてしまう場合には、トラッキングが外れる直前の選手の位置情報と、次に当該選手が検出される位置情報とで補間すればよい。

[0114] また、広角映像において、コートの領域を特定して、人物位置情報をコートの領域上の２次元座標の情報に変換してもよい。これは例えば図１７に示すように、広角映像上のコートの端点と、２次元座標上のコートの端点との対応関係から、ホモグラフィ行列等の変換行列を作成して、広角映像上の選手およびボール位置情報に行列演算を掛けることで、２次元座標に変換する。なお、撮影システム１４１０のカメラのそれぞれをステレオカメラとすれば、広角映像をステレオ画像で生成でき、かつ、奥行き情報を取ることができる。このため、奥行き情報を使うことで、選手やボールの位置情報をより精度高く取ることが可能となる。また、ステレオカメラの代わりに、撮影システム１４１０のカメラにデプスセンサーを搭載すれば、奥行き情報であるデプスマップを取得することが可能となるため、選手やボールの位置情報を精度高く取ることが可能となる。ここで、「デプスセンサー」とは、赤外線等のレーザをターゲットに照射して、往復するまでの時間を計測する方式（ＴＯＦ）などを利用して、ターゲットとの距離を各ピクセル単位で計測する

センサーである。デプスセンサーを利用したカメラは、例えばＭｉｃｒｏｓｏｆｔ社のＫｉｎｅｃｔなどが有名である。こうして生成されたデプスマップを使えば人物位置のみならず、骨格情報も取得できるため、３次元空間上に、撮影対象のイベントをＣＧ等で再現することも可能である。

[0115] 情報取得部１４２５は、通信Ｉ／Ｆ１４２７を介して、ユーザの嗜好情報を取得する。つまり、情報取得部１４２５は、ネットワークを経由して、ユーザの嗜好情報を取得する。ユーザの嗜好情報は、該当映像コンテンツのユーザの好みの見方を記す情報である。例えば、図２の例においては、ユーザの嗜好情報は、「ボール中心の映像」「選手Ａ中心の映像」「選手Ｂ中心の映像」の選択肢の中でユーザ選択した値である。つまり、ユーザの嗜好情報は、ユーザが視聴したい対象である視聴対象を示す情報である。

[0116] 自動映像編集部１４２４は、映像生成部１４２３が生成する広角映像と、位置特定部１４２２が生成する視聴対象の位置を示す被写体位置情報と、情報取得部１４２５により取得されたユーザの嗜好情報とを用いて、ユーザの嗜好にあった映像ストリームを生成する。自動映像編集部１４２４は、領域算出部１４２４ａとクロッピング部１４２４ｂとを有する。

[0117] 領域算出部１４２４ａは、情報取得部１４２５により取得されたユーザの嗜好情報に基づいて、映像生成部１４２３により生成された広角映像のうちの一部の領域であって、当該広角映像の領域よりも小さいクロッピング領域を算出する。より具体的には、領域算出部１４２４ａは、広角映像のうち、位置特定部１４２２により特定された視聴対象の位置を用いて、視聴対象が含まれる領域を、クロッピング領域として算出する。ここで、領域算出部１４２４ａは、広角映像のうち、視聴対象の位置が、広角映像をクロッピングするための予め定められたサイズのクロッピング枠における所定の基準位置に一致させた場合に、当該クロッピング枠で特定される領域をクロッピング領域として算出してもよい。

[0118] そして、クロッピング部１４２４ｂは、映像生成部１４２３により生成された広角映像を、領域算出部１４２４ａにより算出されたクロッピング領域

でクロッピングする。

[0119] 図１８はその例を示している。領域算出部１４２４ａは、ユーザの嗜好情報において、視聴対象が「ボール中心の映像」を示している場合には、広角映像から、例えばボールの位置情報が真ん中に位置するように、クロッピング枠の位置を定める。そして、クロッピング部１４２４ｂは、広角映像を、クロッピング枠で特定されたクロッピング領域でクロッピングすることにより、ユーザの好みの映像を生成する。つまり、図１８の（ａ）の例では、黒枠（クロッピング枠）で囲まれたクロッピング領域がユーザに提供する映像（クロッピング映像）となる。また、ユーザの嗜好情報が「特定の選手中心の映像」を示している場合には、広角映像から、特定の選手の位置情報が真ん中に位置するように、クロッピングを行い、ユーザの好みの映像を生成する。つまり、図１８の（ｂ）の例では、特定の選手（つまり、視聴対象）が選手Ａである場合に、黒枠（クロッピング枠）で囲まれたクロッピング領域のクロッピング映像がユーザに提供する映像となる。クロッピング部１４２４ｂによりクロッピングされた映像は、映像提供部１４２６により、圧縮符号化され、音声と共に多重化されてシステムストリームとして出力する。つまり、映像提供部１４２６は、クロッピング部１４２４ｂによりクロッピングされることにより生成されたクロッピング映像をシステムストリームとしてユーザに提供する。なお、自動映像編集部１４２４によって生成されるシステムストリームを以降で通信ストリームと呼ぶことにする。

[0120] なお、ここで広角映像からのクロッピング方法は、図１９に示すような幾つかの方法があり、用途に応じて適宜選択して利用すればよい。図１９の（ａ）は、広角映像から矩形領域を切り出す方法である。図１９の（ｂ）および（ｃ）の方法は、３次元のオブジェクトを構成して広角映像を表示する方法である。広角映像の表示方法としては一般的にこの方法が利用される。具体的には、ＯｐｅｎＧＬ等の３次元描画ライブラリを利用して、３次元の座標上に円柱モデルを生成し、パノラマ映像をテクスチャとして、図１９の（ｂ）のように円柱モデルの表面内側に貼り付けを行う。広角映像のフレーム

レートに応じて、広角映像をデコードしてテクスチャの更新を行う。図１９の（ｃ）は、図１９の（ｂ）で示す円柱を上からみた図である。図１９の（ｃ）に示すように、ユーザの視点は３次元座標上で円柱の中心に配置しており、この視点位置から矢印で示す視線方向に円柱の３次元モデルをみた映像を透視投影することで、視点からみた映像を広角映像からクロッピングして表示することが可能となる。例として「ボール」を中心に視聴するケースの場合には、広角映像のテクスチャが貼り付けられた円柱の表面における、ボール位置の座標を特定し、このボール位置に視点位置からの向きを設定すれば、ボール位置を中心にしたクロッピング再生が可能となる。なお、円柱モデルではなく、球モデルに、広角映像のテクスチャを貼り付けるように構成してもよい。この場合は、円柱モデルと同じ方法で、視点位置を球の中心に配置し、向きと画角とから透視投影することでクロッピング映像を得ることが可能となる。

[0121] なお、視点位置を円柱の中心に配置して、向きと画角を変えてクロッピングするとしたが、図２０に示すように視点位置は必ずしも中心でなくてもよい。図２０の（ａ）の例では、中心ではなく、その後方の円周上に配置している。この場合には、中心より後方に視点位置を配置することで歪みが小さくなり、映像によってはこちらの方がよいケースがある。また、特に円周上に配置すれば、円周角の定理により画角は中心の半分で計算が容易でなる。この場合には、視点の向きは固定にし、図２０の（ｂ）のように円柱自体を、円の中心と円の中心を結ぶ軸を中心に回転させる。図２０（ｂ）の例では、ボールが広角映像において左側に移動する場合には、円柱を右側に回転する。このように構成すれば、視点向きは固定であってもボール位置に追従してクロッピング映像を生成できる。

[0122] また、音声データの生成は、映像撮影部によって収音された音声データを利用することで生成できる。なお、図１５のように複数台の映像撮影部によって広角映像を生成する場合には、クロッピングされている領域を撮影する映像撮影部の音声データを選択するように構成すると、映像と音声の関係性

が増すため臨場感のある音声データを生成できる。なお、図１５のように複数台の映像撮影部によって広角映像を生成する場合には、クロッピングされている領域の位置に応じて、映像撮影部の音声合成係数を変えて映像データを生成してもよい。具体例を図２１に示している。図２１のイメージは、複数の映像撮影部によって撮影された映像を合成した広角映像であり、イメージの上段の矢印はそれぞれの映像撮影部（第一カメラ、第二カメラ、第三カメラ）の撮影領域を示している。これらカメラによって収音される音声データに対する音声合成係数の例をイメージの下部の矢印に示している。ｋ１は第一カメラの音声データに対する音声合成係数、ｋ２は第二カメラの音声データに対する音声合成係数、ｋ３は第三カメラの音声データに対する音声合成係数である。この音声合成係数の例は、クロッピングされる中心位置に応じて変動するとする。例えば、図２１において、クロッピング領域が黒枠領域の場合で、中心が黒丸の地点である場合には、ｋ１は０．５、ｋ２は０．５、ｋ３は０．０であるため、各音声データにこの係数を掛け合わせて合成することで、合成音声データを生成する。このように構成すれば、映像と音声の関係性が増すためより臨場感のある音声データを生成できる。なお、クロッピング領域の位置と音声合成係数の関係は、図２１は一例にすぎず、コンテンツの特性やユーザ嗜好に応じて変更してもよい。

[0123] なお、音声データの生成は、位置特定部１４２２によって生成される被写体位置情報または映像データを利用し、そのシーンの意味を解析することで、効果音を合成してもよい。例えば、サッカーを例に取ると、ボール位置情報が選手位置情報と近接した直後、選手位置情報からボール位置情報が離れ、ボール位置情報が一定速度以上でゴールに向かう場合には、選手がシュートしたタイミングであることを特定できる。このため、あらかじめ用意した、選手がシュートを行うときの一般的な音を、効果音として合成してもよい。同様にゴールポストに当たったり、キーパーがボールをキャッチしたりすることを被写体位置情報または映像データの画像解析によって判断すれば、その行為に対応する効果音を合成することにより、迫力のある音声をユーザ

に提供することが可能となる。

[0124] 通信Ｉ／Ｆ１４２１、１４２７は、インターネットと接続するためのＩ／Ｆを示しており、例えばＮＩＣであり、ルータ等を通してインターネットと接続するＩ／Ｆである。

[0125] （編集システムの動作）

編集システム１４２０は、映像提供方法として以下の処理を行う。

[0126] 図２２は、編集システム１４２０により行われる映像提供処理の流れを示すフローチャートである。

[0127] まず、映像取得部としての通信Ｉ／Ｆ１４２１は、第一主映像１５１１、第二主映像１５１２および第三主映像１５１３を取得する（Ｓ２２０１：映像取得ステップ）。

[0128] 次に、映像生成部１４２３は、通信Ｉ／Ｆ１４２１により取得された、第一主映像１５１１、第二主映像１５１２および第三主映像１５１３から広角映像を生成する（Ｓ２２０２：映像生成ステップ）。

[0129] また、情報取得部１４２５は、通信Ｉ／Ｆ１４２７を介したネットワークを経由して、ユーザの嗜好情報を取得する（Ｓ２２０３：情報取得ステップ）。

[0130] 位置特定部１４２２は、ユーザの嗜好情報に基づいて広角映像に対して画像認識を行うことで、広角映像における視聴対象の位置を特定する（Ｓ２２０４：位置特定ステップ）。

[0131] 領域算出部１４２４ａは、位置特定部１４２２によって特定された視聴対象の位置を用いて、視聴対象が含まれる領域をクロッピング領域として算出する（Ｓ２２０５：領域算出ステップ）。

[0132] 映像提供部１４２６は、クロッピングによりクロッピングされることにより生成されたクロッピング映像を再生システムに送信することでユーザに提供する（Ｓ２２０６：映像提供ステップ）。

[0133] （再生システム）

再生システム１４３０は、通信Ｉ／Ｆ１４３１、ストリームデコード部１

４３２、アプリケーション実行部１４３４、および入力Ｉ／Ｆ１４３３から構成され、編集システム１４２０が生成する通信ストリームを再生する、例えば、デジタルテレビなどの端末である。なお、再生システム１４３０は、送信装置として機能する編集システム１４２０にネットワークを介して接続される受信装置として機能し、編集システム１４２０から送信される映像を受信する。

[0134] 通信Ｉ／Ｆ１４３１は、例えばＮＩＣでありインターネットと接続するためのＩ／Ｆである。

[0135] ストリームデコード部１４３２は、通信ストリームをデコードする。ストリームデコード部１４３２は通信ストリーム内の圧縮符号化されたビデオストリームをデコードして、非圧縮のイメージ映像を生成して、ビデオプレーンに出力しテレビ等に出力する。ストリームデコード部１４３２は、通信ストリーム内に圧縮符号化されたオーディオストリームをデコードし、非圧縮のＬＰＣＭ状態のオーディオフレームを生成して、テレビなどのスピーカに出力する。

[0136] アプリケーション実行部１４３４は、通信Ｉ／Ｆ１４３１経由で伝送されるアプリケーションを実行する実行制御部である。アプリケーション実行部１４３４は、例えば、アプリケーションがＨＴＭＬコンテンツである場合にはＷｅｂブラウザであり、アプリケーションがＪａｖａ（登録商標）であれば、Ｊａｖａ（登録商標）ＶＭとなり、各種ＡＰＩを経由して、再生装置の各処理部にアクセスすることが可能である。また、アプリケーションは、再生制御のＡＰＩを経由して、ストリームデコード部１４３２の再生、停止等をコントロールする。また、アプリケーションは、グラフィックスの描画ＡＰＩを経由して、グラフィックスデータをグラフィックスプレーンに出力し、ストリームデコード部１４３２が出力するビデオプレーンに合成してテレビ等に出力することで、ユーザにグラフィックスによるメニュー等を提示できる。また、アプリケーションは、入力Ｉ／Ｆ１４３３からのデータを取得して、ユーザの指示に合わせて画面の表示内容を変更させることで、グラフ

ィカルユーザインターフェースを実現する。

[0137] 入力Ｉ／Ｆ１４３３は、再生システムに対して、ユーザの意図を示す情報を入力するＩ／Ｆであり、例えばリモコンである。入力された情報はアプリケーション実行制御部に入力される。

[0138] 以上が、本実施の形態に係る、配信・視聴システム１４００の説明である。

[0139] なお、複数の映像撮影部１４０２による同期の取り方として、図１５の例では、複数のカメラ間でネットワーク接続を行い、撮影制御部１４０１による同期信号をストリームに埋め込む方法を説明したが、同期の合わせ方は下記に示す方法のいずれかを用いてもよい。

[0140] １．ＧＰＳ情報をストリームに埋め込む方法

映像撮影部１４０２にはそれぞれＧＰＳ受信機が搭載すれば、ＧＰＳ衛星からのＧＰＳ情報を受信できる。ＧＰＳ情報には、衛星に搭載された原子時計による時刻データが格納されているため、その情報を使えば、複数の映像撮影部１４０２によって作成されたストリーム間の同期を取ることが可能となる。また、ＧＰＳ情報の場所情報を使うことで、複数の映像撮影部１４０２によって作成されたストリームの関係性が特定できる。つまり、サーバにアップロードされた複数の映像ストリームがある場合に、広角映像を構成するためのストリームの組み合わせを、位置情報を用いて判断することが可能となる。なお、撮影制御部１４０１のみがＧＰＳ情報の受信機を有していてもよく、この場合には、撮影制御部１４０１がＧＰＳ情報を取得して、その情報を無線や有線等の通信部を通じて、各映像撮影部１４０２に伝送するような構成となる。

[0141] ２．複数の映像撮影部１４０２から撮影された映像を別デバイスに蓄積／伝送する方法

図２３は、図１５の構成に対して、同期制御部２３０１が追加されている。同期制御部２３０１は、カメラ１５０１～１５０３から撮影された映像を有線（例えばＨＤＭＩ（登録商標））または無線でそのまま入力し、各映像

ストリームに対して、同期信号を付与して、ＳＤカード等のデバイスに格納したり、通信Ｉ／Ｆを経由して、ネットワーク上の編集システムにアップロードしたりする。このため、各カメラ１５０１～１５０３側で、同期信号を設定することなく、同期を取ることが可能となる。

[0142] ３．カチンコ／時計を複数の映像撮影部１４０２で撮影した後、画角を変える方法

カチンコまたは時計を複数の映像撮影部１４０２で撮影した後に、複数の映像撮影部１４０２を所定の向きになるように画角を変えれば、複数の映像撮影部１４０２によって撮影された映像のそれぞれにカチンコまたは時計が含まれることになる。このため、カチンコや時計が撮影されたストリームに対して、画像解析を行うことで、カチンコであれば、カチンコがたたかれた瞬間、時計であれば同一時刻のフレームを特定することにより、複数のストリーム間の同期を取ることが可能となる。

[0143] ４．強度が変動する光を複数の映像撮影部１４０２に対して照射する方法

強度が変動する光を複数の映像撮影部１４０２に対して照射することで、複数の映像撮影部１４０２によって撮影された映像のそれぞれに同一の光が照射された映像が含まれることになる。つまり、同一の光が照射された複数のストリームに対して、光の時間的な強度の差を特定するための画像解析を行うことで、同一強度のフレームを特定できる。このように同一強度のフレームを特定できるため、複数のストリームの同期を取ることが可能となる。

[0144] ５．映像撮影部１４０２からサーバにアップロードされた時刻情報を用いる方法

複数の映像撮影部１４０２からリアルタイムにサーバへアップロードされている場合には、サーバへの到着時刻を用いて同期を取るための参考値として利用してもよい。

[0145] なお、図１５に示したように、複数の映像撮影部１４０２は、イベント全体が写るように向きや位置を固定するとしたが、ユーザによる複数の映像撮影部１４０２の向きや位置の設定支援のために、次のような方法を導入して

もよい。図２４に示すように、撮影制御部１４０１に複数の映像撮影部１４０２の映像データを伝送し、合成時の映像を確認できる構成にする。図２４は図１５に比べて、撮影制御部であるタブレット型端末２４０４の構成が異なる。図２４のタブレット型端末２４０４は、上述した編集システム１４２０が備える映像生成部１４２３と同じ機能を有する。つまり、タブレット型端末２４０４は、複数の映像撮影部１４０２により撮影された各映像ストリームと、各映像ストリームが映像生成部の機能により合成された広角映像とを表示する。このようにすることで、ユーザは複数の映像撮影部１４０２の位置や向きを、映像を見ながら確認できる。なお、複数の映像撮影部１４０２によって撮影された映像ストリームの表示では、図２４の各映像１５１１、１５１２、１５１３のようにオーバラップ領域（重複領域）を囲み枠や色等で表示するように構成してもよい。なお、タブレット型端末２４０４に表示される映像は、複数の映像撮影部１４０２の向きや位置の設定の確認をするための映像なので、必ずしも動画でなくてもよく、同一時刻の静止画であってもよい。

[0146] なお、広角映像は、タブレット型端末２４０４で作成されなくてもよい。例えば、複数の映像撮影部１４０２で撮影された複数の映像が、ネットワーク上のサーバにアップロードされ、サーバが有する映像生成部によって広角映像を生成してもよい。なお、この場合に、サーバで生成された広角映像をタブレットにダウンロードして表示するように構成してもよい。このように構成すれば、タブレット型端末２４０４の広角映像の生成に係る処理負荷を削減できる。

[0147] なお、上述のようにして広角映像が正しく生成できない場合には、ワーニングメッセージと共に、合わせるためのアドバイスを提示するようにしてもよい。例えば、「右側のカメラのズーム率を変えてください」「左側のカメラの位置を右側に寄せてください」といったメッセージである。このように構成すれば、ユーザは指示に従って、容易にカメラの設定を実現できる。

[0148] なお、映像撮影部１４０２にパンチルトズーム操作制御コードを受けとり

、カメラのパンチルトズーム操作を実行する機能を備えていれば、撮影制御部１４０１がパンチルトズームを計算し、制御コードを各映像撮影部１４０２に伝送することで、最適なカメラの向きおよびズーム率となるようにカメラの設定の自動調整が行われるようにしてもよい。例えば、カメラ間に死角が発生して被写体が隠れてしまう場合のように広角映像が正しく生成できない場合には、当該死角が発生しないように、カメラの向きを内向きに移動するように撮影制御部１４０１によりコードが伝送されることになる。このようなプログラム操作によるカメラの自動パンチルト操作を実現するカメラとしては、ＰＴＺカメラが有名であり、映像撮影部１４０２はこのようなカメラを使うことで実現できる。また、撮影制御部１４０１は、イベント全体を複数の映像撮影部１４０２によって撮影することができない場合には、不足部分をアラーム等やメッセージ等で通知してもよい。

[0149] なお、無線または有線上の信号によって映像撮影部１４０２のカメラパラメータの設定を制御できるようにすれば、撮影制御部１４０１は、映像撮影部１４０２のカメラパラメータを均一にするように制御できる。例えば、映像撮影部１４０２はホワイトバランス等のカメラパラメータを一致させることで、広角映像にしたときの色の違いを小さくできる。なお、カメラパラメータは、複数の映像撮影部１４０２において、最も性能が低いものに合わせるようにしてもよい。例えば、第一カメラが１９２０ｘ１０８０６０ｐ映像が撮影できるカメラ、第二カメラが１９２０ｘ１０８０３０ｐ映像が撮影できるカメラ、第三カメラ３が１２８０ｘ７２０３０ｐ映像が撮影できるカメラである場合には、全てのカメラを１２８０ｘ７２０３０ｐで動作させる。このようにすることで、合成された広角映像の品質の歪みを少なくできると共に、映像のアップコンバートやダウンコンバートといった処理の削減ができる。

[0150] なお、自動映像編集部１４２４が、ユーザの嗜好情報を用いて、広角映像から被写体の位置情報に追従してクロッピングして映像を生成する方法において、下記に示す方法のいずれかを用いれば、事前でより快適な映像を生成

することができる。

[0151] １．視聴対象の画面上の位置情報に対してローパスフィルタをかける方法

図２５の（ａ）は、被写体の位置情報（ここではＸ座標の値）の時間的な推移を示している。位置情報をそのまま使用して、クロッピングを行う場合には、被写体の小刻みな動きにも追従してしまうため、ぶれた映像になってしまい、ユーザにとって見づらい映像となってしまう。そこで図２５の（ｂ）に示すように、前後位置情報を使ってローパスフィルタをかけた位置情報を使って計算された位置情報（黒丸で記した点）をクロッピングに利用することで、画面ぶれの少ない見やすい映像をユーザに提供できる。具体的な位置情報の計算方法は、次のように行う。時刻Ｔの位置座標を求める場合には、時刻（ｔ－Ｎ）から時刻（ｔ＋Ｍ）までの被写体の位置情報の合計を、Ｎ＋Ｍ＋１で割り算をすることで、求められる。計算式は図２１下段に乗せている。ＮおよびＭの値は、例えば、Ｎ＝Ｍ＝０．５秒分等の一定の間隔を与えるようにする。ただし、ｋがマイナス値やストリーム終端を越えないようにＮおよびＭの値は調整される。ＮおよびＭの値は、コンテンツ毎に異なる値に定められていてもよい。ＮおよびＭの値は、例えば、サッカーの場合０．５秒、バスケットボールの場合０．４秒等である。このように構成することで、コンテンツの特性に合わせた制御が可能となる。なお、ユーザがＮおよびＭの値を設定できるようにしてもよい。このように構成すれば、ユーザの好みを反映することが可能となる。

[0152] ２．視聴対象の位置情報として過去の時刻を採用する方法

クロッピングを行う基準となる視聴対象の位置情報として、ストリームの再生時刻（ｔ）よりも時間的に過去の時刻（ｔ－Ｄ）の位置情報を利用する。図２６にボール位置情報に追従してクロッピング領域を設定する場合の例を模式的に示している。図２６の（ａ）は、ボール位置情報の移動と同時にクロッピング領域も移動する場合の例を示している。この場合、ボールの位置情報にクロッピング領域が追従しすぎてしまい、ユーザにとっては違和感がある。クロッピングの移動がボールの動きを予測しているかのような印象

を与えてしまうからである。そこで図２６の（ｂ）のように、ボールの位置情報の移動よりも遅れて、クロッピング領域を移動させる。図２６の（ｂ）の例では、クロッピング領域の基準となる黒丸はビデオフレームの表示時刻（ｔ）よりも少し遅れた時刻（ｔ－Ｄ）地点におけるボールの位置情報を示し、クロッピング領域を特定するための黒枠（クロッピング枠）は、時刻（ｔ－Ｄ）地点におけるボールの位置情報が真ん中に設定されるようにクロッピングされた領域を示す。つまり、領域算出部１４２４ａは、広角映像のうち、処理対象のフレームよりも所定時間（遅延量Ｄ）以前のフレームでの視聴対象の位置（黒点）が、クロッピング枠（黒枠）の所定の基準位置（クロッピング枠の中心）に一致させた場合に当該クロッピング枠で特定される領域をクロッピング領域として算出する。

[0153] このように構成することで、ボールの位置情報とクロッピング領域の関係の緩やかになり、違和感がない人間が撮影したかのような印象を与える映像となる。つまり、人間がカメラをパンさせようとすれば、視聴対象の移動に追従する形となるため、人間によるパン操作は、基本的に被写体の動いた後に行われる。このため、カメラの動きを所定の基準で遅延させることにより、人間が撮影するかのような自然な印象をユーザに提示することができる。なお、この遅延量Ｄは、ユーザが設定できるようにしてもよいし、コンテンツの特性に応じて変更されるようにしてもよい。

[0154] ３．クロッピング領域のサイズを被写体の位置情報に応じて変更する方法

図２７に示すように被写体の位置情報に応じて、チルトおよびズームが行われたかのようにクロッピング領域のサイズを変更するように構成してもよい。図２７の（ａ）は、クロッピング領域のサイズが変更される前の映像を示しており、図２７の（ｂ）は、クロッピング領域のサイズが変更された後の映像を示している。クロッピング領域のサイズ変更は、位置情報における縦方向の座標値を用いることにより変更させることができる。なお、クロッピング領域のサイズは、ユーザによって設定できるようにしてもよい。例えば、タブレット上でピンチ操作によって、クロッピング領域のサイズの拡大

・縮小ができれば、ユーザに分かりやすい。

[0155] なお、ユーザ意図を反映した自動映像編集部１４２４によるクロッピング再生の方法として、図１８の例では、排他的に一つの視聴対象のみを選択する構成を説明したが、ユーザが複数の視聴対象を同時に見たいといった意図を反映させるようなケースには、複数の視聴対象の位置情報の平均値が、画面の真ん中に来るようにクロッピング領域の位置が設定されるようにしてもよい。図２８にその例を示しており、図２８の例では、選手Ａとボールの位置情報の平均値がセンターに来るようにクロッピング領域が設定されている。このように構成すれば、複数の視聴対象にフォーカスした映像を楽しむことが可能となる。なお、興味のある視聴対象の優先順位をつければ、複数の視聴対象の位置情報の平均値ではなく、重み付けづけされた平均値の値としてもよい。例えば、ボールよりも選手Ａの優先度が高い場合には、（ボール位置情報＊２＋選手Ａの位置情報＊１）／３とすることで、重み付けされた平均値にできる。

[0156] なお、ユーザの嗜好情報として、「選手」や「ボール」といったものを指定するとしたが、例えば「俯瞰」および「ズーム」といった好みの映像アングルを指定するための情報であってもよい。例えば、「俯瞰」が選択される場合には、自動映像編集部１４２４は、広角映像からクロッピングする場合にコート全体が俯瞰できるような映像を中心に配信することになる。また、例えば、「ズーム」が選択される場合には、自動映像編集部１４２４は、広角映像からクロッピングする場合にややズームした映像を中心に配信することになる。ユーザの嗜好情報として、ユーザがより具体的にクロッピング領域を通知してもよい。この場合には、タブレット型端末等に広角映像とクロッピング領域を示すクロッピング枠との両方を表示させ、ユーザはクロッピング領域をピンチイン／ピンチアウトしてそのサイズおよび／または位置を変更して、変更したクロッピング領域の領域情報を編集システム１４２０に通知してもよい。このように構成すれば、ユーザは、ターゲットとしての視聴対象の好みだけでなく、映像の種類の好みも反映できる。また、ユーザの

嗜好情報としてテレビのサイズを編集システム１４２０に通知してもよい。より具体的には、テレビのサイズが大きければ俯瞰映像となるようにクロッピング領域を変更し、また、テレビのサイズが小さければズーム映像となるようにクロッピング領域を変更すれば、ユーザはデバイスの大きさに適した映像視聴が実現できる。

[0157] なお、本実施の形態１に係る配信・視聴システム１４００においては、編集システム１４２０は、ユーザの嗜好情報に合わせて、広角映像からクロッピングした映像を圧縮および符号化してユーザの端末に伝送するとしたが、広角映像自体を圧縮および符号化してユーザの端末に伝送して、クロッピングにかかる処理をユーザの端末である再生システム１４３０により行わせるように構成してもよい。この場合には、ユーザに伝送するストリームのビデオストリームの補足データ等には、各フレームに対するボール位置や人物位置を示す座標情報が格納されている。この場合に、再生システム１４３０が自動映像編集部を有していれば、ユーザの嗜好情報に応じて、広角映像とストリームに埋め込まれた座標情報を使い、再生システム内でクロッピング処理を行い、テレビ等の表示端末に表示させることができる。このように構成すれば、ユーザの嗜好情報をネットワーク上に伝送する必要がなく、レスポンスを高速化することが可能となる。なお、ストリームに位置情報を埋め込む場合には、ＧＯＰ先頭やシーン先頭に、ＩＤと人物名や静止画が紐づけて格納し、それ以降のフレームにＩＤと位置情報とを格納するようにしてもよい。このように格納すれば、全てのフレームに、人物名や静止画を格納するよりもデータ量が少なく効率的である。なお、このストリームの構造および再生方法は放送波で実現してもよいことは言うまでもない。

[0158] なお、通信Ｉ／Ｆ１４０３、１４２１、１４２７、１４３１によって接続されるネットワークの一部または全ては、インターネットでなく、ローカルエリア上のネットワークであってもよいことは言うまでもない。

[0159] なお、映像生成部１４２３は、撮影システム上に存在してもよい。その場合には、撮影システム１４１０が広角映像のストリームを生成し、生成した

広角映像のストリームを編集システムに伝送することになる。このため、編集システムは広角映像として、伝送されたストリームを使用するような構成となる。また、例えば、映像撮影部１４０２が、広角映像を撮影可能なワイドレンズを格納し、８Ｋ４Ｋのような高解像度カメラである場合には、映像生成部１４２３は不要となり、映像撮影部１４０２が撮影した映像ストリームを編集システムに伝送するような構成にしてもよい。つまり、このような場合には、映像生成部は撮影システムにも編集システムにも搭載しなくてもよい。

[0160] なお、映像撮影部１４０２は、ビデオカメラに限らず、カメラ機能を搭載するスマートフォン等によって構成されていてもよい。スマートフォンを並べて撮影する場合の問題点は、イベントの撮影途中で、該当端末に電話がかかってくるケースである。この場合には、該当スマートフォンと認証がなされた友人や奥さん等の携帯を代理受話器として、代理受話器にかかってきた電話を転送するようにしてもよい。例えば、パパのスマートフォンを撮影に利用する場合に、電話がかかってきた場合には、ママの携帯に「パパにＸＸさんから電話」と表示して知らせ、その携帯で通話できるようにしてもよい。

[0161] なお、複数の映像撮影部１４０２で撮影された映像コンテンツを編集システムに伝送する場合には、それぞれの端末からネットワーク経由で送るのではなく、一度１つの端末に集めて伝送するように構成してもよい。例えば、Ａ、Ｂ、およびＣのスマートフォンで撮影した場合には、ＡにＢおよびＣの映像コンテンツを集めてもよい。集める方法は、Ｗｉ－ＦｉやＷｉＧｉｇなどの無線伝送やＳＤカードでのデータ受け渡しなどが考えられる。そして、Ａのスマートフォンで、一括で映像コンテンツを編集システムに伝送する。伝送の方法は、Ｗｉ－ｆｉ、ＬＴＥ、３Ｇなどの無線伝送、有線ＬＡＮなどの有線伝送が考えられる。このように、映像コンテンツをばらばらに伝送するのではなく、一括で伝送することで、コンテンツの管理、紐付けが容易となる。

[0162] （実施の形態２）

本実施の形態１では、個人嗜好を反映した映像コンテンツの視聴を実現するための配信・視聴システム１４００について説明を行ったが、本実施の形態では、より視聴の楽しさを演出する、映像コンテンツの高度な編集を実現するための配信・視聴システム２９００の実現方法について説明する。

[0163] 図２９は、本実施の形態２に係る配信・視聴システムを示している。基本的な構成は図１４で説明したシステムと同じであるため説明を省略し、差分のみを説明する。

[0164] 撮影システム２９１０は、撮影制御部１４０１、広角映像を生成する映像撮影部１４０２、および通信Ｉ／Ｆ１４０３の他に、スポット映像撮影部２９０１が追加されている。

[0165] スポット映像撮影部２９０１は、イベントを広角映像とは別視点からの映像を撮影する固定カメラである。図３０の例では、イベント全体の広角映像を撮影する映像撮影部としてカメラ１５０１～１５０３が配置されているが、そのカメラとは別に、スポット映像撮影部２９０１としてカメラ３００１－３００７が配置されている。カメラ３００１－３００７は、それぞれの視点から映像を撮影する。スポット映像撮影部２９０１は、映像撮影部１４０２と同様に撮影制御部１４０１によって撮影が制御され、撮影された映像ストリームは通信Ｉ／Ｆ１４０３を経由して編集システム２９２０に伝送される。スポット映像撮影部２９０１によって撮影された映像ストリームは、映像撮影部１４０２によって撮影された映像ストリームと同様の手段により同期を取ることが可能である。スポット映像撮影部２９０１は、第一主映像１５１１、第二主映像１５１２、および第三主映像１５１３と同じタイミングで、撮影空間のうちの少なくとも一部の空間が第一主映像１５１１、第二主映像１５１２、および第三主映像１５１３とは異なる角度で撮影された副映像としてのスポット映像を撮影する。

[0166] 編集システム２９２０は、図１４の編集システム１４２０とは、自動映像編集部１４２４の代わりに、自動映像選択編集部２９０２となっていること

が異なる。また、通信Ｉ／Ｆ１４２１が、第一主映像１５１１、第二主映像１５１２、および第三主映像１５１３の他に、さらに、スポット映像を取得することが異なる。

[0167] 自動映像選択編集部２９０２は、映像生成部１４２３が生成する広角映像と、スポット映像撮影部２９０１によって撮影されたスポット映像と、位置特定部１４２２が生成する被写体位置情報と、情報取得部１４２５が取得するユーザの嗜好情報を用いて、ユーザの嗜好にあった映像ストリームを生成する。自動映像選択編集部２９０２は、図１４の自動映像編集部１４２４と比較して、領域算出部１４２４ａと、クロッピング部１４２４ｂとを有する他に、さらに、シーン分割部２９０２ａと、評価部２９０２ｂと、映像選択部２９０２ｃとを有する点が異なる。つまり、自動映像選択編集部２９０２は、スポット映像撮影部２９０１によって撮影される映像を提供映像の生成に利用する点が自動映像編集部１４２４とは異なり、ユーザ意図を反映する映像を生成する際に、広角映像とスポット映像との中で、最適な映像を選択して、通信ストリームを生成する点が異なる。シーン分割部２９０２ａは、クロッピング部１４２４ｂによりクロッピングされたクロッピング映像と、映像取得部としての通信Ｉ／Ｆ１４２１により取得されたスポット映像とのそれぞれを、所定のアルゴリズムに基づいて複数のシーンに分割する。評価部２９０２ｂは、シーン分割部２９０２ａにより分割された複数のシーンのそれぞれを、情報取得部１４２５により取得されたユーザの嗜好情報と、所定の評価指標とに基づいて評価する。映像選択部２９０２ｃは、情報取得部１４２５により取得されたユーザの嗜好情報に基づいて、シーン分割部２９０２ａにより分割された複数のシーンのそれぞれについて、クロッピング映像およびスポット映像のいずれかを選択する。また、映像選択部２９０２ｃは、評価部２９０２ｂによって評価された結果に基づいて、複数のシーンのそれぞれについて、クロッピング映像およびスポット映像のいずれかを選択してもよい。

[0168] 図３１は、その自動映像選択編集部２９０２による編集例である。図３１

の左は、撮影対象のシーンとカメラとの構成を示す。図３１は、サッカーの１シーンの例である。具体的には、選手１および選手２がコート上にいて、選手１が下方向にドリブルで攻め上がり、かつ、選手２が選手１の攻め上がりに対する守りを行うシーンであって、選手１が選手２をドリブルで抜くシーンを示している。ここで、ユーザの嗜好情報は、「選手１」と「ボール」との両方が選択されているとする。

[0169] ここで、映像撮影部１４０２としてコート全体を俯瞰するように複数台のカメラからなるカメラＣが配置されており、映像生成部１４２３は、カメラＣによって撮影された複数の映像から広角映像を生成する。自動映像選択編集部２９０２は、映像生成部１４２３により生成された広角映像に対して、ユーザの嗜好情報からクロッピング処理を行い、クロッピング映像からなる映像ストリームを生成する。ユーザの嗜好情報は、選手１およびボールが選択されているため、選手１およびボールの位置情報の平均値が画面の真ん中になるように広角映像からクロッピングされた映像が生成される。その画面イメージは、３１０３の列に示している。また、スポット映像撮影部２９０１としてカメラＡおよびカメラＢが配置されており、固定された位置でイベント映像を撮影する。これらの映像は、自動映像選択編集部２９０２に伝送される。その画面イメージは、それぞれ３１０１、３１０２に示している。

[0170] 自動映像選択編集部２９０２は、これらの映像から１つの映像を選択して、通信ストリームを生成する。広角映像からクロッピングしたクロッピング映像とスポット映像撮影部２９０１により撮影されたスポット映像とは同期されているため、これらの複数の映像から１つの映像を選択してつなぎ合わせれば、時間の流れが一定のコンテンツを生成できる。つまり、時間が過去に戻ったり、未来に飛んだりしない映像コンテンツを生成できる。

[0171] 自動映像選択編集部２９０２は、複数の映像（クロッピング映像および複数のスポット映像）から１つの映像を選択するために、位置特定部１４２２が生成する被写体位置情報を利用する。例えば、図３１の例において、時刻ｔ１における位置情報を３１０４、時刻ｔ２における位置情報を３１０５、

時刻ｔ３における位置情報を３１０６で示している。各丸で記した物体は、人物とボールおよびカメラの位置を示す。つまり、数字の「１」および「２」が人物の位置を示し、アルファベットの「Ａ」、「Ｂ」、および「Ｃ」がカメラの位置を示し、黒丸がボールを示している。ここで、「視聴対象が近く、邪魔がないものを選択する」というロジック（所定のアルゴリズム）によって、映像を選択する場合の自動映像選択編集部２９０２による映像選択処理の例を示す。ここでユーザの嗜好情報によって選択される対象は「選手１およびボール」であるとする。なお、図３１の例では、映像選択部２９０２ｃは、映像選択処理を各映像のフレーム毎に行っている。つまり、図３１では、自動映像選択編集部２９０２のうちで、領域算出部１４２４ａ、クロッピング部１４２４ｂ、および映像選択部２９０２ｃが機能しており、シーン分割部２９０２ａおよび評価部２９０２ｂは機能していない例を示している。つまり、図３１を用いて説明する処理では、シーン分割部２９０２ａおよび評価部２９０２ｂの構成はなくてもよい。

[0172] 時刻ｔ１のフレームにおいては、視聴対象である選手１およびボールには、複数のカメラのうちでカメラＡが一番近く、カメラＡと視聴対象との間には、視聴対象以外の物体は存在しない。このため、映像選択部２９０２ｃは、複数のカメラのうちで視聴対象に一番近いカメラＡの映像を選択する。

[0173] 時刻ｔ２のフレームにおいては、視聴対象である選手１およびボールには、カメラＡが一番近いが、カメラＡおよび視聴対象の間には、視聴対象ではない選手２が存在しており、カメラＡで撮影された映像において視聴対象を確認しにくい。このため、映像選択部２９０２ｃは、複数のカメラにうちで視聴対象に２番目に近いカメラであるカメラＣの映像を選択する。

[0174] 時刻ｔ３のフレームにおいては、視聴対象である選手１およびボールには、カメラＢが一番近く、カメラＢおよび視聴対象の間には、視聴対象以外の物体が存在しない。このため、映像選択部２９０２ｃは、複数のカメラのうちで視聴対象に一番近いカメラＢの映像を選択する。

[0175] このように被写体位置情報を用いて、複数の映像から１つの映像を選択す

ることで、ユーザ嗜好に応じた自動的な映像選択が実現できる。自動映像選択編集部２９０２は、選択した映像を圧縮符号化して、多重化し通信ストリームを生成する。そして、映像提供部１４２６は、自動映像選択編集部２９０２の映像選択部２９０２ｃにより選択されたクロッピング映像およびスポット映像のいずれかを通信Ｉ／Ｆ１４２７を経由してユーザに提供する。

[0176] 自動映像選択編集部２９０２による、複数の映像から１つの映像を選択する上で、図３１の例では、各時刻のフレームに応じて、使用する画像を選択するとしたが、この場合には、複数の映像の切り替わりが多くなりすぎ、ユーザにとっては見づらい映像となってしまう可能性が高い。そこで、図３２のように、コンテンツを時間軸で複数のシーンに区切り、各シーンに対して被写体の位置関係に基づき、同期する複数の映像を評価し、評価結果に基づき、該当シーンを表示する１つの映像を選択するようにしてもよい。このため、同一シーンで選択される映像は、同じカメラにより撮影される映像となる。図３２で説明する処理では、自動映像選択編集部２９０２のうちの、領域算出部１４２４ａ、クロッピング部１４２４ｂ、シーン分割部２９０２ａ、評価部２９０２ｂおよび映像選択部２９０２ｃの全てが機能する。

[0177] 図３２の構成は、図３１とカメラ、人物、ボールの位置等構成は同じである。ここで、自動映像選択編集部２９０２のシーン分割部２９０２ａは、図３２に示すように、被写体位置情報から時間軸上でシーンを区切る。図３２の例の場合は、シーン分割部２９０２ａは、所定のアルゴリズムを利用して、被写体位置情報から選手１が選手２を抜くまでのシーン１と、被写体位置情報から選手１が選手２を抜いてからのシーン２とに区切っている。シーン分割部２９０２ａは、被写体位置情報に従い、シーンのデータベースを参照することでシーンの区別を行う。なお、「シーンの区切り方」についての詳細は後述する。

[0178] そして、評価部２９０２ｂは、シーン分割部２９０２ａにより区切られたシーン毎に、同期されている複数の映像のそれぞれについて評価を行う。図３２の例では、まず、カメラＡの映像と、カメラＢの映像と、カメラＣの映

像とのそれぞれについて、シーン１の範囲内で評価を行う。評価の仕方は、例えば、「対象が近く、カメラと対象までに邪魔が少ない」というロジックであれば、シーン１の区間の範囲で、被写体位置情報を用いて、「対象が近く、カメラと対象までに邪魔が少ない」のロジックに最も当てはまる映像を選択する。より具体的には、評価部２９０２ｂは、シーン１内での、カメラから視聴対象（選手Ａおよびボールの位置情報）までの距離総和をカウントし、かつ、カメラから対象までに対象以外の物体が入る回数の総和をカウントして、それらの量に基づき評価を行ってもよい。このような同期されている複数の映像から１つの映像を選択する評価方法の具体例は後述する。図３２の例では、映像選択部２９０２ｃは、評価部２９０２ｂの評価結果に基づいて、シーン１においてはカメラＣの映像を選択している。シーン２においても同様に評価部２９０２ｂによる評価が行われ、映像選択部２９０２ｃは評価部２９０２ｂのシーン２に対する評価に基づいて図３２の例ではカメラＢの映像を選択する。このように自動映像選択編集部２９０２が、複数の同期された映像を複数のシーンに分割し、分割されたシーン毎に１つの映像を選択することにより、カメラの切り替わりを押さえ、ユーザに見やすい映像を提供できる。

[0179] 自動映像選択編集部２９０２のシーン分割部２９０２ａのシーンの区切り方について説明する。

[0180] シーンを区切る最もシンプルな方法は、一定間隔で区切る方法である。つまりシーン分割部２９０２ａは、所定のアルゴリズムによって、所定間隔毎に複数の映像のそれぞれを複数のシーンに分割してもよい。例えば、５秒間といった定数を設けて、シーンを５秒単位で区切る。このように構成すれば、カメラの切り替わりを押さえられる。しかし、一方で、コンテンツの中身に連動せずに、アングルが切り替わってしまう場合と、見にくい映像となってしまう可能性がある。例えば、サッカーのシュートを打っているシーンの中でシーンが分割され、俯瞰映像から選手のアップに切り替わってしまう可能性もあり、ユーザはボールや選手の動きや位置関係が分からなくなってし

まう。よってシーンの区切り方は、イベント上で発生している内容と連動することが好ましい。

[0181] そこで、コンテンツ毎にシーンを定義し、被写体位置情報、映像および音声を用いて、シーンの区切りを実現することが考えられる。つまり、シーン分割部２９０２ａによる複数の映像に対するシーンの分割処理に係る所定のアルゴリズムは、撮影空間内で行われているイベントの種類毎に異なっていてもよい。図３３にその例を示している。図３３はサッカーの試合の映像コンテンツにおけるシーンを示している。各シーンは、シーンの内容と、各シーンの切れ目（開始地点および終了地点）の検出アルゴリズムが定義されている。各シーンの切れ目の検出は、検出アルゴリズムを、被写体位置情報、映像または音声に対して実行することで得られる。各シーンは、シーンＩＤ、シーン内容、シーン開始、シーン開始検出アルゴリズム、シーン終了、およびシーン終了検出アルゴリズムから構成される。「シーンＩＤ」はシーンのユニークな番号、「シーン内容」はシーンの内容を記載する情報、「シーン開始」はシーンの開始のコンテンツ上の意味、「シーン開始検出アルゴリズム」はシーンの開始を検出するアルゴリズム、「シーン終了」はシーンの終了のコンテンツ上の意味、「シーン終了検出アルゴリズム」はシーンの終了を検出するアルゴリズムをそれぞれ意味する。

[0182] ここで、サッカーに限らず、スポーツのコンテンツは、大きく二つのシーンに分類できる。一つは「ゲーム中」である。ゲーム中とは、選手の動作によって得点が発生し得る状況、を意味する。サッカーの場合は、制限時間内であり、ボールがコート内にあり、反則等による中断が起こっておらず、選手の動きによって得点することが可能な状況である。バレーボール、テニスの場合は、サービスを開始し、どちらかのチーム、選手が得点を決するまでである。もう１つは「非ゲーム中」である。非ゲーム中とは、選手の動作によって得点が発生し得ない状況であり、ゲーム中と反対の状況を意味する。サッカーの場合には、制限時間を過ぎた、ボールがコートを出る、反則による中断、得点による中断等、どちらの選手がどのような動作をしても決して

得点を決めることができない状況である。バレーボールやテニスの例で言えば、どちらかのチームまたは選手が得点を決めてからサービスを開始するまでの間、もしくはハーフタイムである。このように、スポーツのコンテンツの多くは、「ゲーム中」と「非ゲーム」にシーンを分離でき、その特性は大きく異なる。「ゲーム中」は緊張状態でありユーザは目が放せない一方で、「非ゲーム中」は緩和状態でありユーザは一息つくことができる。この「ゲーム中」「非ゲーム中」の特性に応じた映像を選択することで、効果的な演出が可能となる。つまり、シーン分割部２９０２ａは、撮影空間内で行われているイベントの種類がスポーツである場合には、「ゲーム中」であるか、「非ゲーム中」であるかを所定のアルゴリズムで判定することにより、判定結果がゲーム中および非ゲーム中の一方から他方へ切り替わったタイミングで、複数の映像（クロッピング映像および複数のスポット映像）のそれぞれを、複数のシーンに分割してもよい。

[0183] 図３３のサッカーの試合の場合のシーン例を説明する。つまり、図３３は、撮影空間で行われているイベントがサッカーの試合である場合のシーン分割部２９０２ａによるシーン分割処理で実行される所定のアルゴリズムについて説明するための図である。

[0184] まず、図３３の（ａ）に示す、各シーンを判定するためのアルゴリズムについて説明する。

[0185] シーンＩＤ＝０のシーンは、「ゲーム中」を意味するシーンであり、シーン開始検出アルゴリズムはシーンＩＤ１～４の「非ゲーム中」の終了であり、シーン終了検出アルゴリズムはシーンＩＤ１～４の「非ゲーム中」の開始であると定義される。

[0186] シーンＩＤ＝１のシーンは、「非ゲーム中」を意味し、ボールがコートに出ている期間を定義するシーンである。シーン開始検出アルゴリズムは「ボールの位置情報がコート領域内から出る」、シーン開始検出アルゴリズムは「ボールの位置情報がコート領域内に入る」とある。このアルゴリズムの実行は、被写体位置情報を利用することで実現できる。つまりボールの位置情

報が、コート領域内を出るか入るかを検出することで判断が可能である。あるポイントが、ある多角形領域内に存在するかどうかの判定は、ポイントから発するレイと交差するポリゴンの線分の数を計算することで求められる。この数が奇数となれば、ポイントは内部に存在することを意味し、偶数であればポイントは外部に存在することを意味する。この問題はＰｏｉｎｔ－ｉｎ－Ｐｏｌｙｇｏｎと呼ばれ、ｏｐｅｎｃｖ等のソフトウェアで実装されている。

[0187] シーンＩＤ＝２のシーンは、「非ゲーム中」を意味し、タイムアウト期間を定義するシーンである。シーン開始検出アルゴリズムは「複数の選手位置情報がコート領域から出る」、シーン開始検出アルゴリズムは「複数の選手位置情報がコート領域から出る」とある。このアルゴリズムの実行は、被写体位置情報を利用することで実現できる。つまり複数の選手位置情報が、コート領域内を出るか入るかを検出することで判断が可能である。なお、検出アルゴリズムとしては、「笛の音」でもよい。タイムアウトを知らす笛の音には特徴があるので、音の波長の特徴をパターンマッチングさせて検出することで判断が可能である。

[0188] シーンＩＤ＝３のシーンは、「非ゲーム中」を意味し、反則による中断の期間を定義するシーンである。シーン開始検出アルゴリズムは「審判の笛の音や選手位置情報が一旦静止する」、シーン終了検出アルゴリズムは「ボール位置情報が特定位置に一旦静止され、開始」とある。シーン開始検出アルゴリズムにおいて、審判の笛の音の検出は、音の波長の特徴をパターンマッチングすることで検出でき、選手位置情報の静止は、選手の位置情報の動きを検出することで判断できる。シーン終了検出アルゴリズムにおいて、ボール位置情報の静止は、ボールの位置情報の動きを検出することで判断できる。

[0189] シーンＩＤ＝４のシーンは、「非ゲーム中」を意味し、得点による中断の期間を定義するシーンである。シーン開始検出アルゴリズムは「ボールの位置情報がゴール領域に入る」、シーン終了検出アルゴリズムは「ボール位置

情報が特定位置に一旦静止され、開始」とある。シーン開始検出アルゴリズムにおいて、被写体位置情報を利用して、ゴール領域内に、ボールの位置情報が入っているのであれば、得点がなされたと判断できる。シーン終了検出アルゴリズムにおいて、被写体位置情報を利用して、ボール位置情報の静止は、ボールの位置情報の動きを検出することで判断できる。

[0190] 図３４にアルゴリズムのフローチャートを示している。まず、時刻ｔのシーンを特定する上で、時刻ｔ－ｄのシーンを調べる。ｄはフレーム時間を示し、つまり一つ前のフレームにおけるシーンが何かを調べる。時刻ｔ－ｄのシーンが「ゲーム中」であれば、「非ゲーム中」のシーン（図３３の場合シーンＩＤ＝１～４）のシーン開始検出アルゴリズムを実行する。どれかのシーン開始が検出されれば、時刻ｔのシーンは「非ゲーム中」であると判断し、検出されなければ「ゲーム中」であると判断する。時刻ｔ－ｄのシーンが「非ゲーム中」であれば、その「非ゲーム中」のシーンＩＤのシーン終了検出アルゴリズムを実行する。シーン終了が検出されれば、時刻ｔのシーンは「ゲーム中」であると判断し、検出されなければ「非ゲーム中」であると判断する。

[0191] 図３３の（ｂ）にサッカーの試合におけるシーンの区切りの例を示している。開始後、反則やボールがコートに出たり、得点をするなどすると、「ゲーム中」から「非ゲーム中」に移行し、それ以外は「ゲーム中」という構成となる。

[0192] 以上が自動映像選択編集部２９０２のシーンの区切り方についての説明である。

[0193] 次に、自動映像選択編集部２９０２の評価部２９０２ｂによる同期されている複数の映像から１つの映像を選択するための評価方法（以降ではこの評価方法を単に「映像評価方法」と呼ぶ）の具体的な方法を説明する。映像評価方法は、映像を評価する評価指標を設けて、該当シーンの全てもしくは幾つかの時刻のフレームに対して評価し、評価値が高いものを選ぶ。つまり、評価部２９０２ｂによる評価の基準となる所定の評価指標について、以降で

示す。以降で説明する映像項目は、全てを行う必要はなく、コンテンツの特性やユーザの好みに応じて変更すればよい。

[0194] 評価指標１：ユーザの嗜好情報で指定される視聴対象（例：ボールや選手）の位置情報と映像撮影部（スポット映像撮影部を含む）との距離

評価指標１は、視聴対象が画角内にあり、視聴対象との距離が近い位置にある映像撮影部によって撮影される映像の評価を高くするための指標である。このように構成することで、ユーザは視聴対象が大きく写る映像を視聴することができる。つまり、所定の評価指標は、映像を撮影した複数のカメラのうちで、視聴対象が画角に含まれており、かつ、視聴対象への距離が近いカメラによって撮影された映像のシーンほど高く評価する指標を含む。

[0195] 評価指標２：ユーザの嗜好情報で指定される視聴対象の位置情報と映像撮影部（スポット映像撮影部を含む）の位置情報までの間に存在するオブジェクトの数

評価指標２は、当映像撮影部と視聴対象までの間に、視聴対象以外のオブジェクトが多くあれば、評価を下げるための指標である。特に審判等の選手以外の人物の場合に評価を下げる。このように構成することで、ユーザは邪魔に隠されずに視聴対象が写る映像を視聴することができる。つまり、所定の評価指標は、映像を撮影した複数のカメラのうちで、視聴対象が画角に含まれており、かつ、視聴対象との間にあるオブジェクトの数が少ないカメラによって撮影された映像のシーンほど高く評価する指標を含む。

[0196] 評価指標３：ユーザの嗜好情報で指定される視聴対象が、該当映像撮影部（スポット映像撮影部を含む）によって撮影される映像に写る面積

評価指標３は、該当映像撮影部によって撮影される映像の中に写る視聴対象の面積が大きいものを評価が高いとする。映像データに対して、顔認識等で視聴対象（選手等）を特定し、その人物の面積を求めることで得られる。ただし、評価指標１と同様に、画角に入りきらなければ意味のある映像ではないため、例えば、顔が入っていなければ、逆に評価を下げるとしてもよい。このように構成することで、ユーザは視聴対象が大きく写る映像を視聴す

ることができる。つまり、所定の評価指標は、映像を撮影した複数のカメラのうちで、視聴対象が画角に含まれており、かつ、当該映像に映り込んでいる視聴対象の面積が大きいカメラによって撮影された映像のシーンほど高く評価する指標を含む。

[0197] なお、各評価指標の重み付けはユーザの嗜好やシーンによって変更するように構成してもよい。例えば、サッカーの例でいえば、ゴール前のシーンにおいては、ターゲットとなる選手やボール以外にもゴールの位置情報も重要となるため、選手、ターゲット、ゴールの位置関係が分かる映像の評価を高めるように構成してもよい。つまり、所定の評価指標は、第一指標（評価指標１）と、第二指標（評価指標２）と、第三指標（評価指標３）とのうちの２以上の指標を含んでいる場合に、評価部２９０２ｂは、複数のシーンのそれぞれについて、当該シーンに対する２以上の指標により評価された複数の結果について、２以上に指標に関連付けられている予め定められた重み付けによって重み付加算された加算値に基づいて評価してもよい。

[0198] ここで、前述したとおり自動映像選択編集部２９０２では、シーン分割部２９０２ａがシーンを区切ることによって、コンテンツの内容に応じたシーンの区切りを実現し、この特性を使って、評価部２９０２ｂによる映像評価または映像選択部２９０２ｃによる選択方法を変更することで、ユーザにとって効果的に編集された映像コンテンツを生成することが可能となる。

[0199] 前述したとおりスポーツコンテンツにおいては、大きく「ゲーム中」と「非ゲーム中」の二つのシーンに分離でき、緊張（ゲーム中）と緩和（非ゲーム中）が繰り返される構成となっている。

[0200] 「ゲーム中」と「非ゲーム中」において、ユーザにとって見たい視点の映像は大きく異なる。「ゲーム中」においては、得点するか否かの緊迫した状況のため、選手を一人だけズームで見せるような映像ではなく、選手とボールとコートの位置関係が把握できる映像、例えば俯瞰映像が好ましい。一方で、「非ゲーム中」においては、得点が発生する状況ではない、ユーザが一息つく、あるいは、前のゲーム中を思い返す状況であるため、むしろ選手一

人ひとりにフォーカスする映像やリプレイ映像が好ましい。下記に、この「ゲーム中」と「非ゲーム中」のシーンの特性に合わせた映像選択を実現させる具体的な映像方法を列挙する。以降で説明する映像選択方法は、全てを行う必要はなく、コンテンツの特性やユーザの好みに応じて変更すればよい。

[0201] １．自動映像選択編集部２９０２は、「非ゲーム中」のシーンの区間に、その直前の「ゲーム中」で最後にボールに触れた選手が写る映像に切り替えるとしてもよい。ボールに触れた選手とは、ボール位置情報に接している位置情報の選手である。このように構成することで、「非ゲーム中」に、ユーザは、直前の「ゲーム中」のキープレーヤにフォーカスして視聴することが可能となり、より映像視聴の楽しさが増す。

[0202] ２．自動映像選択編集部２９０２は、「非ゲーム中」のシーンの区間に、その直前の「ゲーム中」で、攻撃チームの中のコート内で最後にボールに触れた選手が写る映像に切り替えるとしてもよい。このように構成することで、多くのユーザにとっての注目選手にフォーカスした映像を提供できる。多くのユーザにとって、注目選手は攻撃側の選手であるからである。

[0203] ３．自動映像選択編集部２９０２は、「非ゲーム中」のシーンの区間に、ユーザの嗜好情報にある自分の好みの選手が写る映像に切り替えるとしてもよい。このように構成することで、「非ゲーム中」の間にユーザ好みの選手にフォーカスした映像を提供できる。

[0204] ４．自動映像選択編集部２９０２は、「非ゲーム中」のシーンの区間に、ユーザの嗜好情報にある自分の好みのチームの選手が写る映像に切り替えるとしてもよい。このように構成することで、「非ゲーム中」の間にユーザ好みのチームの選手にフォーカスした映像を提供できる。

[0205] ５．自動映像選択編集部２９０２は、「非ゲーム中」のシーンの区間に、直前の「ゲーム中」のリプレイ映像に切り替えるとしてもよい。このように構成することで、「非ゲーム中」の間に、ユーザは直前のゲーム中のシーンを思い返すことが可能となる。つまり、自動映像選択編集部２９０２の映像選択部２９０２ｃは、撮影空間内で行われているイベントの種類がスポーツ

である場合、「ゲーム中」から「非ゲーム中」に切り替わったときに、当該「非ゲーム中」のシーンを選択する代わりに直前の「ゲーム中」のシーンの映像の中から選択してもよい。ここで、リプレイ映像に利用する「ゲーム中」のシーンは、直前の「ゲーム中」のシーン終了時刻で終わるように設定される。図３５にリプレイ映像に利用する区間の時間関係を示している。図３５において、ｔ１は「ゲーム中」の開始時刻、ｔ２は「ゲーム中」の終了時刻でありで「非ゲーム中」の開始時、ｔ４は「非ゲーム中」の終了時刻で次の「ゲーム中」の開始時刻である。このときに、「非ゲーム中」の先頭部分では特定選手のズームやリプレイ映像を再生することをユーザに通知するメッセージである場合に、時刻ｔ３からリプレイ映像の再生を開始するとしてもよい。この場合に、リプレイ映像の再生時間は、ｔ４－ｔ３となる。このリプレイ映像を再生スピードで割ることによって、リプレイ映像に利用する直前の「ゲーム中」のシーンの期間であるリプレイシーン時間を特定する。なお、ここでの再生スピードは、予め定められたリアルタイムよりも遅い再生スピードである。直前の「ゲーム中」の終了時刻であるｔ２から、リプレイシーン時間を減算することで、リプレイ映像に利用する直前の「ゲーム中」の先頭時刻（ｔ５）を特定する。このようにすることで、直前の「ゲーム中」の注目シーンをリプレイ再生することが可能である。

[0206] ６．自動映像選択編集部２９０２は、「非ゲーム中」のシーンの区間に、観客を撮影するカメラ映像に切り替えるとしてもよい。このように構成することで、「非ゲーム中」の間にユーザはイベントが行われる周囲の状況を知ることができる。

[0207] ７．自動映像選択編集部２９０２は、「非ゲーム中」から「ゲーム中」へ切り替わるタイミングでは、ボールを持つ選手（ボール位置情報に近接している位置情報の選手）が写る映像に切り替えるとしてもよい。スポーツにおいて「非ゲーム中」から「ゲーム中」への切り替わりにおいては、ある特定の一人の選手によって、再開されるケースが多く、例えば、サッカーの場合のスローインやコーナーキック、バレーボールの場合はサーブなどである。

「非ゲーム中」から「ゲーム中」へ切り替わるタイミングにおいて、ボールを持つ選手が写る映像に切り替わることで、ゲームの再開をユーザに正確に伝えることが可能となる。なお、サッカーの場合のスローインの場合には、スローインをする選手を後ろ側から撮影する映像に切り替えることが好ましい。このような映像を選択することで、ユーザはスローインの選手だけでなく、周りの位置関係を把握することができる。

[0208] ８．自動映像選択編集部２９０２は、「ゲーム中」のシーンの区間に、俯瞰映像に切り替えるとしてもよい。俯瞰映像は、映像生成部によって撮影される広角映像を、ユーザが指定するボールまたは選手の位置情報を元にクロッピングすることで生成できる。俯瞰映像は、ボールと選手、コートの位置関係を把握できる映像であるため、得点が発生しうる緊迫した「ゲーム中」において、得点シーンを見逃すことなくユーザは全体を見渡して視聴できる。

[0209] ９．自動映像選択編集部２９０２は、「ゲーム中」のシーンの区間に、一時的に特定の選手やボールをズームした映像に切り替えるとしてもよい。この場合に、特定の選手やボールをズームした映像によって、ユーザは、ユーザの嗜好を反映した映像を視聴できる。しかし、この場合に特定の選手のズーム映像にフォーカスすることで、得点シーンを見逃すことになっては、ユーザの楽しさが損なわれるため、「ゲーム中」において、得点が発生しない区間に限定することが好ましい。

[0210] １０．自動映像選択編集部２９０２は、「ゲーム中」のシーンの区間に、ボールを持つ選手の動きベクトルの方向を向き、かつその動きベクトルの後方に位置するカメラの映像に切り替えるとしてもよい。例えば、図３５の（ａ）の例の場合には、ボールを持つ選手の動きベクトルが矢印の場合には、動きベクトルの後方に位置するカメラ３００６に切りかえ、図３５の（ｂ）の例の場合には、ボールを持つ選手の動きベクトルが矢印の場合には、動きベクトルの後方に位置するカメラ３００７に切りかえる。このように構成することで、ボールを持って攻撃する方向を奥にした映像を視聴でき、自らが

攻め込んでいるかのような印象を与える迫力ある映像をユーザに提供できる。

[0211] 以上が本実施の形態に係る個人嗜好を反映した映像コンテンツの自動生成・視聴システムの説明である。

[0212] なお、実施の形態１および実施の形態２では、リアルタイム性（イベントをリアルタイムにユーザに提供する）について言及していないが、（１）撮影システムから編集システムへの伝送、（２）編集システムの映像生成処理、（３）編集システムから再生システムへのストリームの伝送、を短くしていけば、リアルタイム性に近づく、つまり生中継が実現できることは言うまでもない。例えば（１）の伝送で１秒、（２）の映像生成処理に１０秒、（３）の伝送に１秒の遅延がある場合には、１２秒遅れではあるが、ユーザは生中継に近くイベントのコンテンツ視聴を楽しむことができる。つまり、シーン分割部２９０２ａは、クロッピング映像および複数のスポット映像のそれぞれを複数のシーンに分割するときに、所定のアルゴリズムとは別に、所定時間毎に分割してもよい。

[0213] ここで、このように生中継に近い形でユーザに提供する場合のシーンの区切り方を説明する。この場合のシーンの分割の仕方を図３７に示している。図３７の（ａ）は、オフライン（つまり、すべてのイベントを撮影した後に編集システムを実行する場合）の例であり、この場合には、イベントの被写体位置情報等に従い、シーンを区切った後に、自動映像選択編集部で映像生成すればよいため、同一ステータス（ゲーム中もしくは非ゲーム中）の区間は、シーンとして区切れることはない。図３７の（ｂ）は、自動映像選択編集部２９０２に５秒の遅延時間が許容されている場合の例である。つまり、自動映像選択編集部２９０２は５秒後にシーンを決定すればよいとする。この場合には、シーンの終了が検出するまで待っていては、処理が遅れてしまう。つまり、５秒間の許容遅延であるにもかかわらず、１０秒間のシーンが存在する場合に１０秒後のシーン区切りまで待っていては、処理が間に合わないことになってしまう。そこで、許容遅延量が決まっている場合には、シ

ーン開始から許容遅延量までに、シーン区切りが発生しない場合には、同一ステータスであってもシーンを区切る。自動映像選択編集部は、区切られたシーン内で、上述した映像評価・選択を行う。このようにすることで、生中継に近い形であってもシーン区切りを実現できる。なお、同一ステータスのシーンが連続する場合（例：図３７の（ｂ）の＃１と＃２）には、同じアングルの映像をできるだけ選択する。このように構成することで、アングル切替の回数を減らすことができる。

[0214] なお、ユーザの嗜好情報として、「選手」や「ボール」といったものを指定するとしたが、例えば「チーム」という単位でもよい。ユーザの嗜好情報として好きなチームが設定されれば、自動映像選択編集部２９０２は、好きなチームの選手が多く写るように映像の選択を行う。このようにすることで、ユーザは映像コンテンツの視聴において、自分の嗜好を反映した視聴が実現できる。

[0215] なお、自動映像選択編集部２９０２は、映像データを選択するとしたが、広角映像からクロッピングして映像を生成するように、興味領域のみを切り出して利用してもよいのは言うまでもない。

[0216] なお、シーンの区切り方の例として、図３３で「ゲーム中」と「非ゲーム中」を取り上げたが、図３８に示すとおり、「ゲーム中」において、攻守交替地点をシーンの区切りとしてもよい。この場合には、例えば、Ａチームの攻撃中のシーンにおいての、検出アルゴリズムは、「（１）ボールに触っている選手がＡチームの選手に変わった。」、「（２）Ｂチームの陣地内の選手の数が過半数を超える。」、および「（３）選手の移動ベクトルの過半数がＢチーム側。」が考えられる。（１）においては、ボール位置情報の近接する選手位置情報の選手が、Ａチームかどうかで判断できる。（２）においては、Ｂチームの陣地領域内に存在する位置情報を持つ選手の数を数えることで判断できる。（３）においては、選手位置情報における移動ベクトルを計算することで判断できる。このように「攻守交替」でシーンを分離することにより、例えば、攻撃側の選手が写る映像を選択することで、基本的にス

ポーツは攻撃側に注目が集まるため、多くのユーザの注目する選手がよく写る映像をユーザに提供することが可能となる。

[0217] なお、本実施の形態におけるコンテンツの例として、スポーツイベントを取り上げたが、それ以外のイベントに応用できることはいうまでもない。例えば、図３９に示すような、コンサート映像であってもよい。複数人のアイドルグループのコンサートの場合には、ファンは、グループのメンバー全てが写る映像が必ずしも見たいわけではなく、グループの特定のメンバーにフォーカスした映像が見たいという強い要望がある。このようなケースにおいても、本実施の形態における個人嗜好を反映した映像コンテンツの自動生成・視聴システムは有効であり、スポーツの時と同じ構成で実現できる。コンサートにおけるシーンの区切り方について説明する。コンサートの構成はまず大きく２つのシーンに分離できる。「演奏中」と「非演奏中」である。つまり、シーン分割部２９０２ａは、撮影空間内で行われているイベントの種類がコンサートである場合には、「演奏中」であるか、「非演奏中」であるかを所定のアルゴリズムで判定することにより、クロッピング映像と複数のスポット映像とのそれぞれを、複数のシーンに分割してもよい。

[0218] この場合、「演奏中」は、アイドルグループにおいては歌やダンスをしている時間を示す。「非演奏中」は、アイドルグループにおいては歌やダンスをしている以外の時間を示す。演奏以外においては、ＭＣと呼ばれ、アイドルグループであれば、グループ同士で会話をしたり、観客に話しかけをおこなったりする。「演奏中」と「非演奏中」の判別は、映像撮影部によって収音される音声を解析することよってなされる。音声解析は、例えば、デジタル化された音に対して音程を計測し、その音程の特徴から判断する方式である。あらかじめコンサートを行う楽曲の音程のデータベースに登録しておき、収音される音声の音程とデータベース上の音程とのマッチングを取ることで、現在「演奏中」なのか「非演奏中」なのかを判定することが可能となる。こうした音声解析技術は、例えば、ソニー株式会社の「１２音解析」などがよく知られている。

[0219] その他に「演奏中」「非演奏中」を分離する情報としては、照明や観客の声がある。照明については画像を解析し輝度値を計測することで、シーンの分離の補助として活用できる。また、観客の声については、音の大きさを計測することで、シーンの分離の補助として活用できる。このようにすることで、「演奏中」「非演奏中」のシーンを分離することで、そのシーンにあった映像選択を実現できる。例えば、「非演奏中」においては、トークを行うことが多いため、実際に声を発生している人物を特定し、その人物にフォーカスした映像に切り替えるとしてもよい。声を発生している人物の特定は、各人のマイクの音量を計測することで実現できる。このように構成することで、ユーザは実際に話している人の声と映像とを同時に視聴できるため、わかりやすい映像をユーザに提供できる。

[0220] また、「非演奏中」においては、トークを音声解析し、その内容を画面上に字幕としてオーバレイ表示してもよい。このようにすることで、ユーザに分かりやすく話の内容を提示できる。なお、「演奏中」においては、音声解析を更に活用して、データベース上の楽曲の音程とマッチングさせて、「イントロ」「Ａメロ」「Ｂメロ」「サビ」「間奏」といった単位でシーンを分離してもよい。このように構成して、シーン毎に映像を切り替えることで、ユーザに飽きさせない迫力ある映像コンテンツの提供が可能となる。なお、演奏中のシーン分離としては、歌詞の単位（例えば、１行ごとなど）で分離しても同様の効果が得られる。

[0221] なお、本実施の形態におけるコンテンツの例として、スポーツイベントを取り上げたが、それ以外のイベントに応用できることはいうまでもない。例えば、図４０に示すような、討論会や会議のイベント撮影であってもよい。この場合は、図４０の例では、スポット映像撮影部は、各討論者に用意され、広角映像を生成するための映像撮影部が別途用意される。このような討論会や会議におけるシーンの区切り方について説明する。討論会や会議におけるシーンの区切りは、話者の交代である。つまり、シーン分割部２９０２ａは、撮影空間内で行われているイベントの種類が討論会である場合には、当

該討論会に参加している複数の参加者のうちで当該討論会における話者の交代を所定のアルゴリズムで判定することにより、クロッピング映像と複数のスポット映像とのそれぞれを、複数のシーンに分割してもよい。

[0222] 例えば、Ａさん、Ｂさん、Ｃさんによる討論が行われる場合には、Ａさんが話すシーン、Ｂさんが話すシーン、およびＣさんが話すシーンが交互に発生することになる。話者の切替を判定するには、各人のマイクの音量の計測結果、話者の声の音程の特徴量などをあらかじめデータベースに登録しておく。そして、現在発生される声に対してデータベースとマッチングを取ることで、誰が現在話しているのかを特定できる。または、画像認識で人物の口の動きの検出することにより現在の話者を特定してもよい。このようにして、話者ごとにシーンを区切り、自動映像選択編集部２９０２は、シーンの話者にフォーカスした映像を選択するとすれば、ユーザは実際に話している人の声と映像を同時に視聴できる。このため、わかりやすい映像をユーザに提供できる。

[0223] なお、シーンの区切り方の例として、図３３で「ゲーム中」と「非ゲーム中」を取り上げたが、例えば、フィギュアスケートであれば、技ごとにシーンを区切るようにしてもよい。このようにして、シーン毎に映像を変えることで、ユーザに飽きさせない迫力ある映像コンテンツの提供が可能となる。例えば、選手位置情報において人物の骨格情報が取得すれば、その骨格の動きのパターンから技の切れ目を判断することができる。

[0224] なお、オペレータが広角映像等の映像を視聴しながら手動でシーンを設定できるようにしてもよい。例えば、オペレータは、タブレット等のディスプレイデバイス上に「ゲーム中」および「非ゲーム中」を示すボタンが用意され、広角映像を視聴しながら、ボタンを押すとシーンの設定が行われ、その情報を編集システムの自動映像選択編集部２９０２に通知する。このようにすることで、自動映像選択編集部は、位置情報の解析を行わなくてもシーンの設定が可能となる。

[0225] なお、生成したシーン情報は当然ながらユーザによる映像視聴のチャプタ

情報として利用してもよい。テレビ等にチャプタメニューとして表示され、ユーザがリモコン等で、チャプタ情報を選択した場合、そのチャプタ位置から再生を開始させれば、ユーザは興味の映像を簡単に検索して、簡単に再生することが可能となる。

[0226] なお、生成したシーン情報を利用して、編集システム２９２０はハイライト映像を生成してユーザに提供するように構成してもよい。例えば、Ａチームおよびチームの対戦において、ユーザがＡチームのファンであれば、「ゲーム中」でかつ「Ａチームが攻撃中」のシーンだけを選択し、映像ストリームを生成して、ユーザに提供してもよい。このようにシーン情報を使ってハイライト映像を生成すれば、ユーザは興味のあるシーンのみを短い時間で再生することが可能となる。

[0227] なお、生成したシーン情報を「非ゲーム中」等のシーン区間で再生するリプレイ映像の再生に利用してもよい。例えば、攻守交替の地点からリプレイ映像を行うとする。このように構成すれば、鍵となるシーンの映像をユーザはリプレイ映像として視聴できる。

[0228] なお、本実施の形態１および２における配信・視聴システム１４００、２９００では、入力Ｉ／Ｆとしてリモコンを使って説明したが、図４１のタブレット型端末４１０１に示すように、タブレットやスマートフォンのようなディスプレイ付きのデバイス（情報端末）としてもよい。タブレット型端末上には、ボールや選手の名前等のボタン（アイコン）が配置されており、ユーザがボタンを選択すると、その選択情報がユーザの嗜好情報として編集システム１４２０、２９２０に伝送されるようにしてもよい。つまり、情報取得部１４２５は、コンピュータにネットワークを介して接続される情報端末に対して、ユーザにより入力されたユーザの嗜好情報を、ネットワークを経由して取得してもよい。編集システム１４２０、２９２０は、ユーザの嗜好情報のターゲットにフォーカスする広角映像のクロッピングやスポット映像撮影部の映像から映像生成して、再生システム１４３０に伝送し、テレビ等に表示する。この際に、編集システム１４２０、２９２０は、タブレット型

端末４１０１の解像度に合わせた広角映像のストリームを生成し、タブレット型端末４１０１上で再生し、現在どこがクロッピングされているのかの矩形領域を合成して表示してもよい。このように構成することで、ユーザは分かりやすい操作で快適に視聴ができる。

[0229] なお、図４１のタブレット型端末４１０１に示した矩形領域をユーザがピンチイン／ピンチアウトやスライド操作などで動かすと、その位置や矩形サイズ情報を編集システム１４２０、２９２０に通知し、編集システム１４２０、２９２０は、その位置や矩形サイズの領域広角映像からクロッピングして、ストリームを生成して、視聴システムに伝送し、テレビ等で表示を行ってもよい。このようにすることで、ユーザはよりダイレクトに自分の見たい視点を操作することが可能である。なお、スライドバーをタブレット上に配置して、スライドバーを移動させることで、クロッピング領域の位置を変更するようにしても同様の効果が得られる。

[0230] また、本実施の形態１および２における配信・視聴システム１４００、２９００では、人物やボール等の「名前」のボタンを用意して、それを選択することで、ユーザの嗜好情報を反映させるとしたが、図４２に示すように、「名前」ではなく、被写体位置情報を用いて、広角映像から切り抜いた「静止画」のボタンをタブレット型端末４１０１等の情報端末に表示して、選択するように構成してもよい。被写体位置情報には被写体ごとにＩＤが付与されており、タブレット型端末４１０１には、静止画とＩＤのペアで情報が伝送される。ユーザが静止画を選択すると、その静止画に対応するＩＤがユーザの嗜好情報として、編集システムに伝送される。編集システムは、該当ＩＤの被写体位置情報を利用して、広角映像のクロッピングやスポット映像の選択を行い、映像コンテンツを生成して、視聴システムに伝送し、テレビ等に表示する。このように構成することで、広角映像上の被写体（人物等）と名前の関連付けが取れない場合であっても、ユーザはユーザの嗜好を反映した映像を再生できる。

[0231] また、本実施の形態１および２における配信・視聴システム１４００、２

９００では、被写体位置情報を利用して広角映像からクロッピングを行い生成した映像を受信し、ユーザに提示するとしたが、図４３に示すように、広角映像を受信して、複数並べられたテレビに表示する構成であってもよい。図４３の構成では、ＳＴＢはインターネットと接続しており、ネットワーク上の撮影システムから広角映像を受け取り、ＳＴＢは、広角映像の左半分の映像をクロッピングしてＨＤＭＩ（登録商標）１からテレビ１に出力し、広角映像の右半分の映像をクロッピングしてＨＤＭＩ（登録商標）２からテレビ２に出力する。このようにテレビを連結させることで、ユーザは超ワイドの映像を視聴することができる。なお、それぞれのＨＤＭＩ（登録商標）から出力する映像の選択は、ユーザがＧＵＩ等によって設定できるようにしてもよい。

[0232] また、図４４に示すように、それぞれのテレビがネットワークに接続されており、左側に設置されるテレビ１は撮影システムから広角映像の左側の映像ストリームを受信し、右側に設置されるテレビ２は撮影システムから広角映像から右側の映像ストリームを受信し、表示する構成にしてもよい。同期を取るために、テレビ１とテレビ２の実時刻情報はＮＴＰサーバ等で同期しており、各映像ストリームの中には、各フレームに表示する実時刻情報が補足データ等に付加されており、各テレビはその実時刻情報に該当フレームを表示するように構成することで、複数のテレビ間の同期を取ることが可能となる。このように構成することで、広角映像を複数のテレビで表示することが可能となると共に、図４３のＳＴＢが不要となる。

[0233] また、本実施の形態１および２における配信・視聴システム１４００、２９００では、ユーザの嗜好情報として、好みの選手や見かた等を設定するとしたが、ここに「設定再生時間」を指定するとしてもよい。例えば、１２０分の試合を１０分の再生時間で見たいといった要望である。この場合には、編集システムは、ユーザの嗜好情報（好みの選手）や被写体位置情報、シーン情報を元に重み付けを行い、設定再生時間の範囲に収まるようにハイライトシーンを生成して映像を生成してユーザに提供する。このように構成する

ことで、ユーザは、設定時間範囲内で、興味の部分のみを抽出された映像を視聴できる。図４５でスポーツによって自動編集方法の例を具体的に説明する。図４５の例は、サッカーの試合の例である。まず、編集システムは、図４５の（ａ）に示す非ゲーム中のシーンのうち、ゴール中断以外の「非ゲーム中」を削除する。結果を図４５の（ｂ）に示している。非ゲーム中に対する関心は少ないため、ここのシーンをカットすることでユーザの興味領域を抽出できる。次に、編集システムは、図４５の（ｂ）に示す攻守交替のシーン内、自分の好みのチームの攻撃側のシーンを抽出する。ユーザの好みがＢチームである場合の結果を図４５の（ｃ）に示している。一般にユーザは、好みのチームの攻撃側のシーンを好んでみるため、ユーザにとっての興味領域を抽出できる。次に、編集システムは、図４５の（ｃ）に示すユーザの好みの選手がボール位置情報に近いシーンを抽出する。ユーザの好みがＸ選手である場合の結果を図４５の（ｄ）に示している。一般にユーザはボール位置が興味領域となるため、好みの選手が、ボール位置情報に近いシーンを抽出することで、ユーザにとっての興味領域を抽出できる。このようにして、ユーザの興味を反映した短い時間のハイライト再生コンテンツを生成することが可能となる。図４５は一例であり、例えば、シーンの抽出順を代えてもよいし、１つだけでもよい。

[0234] なお、本実施の形態１および２における配信・視聴システム１４００、２９００では、位置特定部１４２２は、広角映像に対して画像認識処理を行うことで、被写体位置情報を特定するとしたが、映像生成部１４２３によって生成された広角映像だけでなく、スポット映像撮影部２９０１によって撮影された複数のスポット映像や、図４６に示すように広角映像とは別の視点から撮影される第２広角映像を、画像認識処理に利用してもよい。画像認識処理における人物認識の課題の１つとして、画像解像度が上げられる。図４６の例の場合、コート下部（手前側）に人物がいれば、カメラ１５０１～１５０３が人物を大きく撮影できるので、顔等を認識することができるが、人物が上部（コートの奥側）に移動すればカメラ１５０１～１５０３の撮影では

人物を大きくできず、顔認識等が困難になる可能性がある。そこで、カメラ４４０１、４４０２、４４０３のように反対側から撮影する手段を設ければ、コート上部に移動する人物も大きく撮影できる。そこで、人物位置が上部にいる場合にはカメラ４４０１、４４０２、４４０３の広角映像による解析結果を優先、人物位置が下部にいる場合にはカメラ１５０１～１５０３の広角映像による解析結果を優先するとすれば、人物認識の解析結果の精度を向上できる。

[0235] また、例えば顔で人物認識を行う場合には、人物が図４６のコート下方向を向いていれば顔を認識できるが、反対を向いている場合は認識できない。そこで、カメラ４４０１、４４０２、４４０３のように反対側から撮影する手段を設ければ、人物がコート上方向（コート奥側の方向）を向いていても顔認識を行うことができる。同様に、例えば背中にある背番号で人物認識を行う場合には、人物が図４６のコート上方向を向いていればユニフォームの番号認識できるが、反対を向いている場合は認識できない。そこで、カメラ４４０１、４４０２、４４０３のように反対側から撮影する手段を設ければ、人物がコート下方向（コート手前側の方向）を向いていてもユニフォームの番号認識を行うことができる。そこで、複数の視点の映像に対して同時に画像認識をかけことで、どこかの映像で認識できれば、その情報を被写体位置情報として利用する。このように構成することで、被写体位置情報の精度を向上できる。

[0236] また、本実施の形態１および２における配信・視聴システム１４００、２９００の一形態として、個人コンテンツの活用があり、例えば、自分の子供の幼稚園・小学校のお遊会の撮影や運動会、各種スポーツイベントなどが考えられる。このようなケースでは、親は自分の子供を中心に視聴するといった視聴が要望される。この際に、編集システム１４２０、２９２０の位置特定部１４２２は、子供の顔認識データベースを保持しておき、その顔認識データベースを使って、人物を特定・トラッキングし、自動映像選択編集部は、ユーザの嗜好情報で指定される子供の位置情報に従い、広角映像からのク

ロッピングや、スポット映像撮影部の映像に切り替えることで、この要望にこたえることが可能となる。ここで、顔認識データベースは、図４７に示すように、ユーザが撮影した個人コンテンツから年代毎に管理するようにしてもよい。図４７の矢印は撮影時期の時間軸を示し、動画や写真の個人コンテンツを白四角で示している。この場合に、顔認識用のデータベースは、動画や写真の撮影時期に応じて生成される。図４７の例では、顔認識データベース１（顔ＤＢ１）は、２０１２－２０１３年に撮影される動画や写真によって生成され、顔認識データベース２（顔ＤＢ２）は、２０１３－２０１４年に撮影される動画や写真によって生成され、顔認識データベース３（顔ＤＢ３）は、２０１４－２０１５年に撮影される動画や写真によって生成され、顔認識データベース４（顔ＤＢ４）は、２０１５－２０１６年に撮影される動画や写真によって生成される。ここで、例えば、撮影時期が２０１３年１０月の映像コンテンツが追加される場合には、その撮影時期に対応するデータベース、図４７の場合では顔ＤＢ２を参照して、顔認識を行う。このように年代ごとで顔認識データベースを変えることで、成長する人物の顔であっても、位置特定部１４２２は、精度良く、顔認識をして人物の特定を行うことができる。

[0237] また、本実施の形態１および２においては、位置特定部１４２２は、顔認識データベースを利用して被写体位置情報の生成を行うとしたが、この被写体位置情報の生成では、１度だけでなく、時間を置いた後に追加で行うようにしてもよい。つまり、顔認識データベースは、写真が増えれば、増えるほど学習により精度が上がっていくため、時間を置き個人コンテンツが増えていくことで、顔認識データベースの精度は向上する。このように構成することで、ユーザは最初に見たときは、自分の子供への追従が悪かったとしても、次に時間を置いてみるときには、追従精度があがるということになる。

[0238] また、本実施の形態２における配信・視聴システム２９００では、スポット映像撮影部２９０１は、固定カメラで撮影するとしたが、スポット映像撮影部２９０１をパンチルトズーム可能なＰＴＺカメラ等で構成し、編集シス

テム２９２０の位置特定部１４２２の解析をリアルタイム化し、その被写体位置情報を撮影制御部１４０１に伝送し、撮影制御部１４０１が、人物やボールの周辺をフォーカスするようにスポット映像撮影部２９０１のパンチルトズームを制御してもよい。このように構成することで、より迫力ある映像を撮影し、ユーザに提供することができる。

[0239] なお、本実施の形態１および２における配信・視聴システム１４００、２９００では、編集システム１４２０、２９２０は、個人嗜好にあった映像ストリームを生成しユーザに提供するとしたが、図４８に示すように映像ストリームではなく、電子書籍、特にコミックとして提供してもよい。これを実現するためには、自動映像編集部１４２４または自動映像選択編集部２９０２は、シーン情報、被写体位置情報、ユーザの嗜好情報、から、代表静止画リストを生成する。自動映像編集部１４２４または自動映像選択編集部２９０２は、代表静止画情報から漫画のコマワリを行い、静止画を挿入する。このときに試合経過情報等を図４８のナレーションとして提示してもよい。また、図４８の（ａ）の最後のコマのように、被写体位置情報からシュートシーンを判別した場合には、イフェクトや擬声語（オノマトペ）を合成してもよい。また、会議等のイベントを撮影する場合には、会話を文字情報化して、図４８の（ｂ）のように人物位置の横に吹き出しで文字情報を合成してもよい。このように構成することで、ユーザはイベントを映像だけでなくコミックという別形態で楽しむことが可能となる。

[0240] なお、上記各実施の形態において、各構成要素は、専用のハードウェアで構成されるか、各構成要素に適したソフトウェアプログラムを実行することによって実現されてもよい。各構成要素は、ＣＰＵまたはプロセッサなどのプログラム実行部が、ハードディスクまたは半導体メモリなどの記録媒体に記録されたソフトウェアプログラムを読み出して実行することによって実現されてもよい。ここで、上記各実施の形態の画像復号化装置などを実現するソフトウェアは、次のようなプログラムである。

[0241] すなわち、このプログラムは、コンピュータに、コンピュータを用いて、

ユーザの嗜好情報に基づいて編集された映像を提供する、映像提供方法であって、（ｉ）撮影空間のうちの一部である第一撮影空間が撮影された第一主映像と、（ｉｉ）前記撮影空間のうちの空間であって、前記第一空間以外の空間を含む第二撮影空間が撮影された第二主映像と、を取得する映像取得ステップと、前記映像取得ステップにおいて取得された前記第一主映像および前記第二主映像を合成することにより広角映像を生成する映像生成ステップと、ネットワークを経由して、前記ユーザの嗜好情報を取得する情報取得ステップと、前記情報取得ステップにおいて取得された前記ユーザの嗜好情報に基づいて、前記広角映像のうちの領域であって、前記広角映像の領域よりも小さいクロッピング領域を算出する領域算出ステップと、前記映像生成ステップにおいて生成された前記広角映像を、前記領域算出ステップにおいて算出された前記クロッピング領域でクロッピングするクロッピングステップと、前記クロッピングステップにおいてクロッピングされることにより生成されたクロッピング映像をユーザに提供する映像提供ステップと、を含む映像提供方法を実行させる。

[0242] 以上、本発明の一つまたは複数の態様に係る・・・装置について、実施の形態に基づいて説明したが、本発明は、この実施の形態に限定されるものではない。本発明の趣旨を逸脱しない限り、当業者が思いつく各種変形を本実施の形態に施したものや、異なる実施の形態における構成要素を組み合わせて構築される形態も、本発明の一つまたは複数の態様の範囲内に含まれてもよい。

[0243] なお、本発明を上記実施の形態に基づいて説明してきたが、本発明は、上記の実施の形態に限定されないのはもちろんである。以下のような場合も本発明に含まれる。

[0244] （１）上記の各装置は、具体的には、マイクロプロセッサ、ＲＯＭ、ＲＡＭ、ハードディスクユニット、ディスプレイユニット、キーボード、マウスなどから構成されるコンピュータシステムで実現され得る。ＲＡＭまたはハードディスクユニットには、コンピュータプログラムが記憶されている。マ

イクロプロセッサが、コンピュータプログラムにしたがって動作することにより、各装置は、その機能を達成する。ここでコンピュータプログラムは、所定の機能を達成するために、コンピュータに対する指令を示す命令コードが複数個組み合わされて構成されたものである。

[0245] （２）上記の各装置を構成する構成要素の一部または全部は、１個のシステムＬＳＩ（Ｌａｒｇｅ　Ｓｃａｌｅ　Ｉｎｔｅｇｒａｔｉｏｎ：大規模集積回路）から構成されているとしてもよい。システムＬＳＩは、複数の構成部を１個のチップ上に集積して製造された超多機能ＬＳＩであり、具体的には、マイクロプロセッサ、ＲＯＭ、ＲＡＭなどを含んで構成されるコンピュータシステムである。ＲＯＭには、コンピュータプログラムが記憶されている。マイクロプロセッサが、ＲＯＭからＲＡＭにコンピュータプログラムをロードし、ロードしたコンピュータプログラムにしたがって演算等の動作することにより、システムＬＳＩは、その機能を達成する。

[0246] （３）上記の各装置を構成する構成要素の一部または全部は、各装置に脱着可能なＩＣカードまたは単体のモジュールから構成されてもよい。ＩＣカードまたはモジュールは、マイクロプロセッサ、ＲＯＭ、ＲＡＭなどから構成されるコンピュータシステムである。ＩＣカードまたはモジュールには、上記の超多機能ＬＳＩが含まれてもよい。マイクロプロセッサが、コンピュータプログラムにしたがって動作することにより、ＩＣカードまたはモジュールは、その機能を達成する。このＩＣカードまたはこのモジュールは、耐タンパ性を有してもよい。

[0247] （４）本発明は、上記に示す方法で実現されてもよい。また、これらの方法をコンピュータにより実現するコンピュータプログラムで実現してもよいし、コンピュータプログラムからなるデジタル信号で実現してもよい。

[0248] また、本発明は、コンピュータプログラムまたはデジタル信号をコンピュータ読み取り可能な記録媒体、例えば、フレキシブルディスク、ハードディスク、ＣＤ－ＲＯＭ、ＭＯ、ＤＶＤ、ＤＶＤ－ＲＯＭ、ＤＶＤ－ＲＡＭ、ＢＤ（Ｂｌｕ－ｒａｙ（登録商標）　Ｄｉｓｃ）、半導体メモリなどに記録し

たもので実現してもよい。また、これらの記録媒体に記録されているデジタル信号で実現してもよい。

[0249] また、本発明は、コンピュータプログラムまたはデジタル信号を、電気通信回線、無線または有線通信回線、インターネットを代表とするネットワーク、データ放送等を経由して伝送してもよい。

[0250] また、本発明は、マイクロプロセッサとメモリを備えたコンピュータシステムであって、メモリは、コンピュータプログラムを記憶しており、マイクロプロセッサは、コンピュータプログラムにしたがって動作してもよい。

[0251] また、プログラムまたはデジタル信号を記録媒体に記録して移送することにより、またはプログラムまたはデジタル信号をネットワーク等を経由して移送することにより、独立した他のコンピュータシステムにより実施するとしてもよい。

[0252] （５）上記実施の形態及び上記変形例をそれぞれ組み合わせるとしてもよい。

産業上の利用可能性

[0253] 本発明にかかるデータ作成装置を用いた映像コンテンツの配信・視聴システムは、従来放送局による配信では実現し得なかった、ユーザの嗜好情報を反映した新しい映像コンテンツの配信が可能となる。故に本発明は、ネット動画配信事業等の映像配信産業、テレビ等の民生機器産業において高い利用可能性を持つ。

符号の説明

[0254] １０ 配信・視聴システム

１００ 放送システム

１０１ 放送映像撮影部

１０２ 放送映像編集部

１０３ 放送ストリーム作成部

１０４ 放送ストリーム

１１０ 再生装置

１１１　チューナ

１１２　放送ストリームデコード部

２０１　リモコン

２０２　デジタルテレビ

５０１　ビデオストリーム

５０２　ＰＥＳパケット列

５０３　ＴＳパケット

５０４　オーディオストリーム

５０５　ＰＥＳパケット列

５０６　ＴＳパケット

５０７　字幕ストリーム

５０８　ＰＥＳパケット列

５０９　ＴＳパケット

５１３　トランスポートストリーム

１４００、２９００　配信・視聴システム

１４０１　撮影制御部

１４０２　映像撮影部

１４０３　通信Ｉ／Ｆ

１４１０　撮影システム

１４２０　編集システム

１４２１、１４２７　通信Ｉ／Ｆ

１４２２　位置特定部

１４２３　映像生成部

１４２４　自動映像編集部

１４２４ａ　領域算出部

１４２４ｂ　クロッピング部

１４２５　情報取得部

１４２６　映像提供部

１４３０　再生システム

１４３１　通信Ｉ／Ｆ

１４３２　ストリームデコード部

１４３３　入力Ｉ／Ｆ

１４３４　アプリケーション実行部

１５０１　第一カメラ

１５０２　第二カメラ

１５０３　第三カメラ

１５０４　タブレット型端末

１５１１　第一主映像

１５１２　第二主映像

１５１３　第三主映像

２３０１　同期制御部

２４０４　タブレット型端末

２９０１　スポット映像撮影部

２９０２　自動映像選択編集部

２９０２ａ　シーン分割部

２９０２ｂ　評価部

２９０２ｃ　映像選択部

２９１０　撮影システム

２９２０　編集システム

３００１～３００７　カメラ

４１０１　タブレット型端末

４４０１　カメラ

請求の範囲

［請求項1］ コンピュータが映像をユーザに提供するための映像提供方法であって、

（ｉ）撮影空間のうちの一部の第一撮影空間が撮影された第一主映像と、（ｉｉ）前記撮影空間のうちの一部の空間であって、前記第一空間以外の空間を含む第二撮影空間が撮影された第二主映像と、を取得する映像取得ステップと、

前記映像取得ステップにおいて取得された前記第一主映像および前記第二主映像を合成することにより広角映像を生成する映像生成ステップと、

ネットワークを経由して、前記ユーザの嗜好情報を取得する情報取得ステップと、

前記情報取得ステップにおいて取得された前記ユーザの嗜好情報に基づいて、前記広角映像のうちの一部の領域であって、前記広角映像の領域よりも小さいクロッピング領域を算出する領域算出ステップと、

前記映像生成ステップにおいて生成された前記広角映像を、前記領域算出ステップにおいて算出された前記クロッピング領域でクロッピングするクロッピングステップと、

前記クロッピングステップにおいてクロッピングされることにより生成されたクロッピング映像を前記ユーザに提供する映像提供ステップと、を含む

映像提供方法。

［請求項2］ 前記ユーザの嗜好情報は、ユーザが視聴したい対象である視聴対象を示し、

前記映像提供方法は、さらに、

前記ユーザの嗜好情報に基づいて前記広角映像に対して画像認識を行うことで、前記広角映像における前記視聴対象の位置を特定する位

置特定ステップを含み、

前記領域算出ステップでは、前記広角映像のうち、前記位置特定ステップにおいて特定された前記視聴対象の位置を用いて、前記視聴対象が含まれる領域を、前記クロッピング領域として算出する

請求項１に記載の映像提供方法。

［請求項3］ 前記領域算出ステップでは、

前記広角映像のうち、前記視聴対象の位置を、前記広角映像をクロッピングするための予め定められたサイズのクロッピング枠における所定の基準位置に一致させた場合に前記クロッピング枠で特定される領域を前記クロッピング領域として算出する

請求項２に記載の映像提供方法。

［請求項4］ 前記領域算出ステップでは、

前記広角映像のうち、処理対象のフレームよりも所定時間以前のフレームでの前記視聴対象の位置が、前記クロッピング枠の前記所定の基準位置に一致させた場合に前記クロッピング枠で特定される領域を前記クロッピング領域として算出する

請求項３に記載の映像提供方法。

［請求項5］ 前記映像取得ステップでは、さらに、

前記第一主映像および前記第二主映像と同じタイミングで、前記撮影空間のうちの少なくとも一部の空間が前記第一主映像及び前記第二主映像とは異なる角度で撮影された副映像を取得し、

前記映像提供方法は、さらに、

前記クロッピングステップでクロッピングされた前記クロッピング映像と、前記映像取得ステップで取得された前記副映像とのそれぞれを、所定のアルゴリズムに基づいて複数のシーンに分割するシーン分割ステップと、

前記情報取得ステップにおいて取得された前記ユーザの嗜好情報に基づいて、前記複数のシーンのそれぞれについて、前記クロッピング

映像および前記副映像のいずれかを選択する映像選択ステップと、を含み、

前記映像提供ステップでは、前記映像選択ステップにおいて選択された前記クロッピング映像および前記副映像のいずれかをユーザに提供する

請求項１から４のいずれか１項に記載の映像提供方法。

[請求項6] 前記シーン分割ステップでは、

前記クロッピング映像および前記副映像のそれぞれを前記複数のシーンに分割するときに、前記所定のアルゴリズムとは別に、所定時間毎に分割する

請求項５に記載の映像提供方法。

[請求項7] 前記所定のアルゴリズムは、前記撮影空間内で行われているイベントの種類毎に異なる

請求項６に記載の映像提供方法。

[請求項8] 前記シーン分割ステップでは、

前記撮影空間内で行われているイベントの種類がスポーツである場合には、当該イベントの状態が「ゲーム中」であるか、「非ゲーム中」であるかを前記所定のアルゴリズムで判定することにより、判定結果が前記「ゲーム中」および前記「非ゲーム中」の一方から他方へ切り替わったタイミングで、前記クロッピング映像と前記副映像とのそれぞれを、複数のシーンに分割する

請求項７に記載の映像提供方法。

[請求項9] 前記映像選択ステップでは、前記撮影空間内で行われているイベントの種類がスポーツである場合、前記「ゲーム中」から前記「非ゲーム中」に切り替わったときに、当該「非ゲーム中」のシーンを選択する代わりに直前の「ゲーム中」のシーンの映像の中から選択する

請求項８に記載の映像提供方法。

[請求項10] 前記シーン分割ステップでは、

前記撮影空間内で行われているイベントの種類がコンサートである場合には、当該イベントの状態が「演奏中」であるか、「非演奏中」であるかを前記所定のアルゴリズムで判定することにより、前記クロッピング映像と前記副映像とのそれぞれを、複数のシーンに分割する

請求項７から９のいずれか１項に記載の映像提供方法。

[請求項11] 前記シーン分割ステップでは、

前記撮影空間内で行われているイベントの種類が討論会である場合には、当該討論会に参加している複数の参加者のうちで当該討論会における話者の交代を前記所定のアルゴリズムで判定することにより、前記クロッピング映像と前記副映像とのそれぞれを、複数のシーンに分割する

請求項７から１０のいずれか１項に記載の映像提供方法。

[請求項12] 前記映像提供方法は、さらに、

前記シーン分割ステップにおいて分割された前記複数のシーンのそれぞれを、前記情報取得ステップにおいて取得された前記ユーザの嗜好情報と、所定の評価指標とに基づいて評価する評価ステップを含み、

前記映像選択ステップでは、前記評価ステップにおいて評価された結果に基づいて、前記複数のシーンのそれぞれについて、前記クロッピング映像および前記副映像のいずれかを選択する

請求項５から１１のいずれか１項に記載の映像提供方法。

[請求項13] 前記所定の評価指標は、映像を撮影した複数のカメラのうちで、前記視聴対象が画角に含まれており、かつ、前記視聴対象への距離が近いカメラによって撮影された映像のシーンほど高く評価する指標を含む

請求項１２に記載の映像提供方法。

[請求項14] 前記所定の評価指標は、映像を撮影した複数のカメラのうちで、前記視聴対象が画角に含まれており、かつ、前記視聴対象との間にある

オブジェクトの数が少ないカメラによって撮影された映像のシーンほど高く評価する指標を含む

請求項１２または１３に記載の映像提供方法。

[請求項15] 前記所定の評価指標は、映像を撮影した複数のカメラのうちで、前記視聴対象が画角に含まれており、かつ、当該映像に映り込んでいる前記視聴対象の面積が大きいカメラによって撮影された映像のシーンほど高く評価する指標を含む

請求項１２から１４のいずれか１項に記載の映像提供方法。

[請求項16] 前記所定の評価指標は、

映像を撮影した複数のカメラのうちで、前記視聴対象が画角に含まれており、かつ、前記視聴対象への距離が近いカメラによって撮影された映像のシーンほど高く評価する第一指標と、

映像を撮影した複数のカメラのうちで、前記視聴対象が画角に含まれており、かつ、前記視聴対象との間にあるオブジェクトの数が少ないカメラによって撮影された映像のシーンほど高く評価する第二指標と、

映像を撮影した複数のカメラのうちで、前記視聴対象が画角に含まれており、かつ、当該映像に映り込んでいる前記視聴対象の面積が大きいカメラによって撮影された映像のシーンほど高く評価する第三指標と、のうちの２以上の指標を含み、

前記評価ステップでは、前記複数のシーンのそれぞれについて、当該シーンに対する前記２以上の指標により評価された複数の結果について、前記２以上に指標に関連付けられている予め定められた重み付けによって重み付加算された加算値に基づいて評価する

請求項１２に記載の映像提供方法。

[請求項17] 前記情報取得ステップでは、

前記コンピュータに前記ネットワークを介して接続される情報端末に対して、前記ユーザにより入力された前記ユーザの嗜好情報を、前

記ネットワークを経由して取得する

請求項１から１６のいずれか１項に記載の映像提供方法。

[請求項18] ユーザの嗜好情報に基づいて編集された映像を送信することで提供する送信装置であって、

（ｉ）撮影空間のうちの一部の第一撮影空間が撮影された第一主映像と、（ｉｉ）前記撮影空間のうちの一部の空間であって、前記第一空間以外の空間を含む第二撮影空間が撮影された第二主映像と、を取得する映像取得部と、

前記映像取得部により取得された前記第一主映像および前記第二主映像を合成することにより広角映像を生成する映像生成部と、

ネットワークを経由して、前記ユーザの嗜好情報を取得する情報取得部と、

前記情報取得部により取得された前記ユーザの嗜好情報に基づいて、前記広角映像のうちの一部の領域であって、前記広角映像の領域よりも小さいクロッピング領域を算出する領域算出部と、

前記映像生成部により生成された前記広角映像を、前記領域算出部により算出された前記クロッピング領域でクロッピングするクロッピング部と、

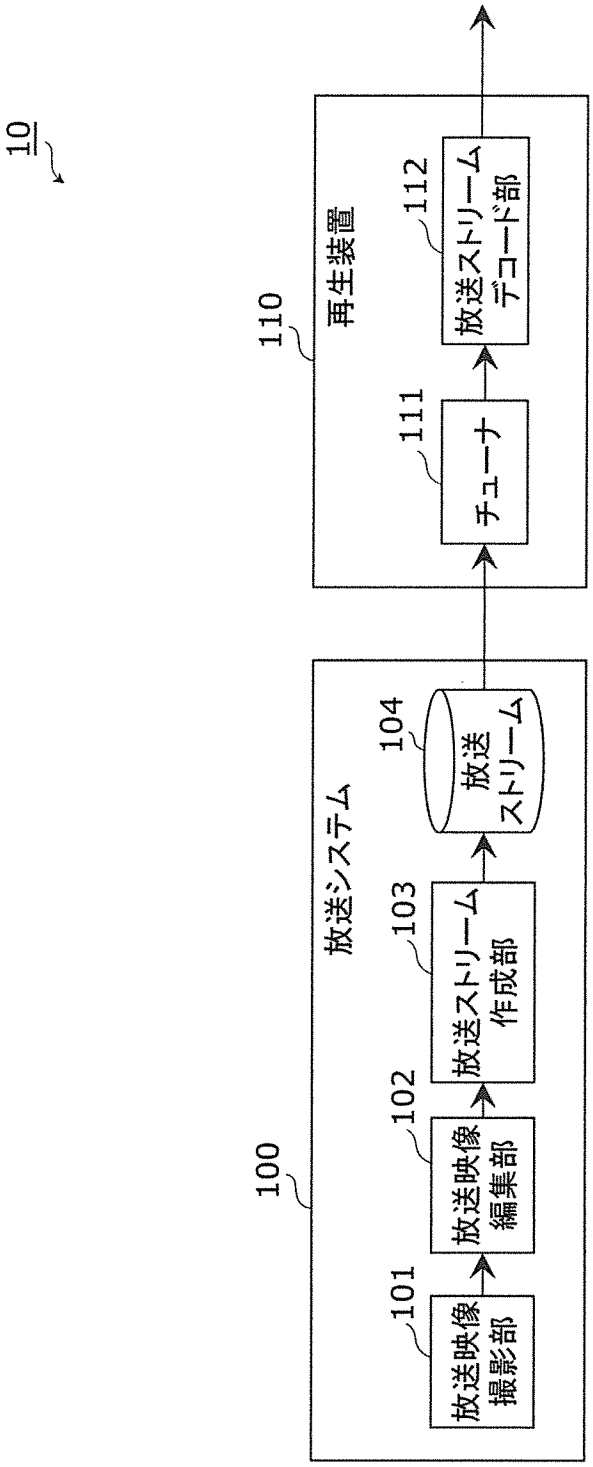
前記クロッピング部によりクロッピングされることにより生成されたクロッピング映像を前記ユーザに提供する映像提供部と、を備える

送信装置。

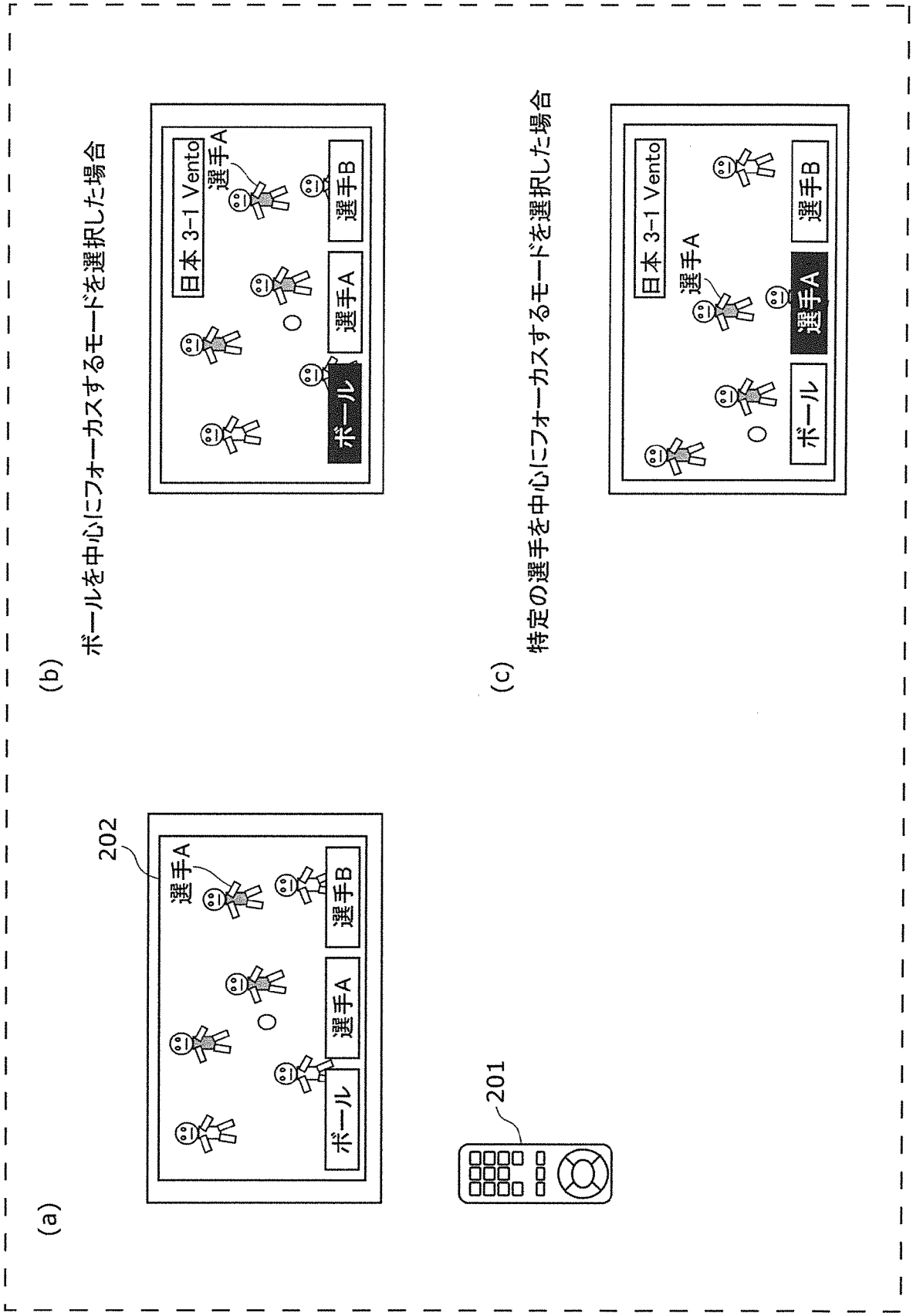
[請求項19] 請求項１８に記載の送信装置にネットワークを介して接続される受信装置であって、

前記送信装置から送信される前記映像を受信する受信装置。

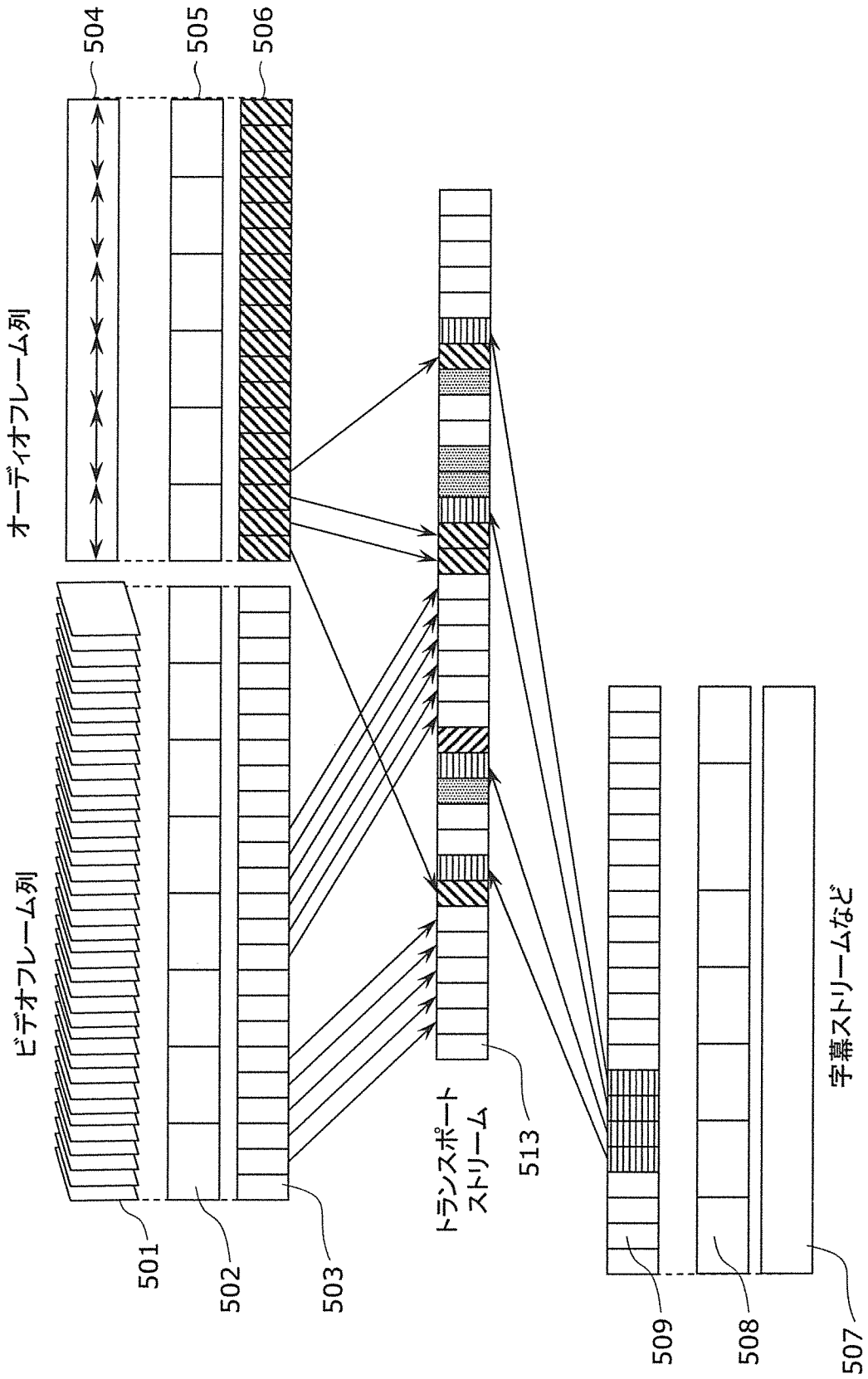
[図1]


10
100
放送システム
101
放送映像撮影部
102
放送映像編集部
103
放送ストリーム作成部
104
放送ストリーム
110
再生装置
111
チューナ
112
放送ストリームデコード部

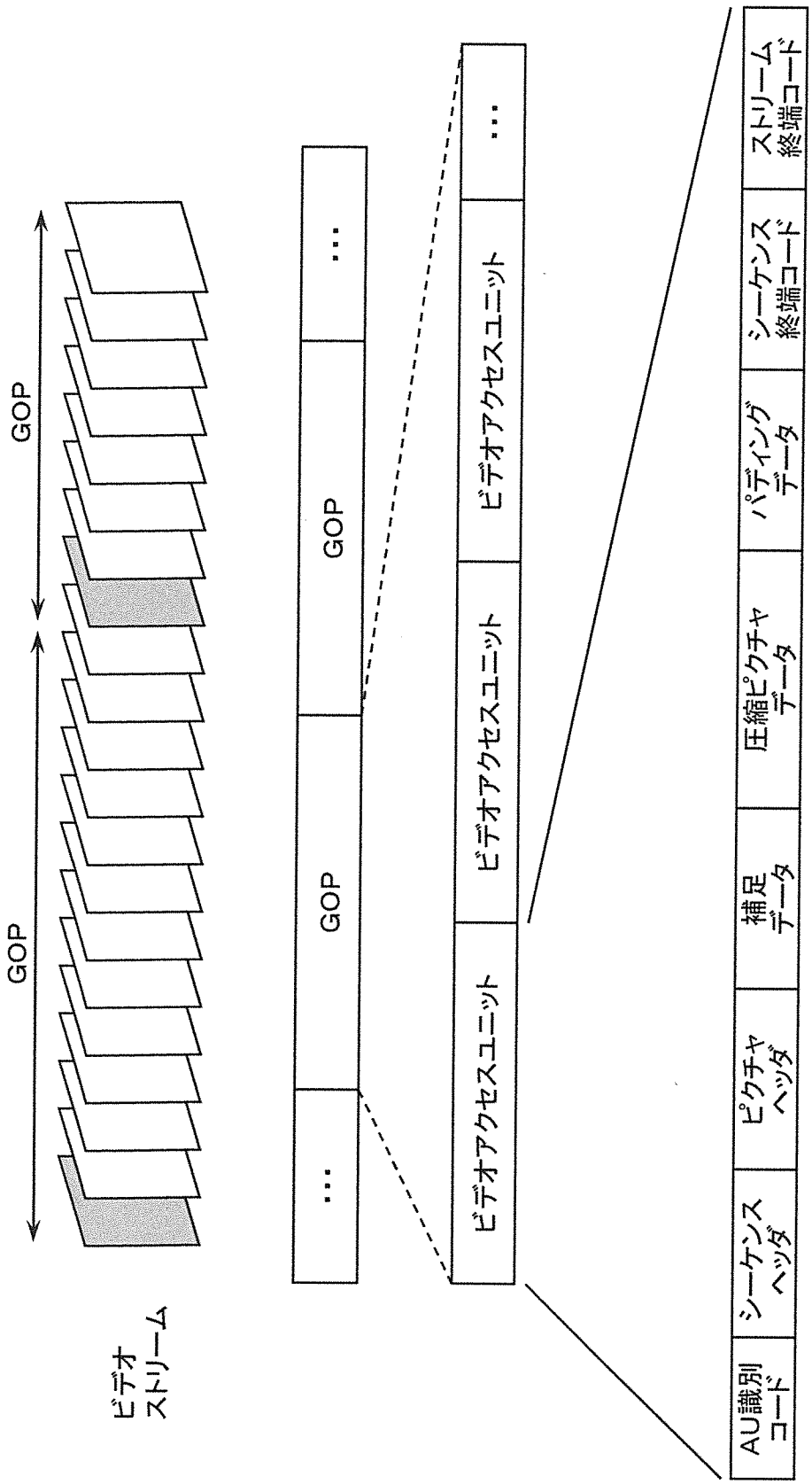
[図2]


(a)
202
選手A
ボール
選手A
選手B
201
(b)
ボールを中心にフォーカスするモードを選択した場合
日本 3-1 Vento
選手A
ボール
選手A
選手B
(c)
特定の選手を中心にフォーカスするモードを選択した場合
日本 3-1 Vento
選手A
ボール
選手A
選手B

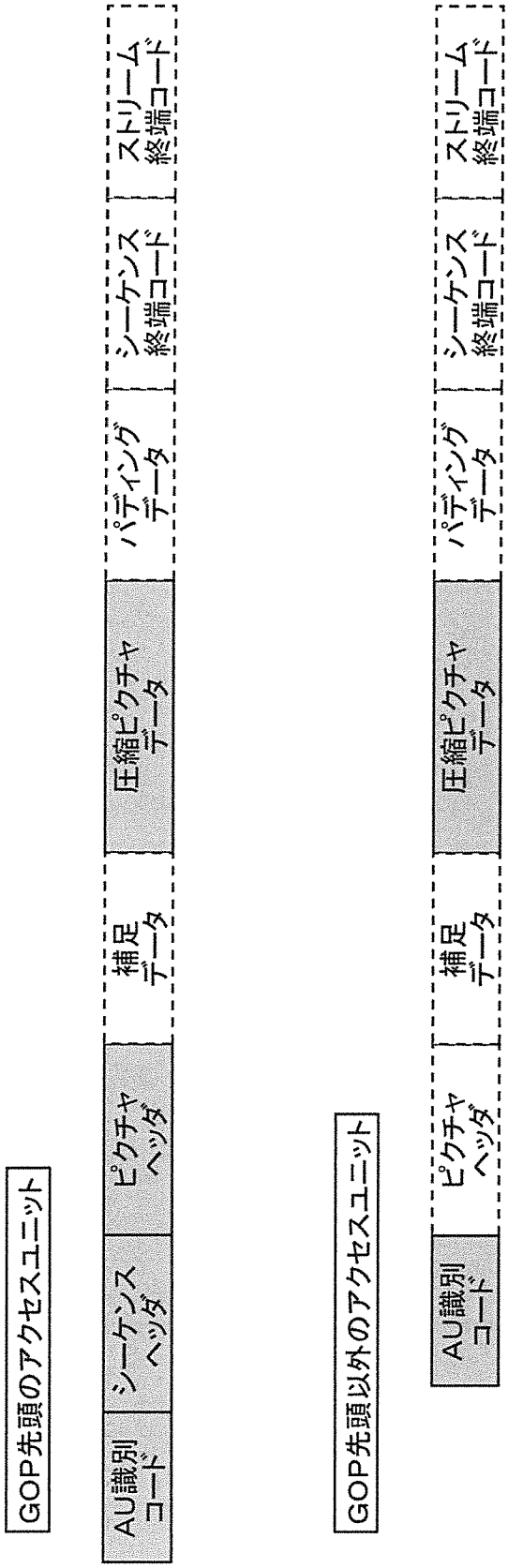
[図3]


504
505
506
オーディオフレーム列
ビデオフレーム列
501
502
503
トランスポートストリーム
513
509
508
507
字幕ストリームなど

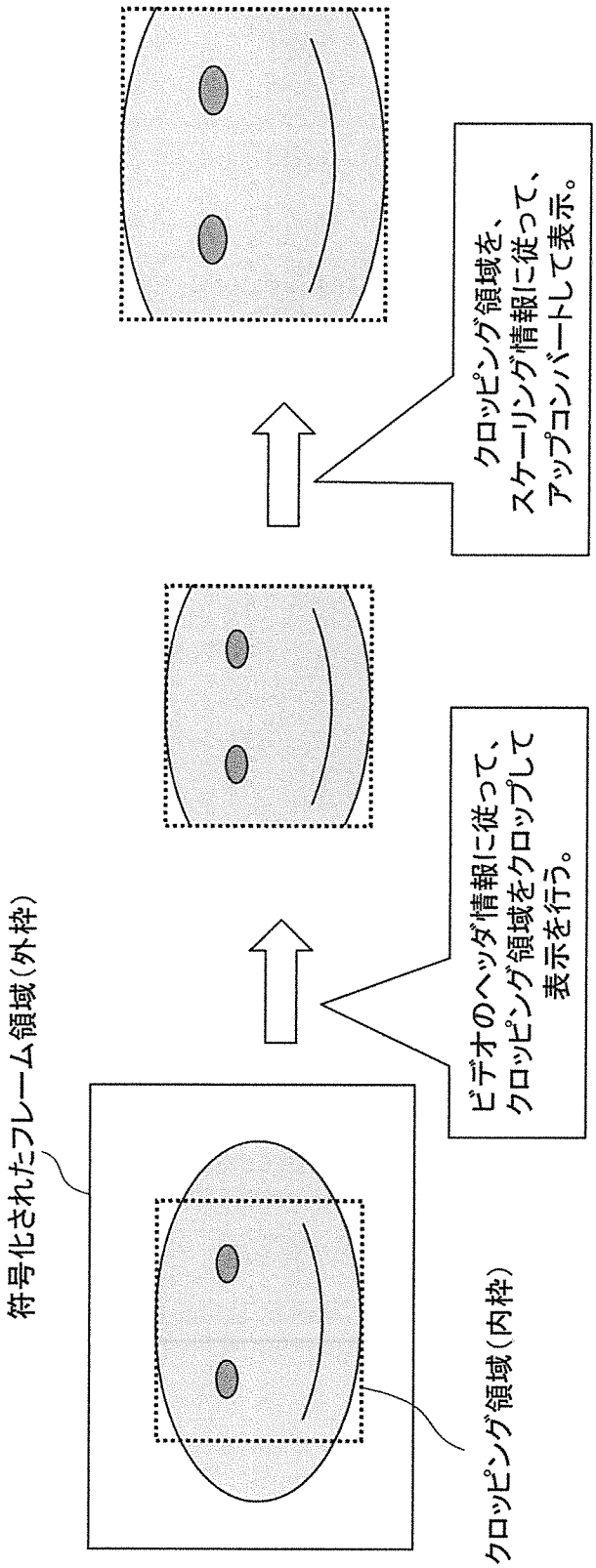
[図4]


ビデオストリーム
GOP
GOP
GOP
GOP
...
...
ビデオアクセスユニット
ビデオアクセスユニット
AU識別コード
シーケンスヘッダ
ピクチャヘッダ
補足データ
圧縮ピクチャデータ
パディングデータ
シーケンス終端コード
ストリーム終端コード

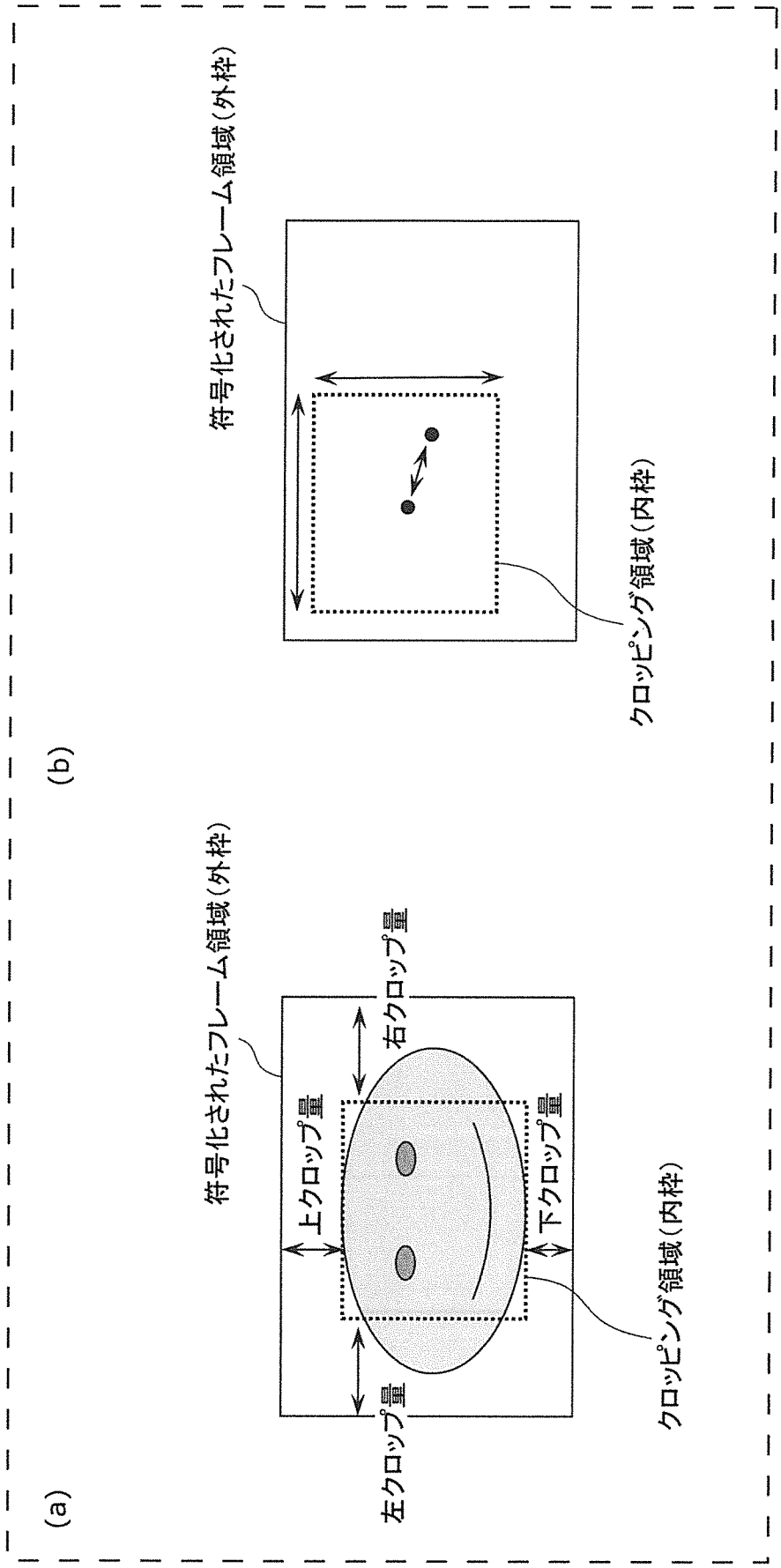
[図5]


GOP先頭のアクセスユニット
AU識別コード
シーケンスヘッダ
ピクチャヘッダ
補足データ
圧縮ピクチャデータ
パディングデータ
シーケンス終端コード
ストリーム終端コード
GOP先頭以外のアクセスユニット
AU識別コード
ピクチャヘッダ
補足データ
圧縮ピクチャデータ
パディングデータ
シーケンス終端コード
ストリーム終端コード

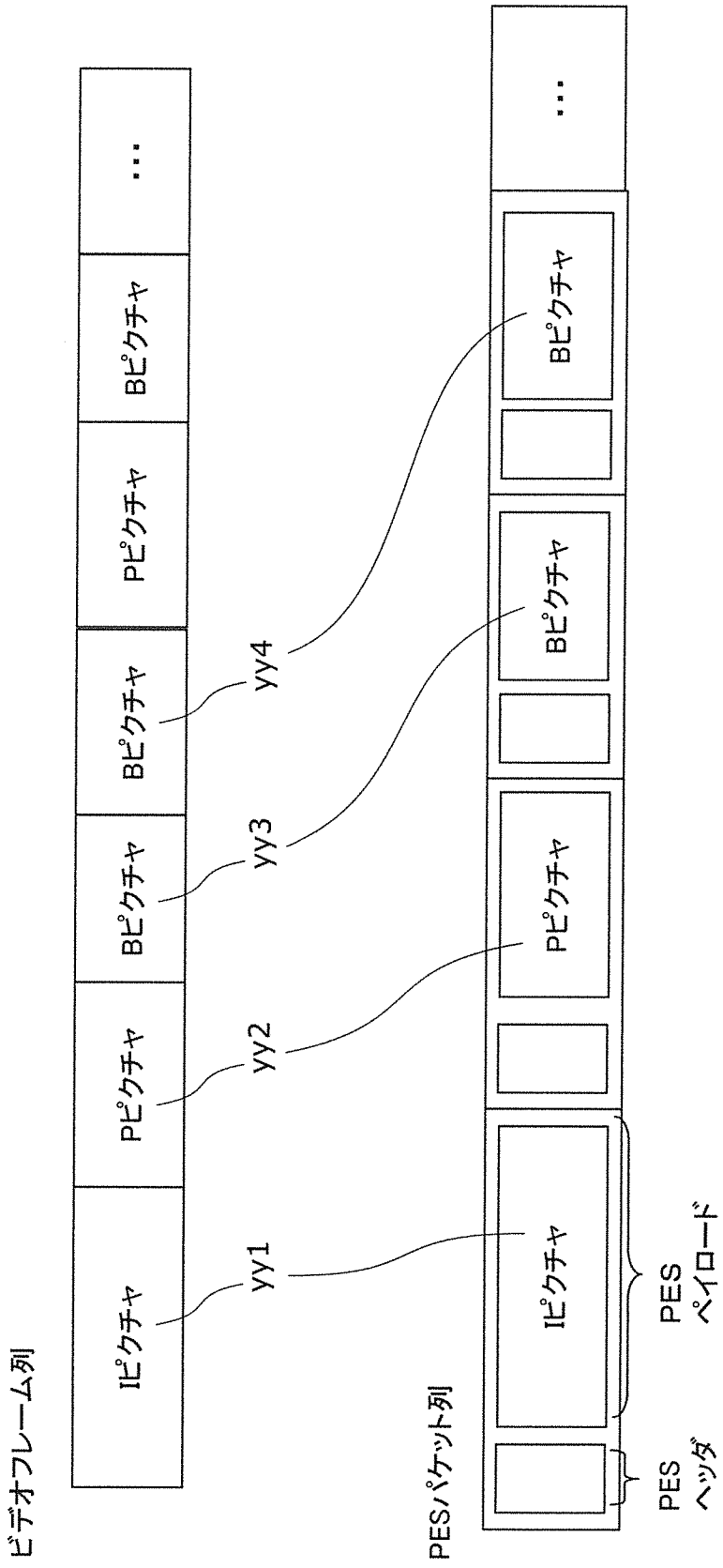
[図6]


符号化されたフレーム領域（外枠）
クロッピング領域（内枠）
ビデオのヘッダ情報に従って、
クロッピング領域をクロップして
表示を行う。
クロッピング領域を、
スケーリング情報に従って、
アップコンバートして表示。

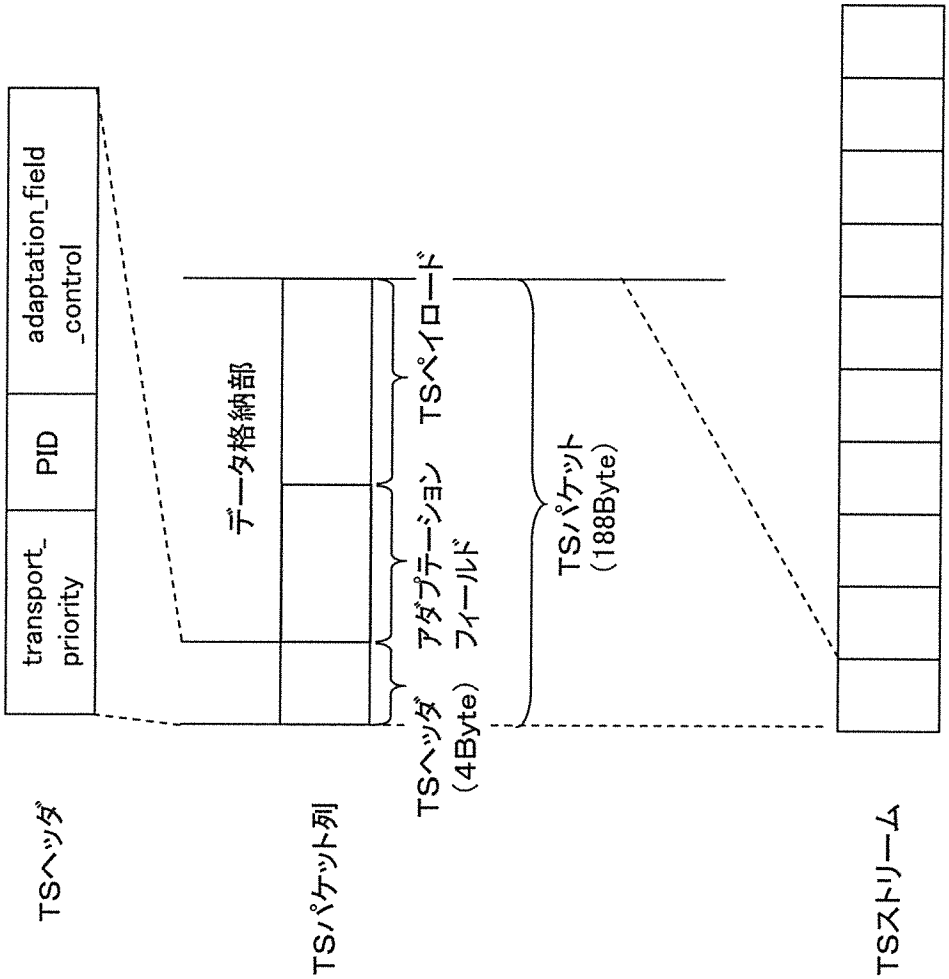
[図7]


(a)
符号化されたフレーム領域(外枠)
上クロップ量
右クロップ量
左クロップ量
下クロップ量
クロッピング領域(内枠)
(b)
符号化されたフレーム領域(外枠)
クロッピング領域(内枠)

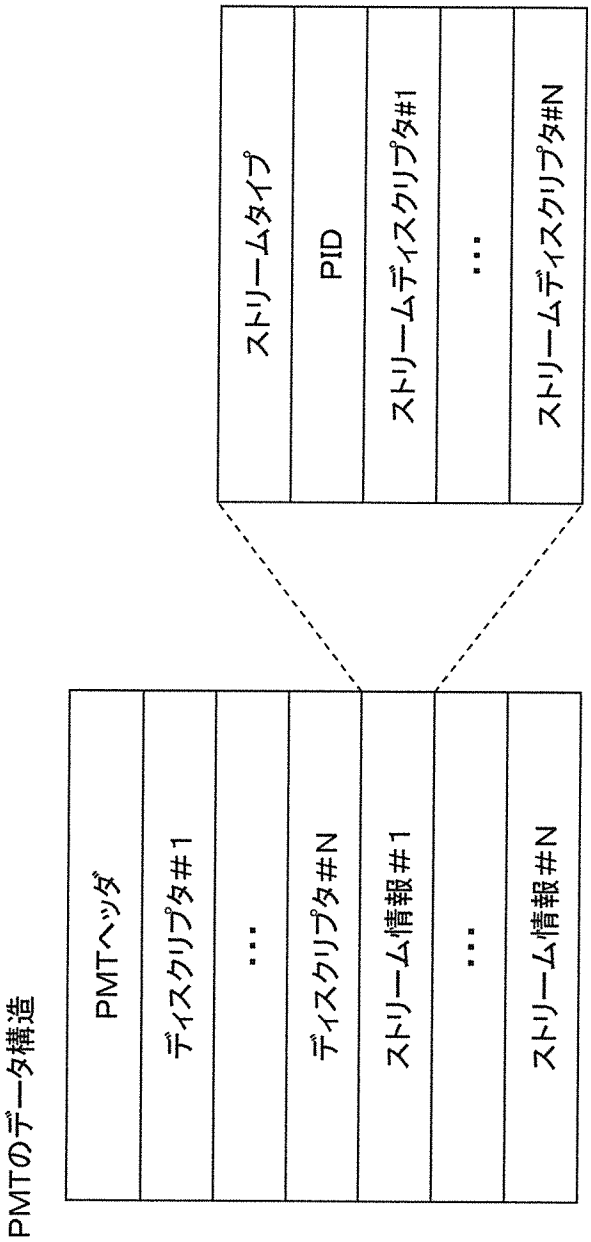
[図8]


ビデオフレーム列
Iピクチャ
Pピクチャ
Bピクチャ
Bピクチャ
Pピクチャ
Bピクチャ
...
yy1
yy2
yy3
yy4
PESパケット列
PES ヘッダ
PES ペイロード

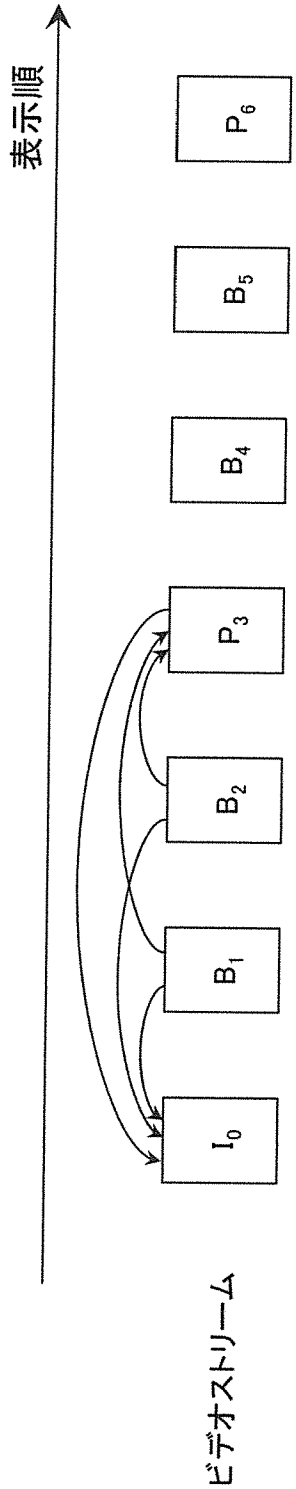
[図9]


TSヘッダ
TSパケット列
TSヘッダ
(4Byte)
アダプテーション
フィールド
TSペイロード
TSパケット
(188Byte)
TSストリーム
transport_
priority
PID
adaptation_field
_control
データ格納部

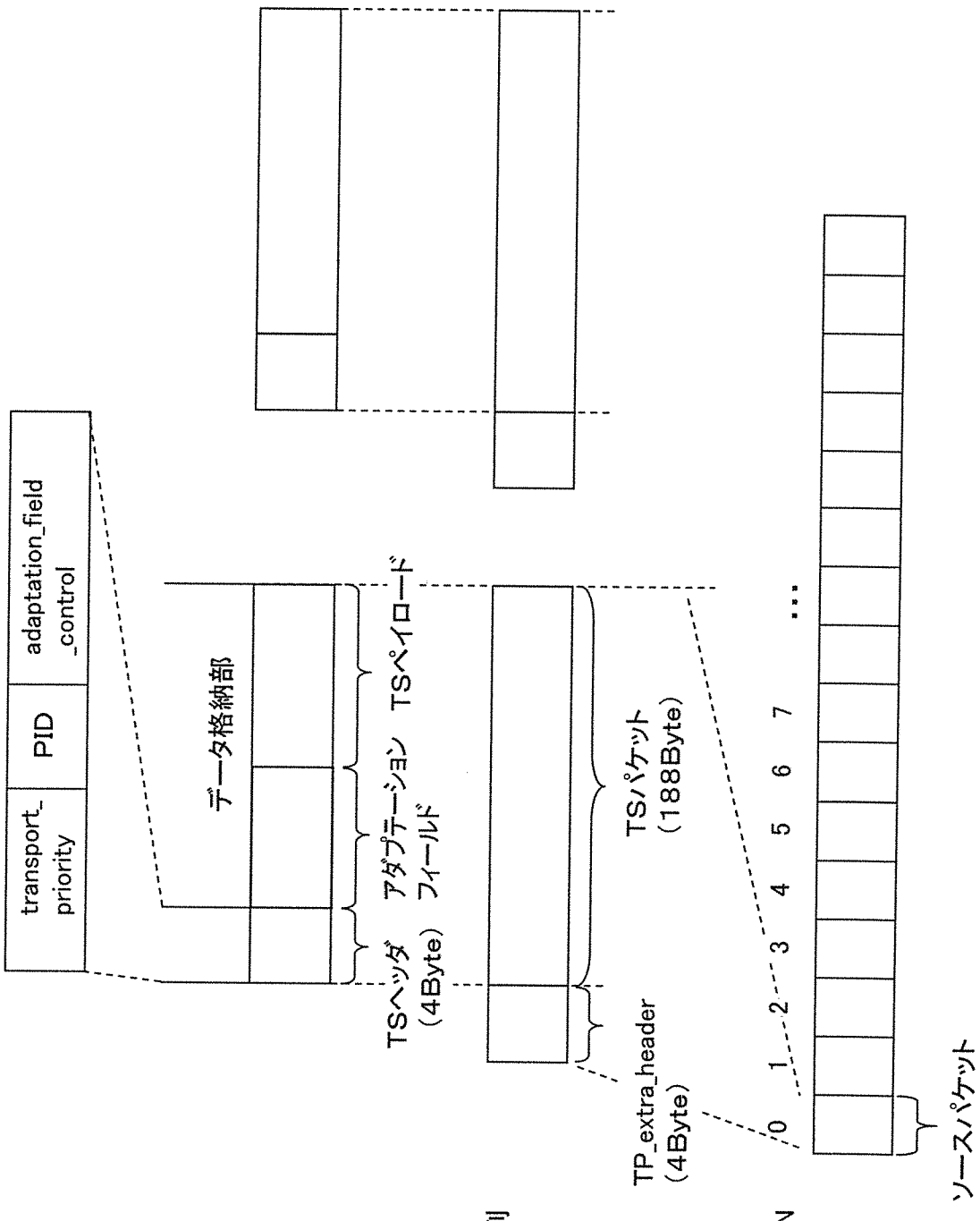
[図10]


PMTのデータ構造
PMTヘッダ
ディスクリプタ#1
...
ディスクリプタ#N
ストリーム情報#1
...
ストリーム情報#N
ストリームタイプ
PID
ストリームディスクリプタ#1
...
ストリームディスクリプタ#N

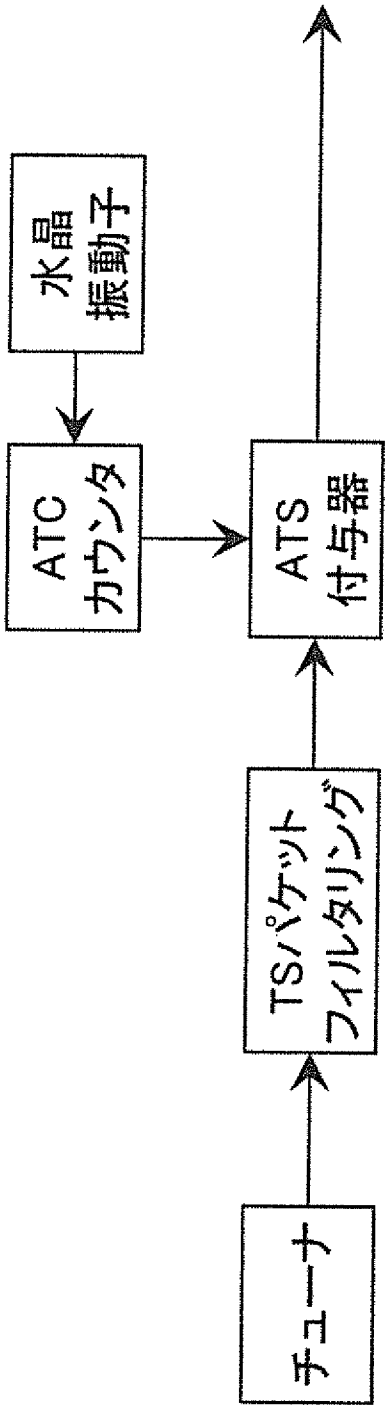
[図11]



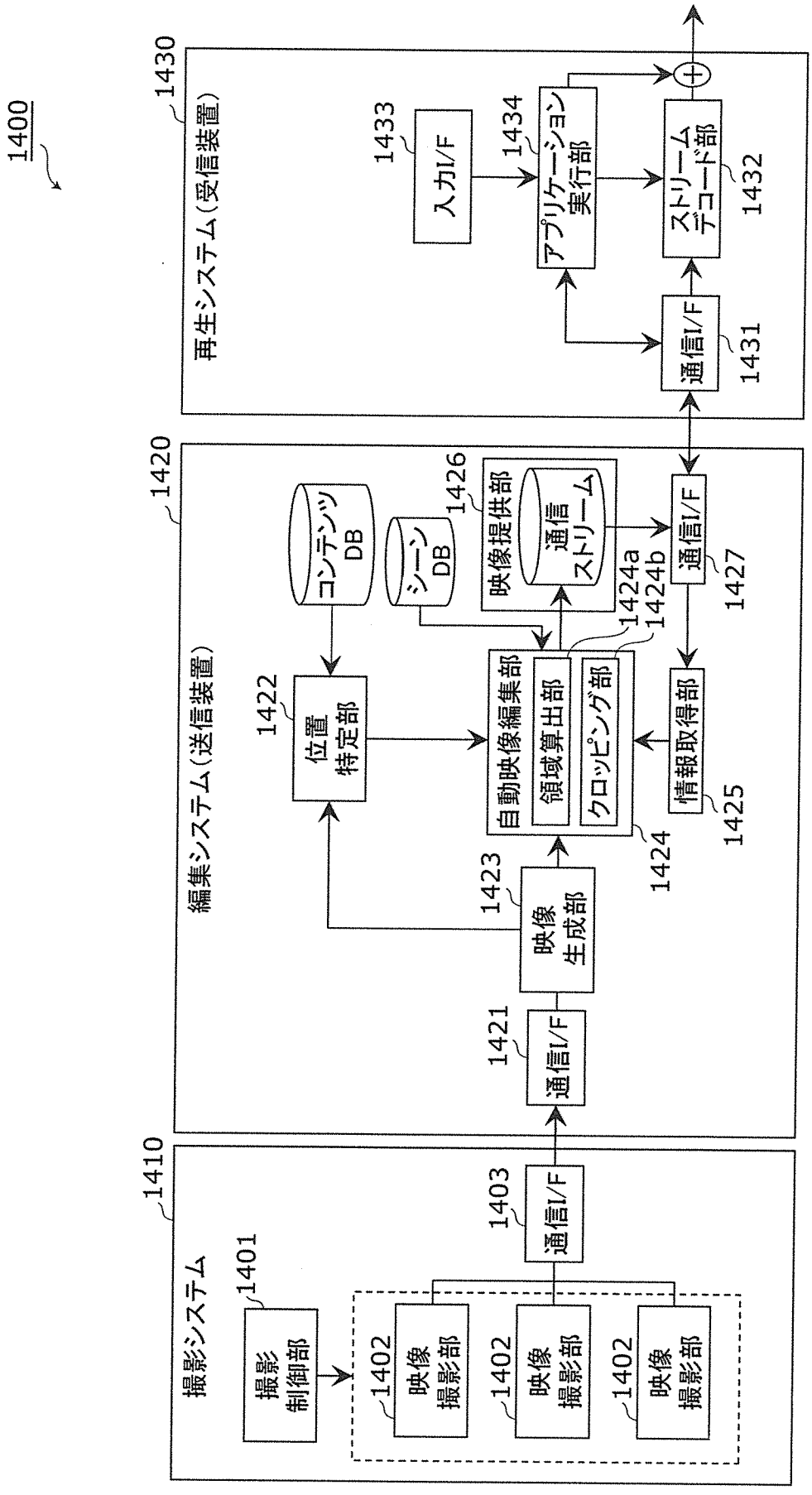
[図12]


TSヘッダ
transport_ priority
PID
adaptation_field _control
TSパケット列
データ格納部
TSヘッダ (4Byte)
アダプテーション フィールド
TSペイロード
ソースパケット列
TP_extra_header (4Byte)
TSパケット (188Byte)
TTSストリーム
SPN
0
1
2
3
4
5
6
7
...
ソースパケット

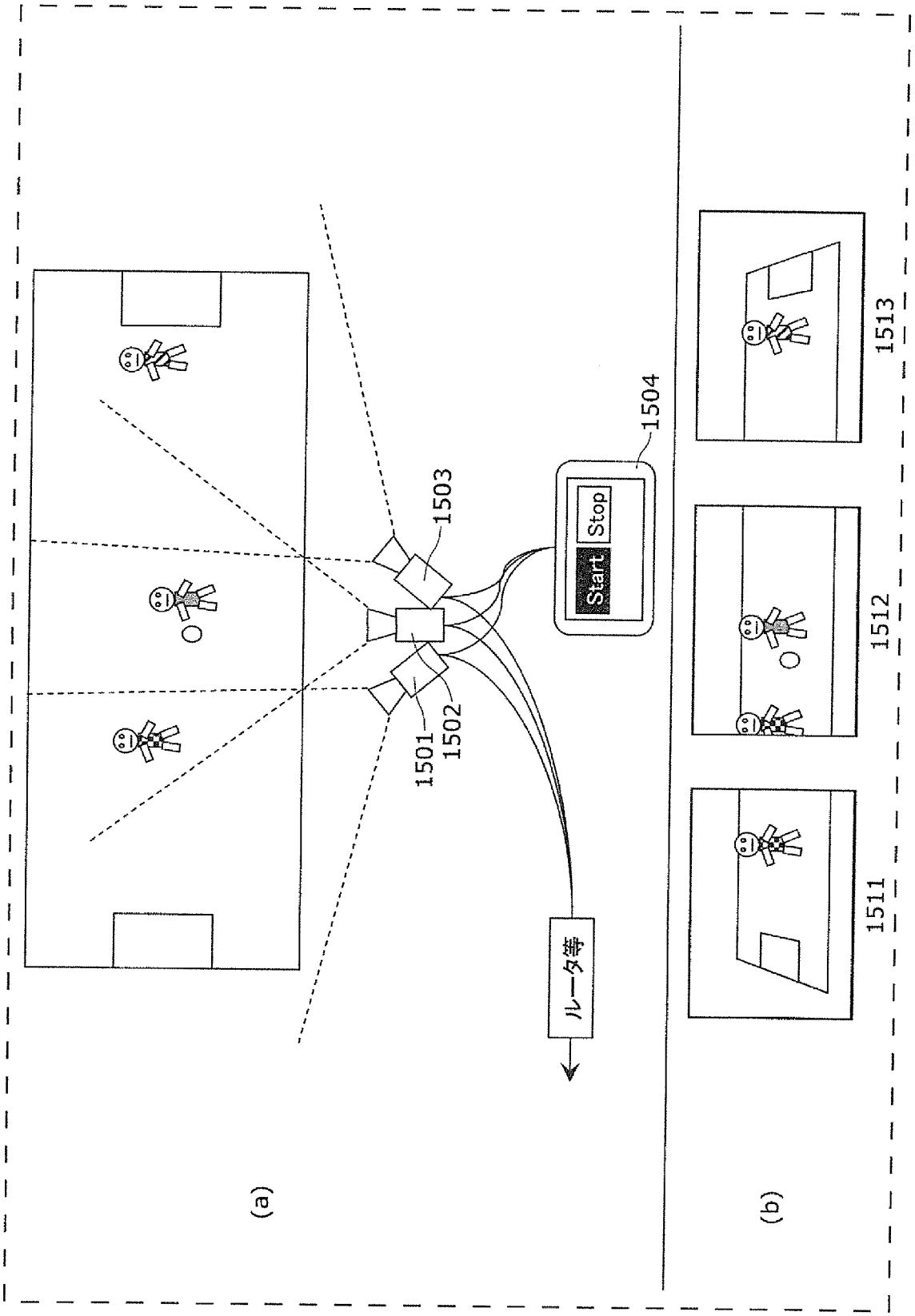
[図13]



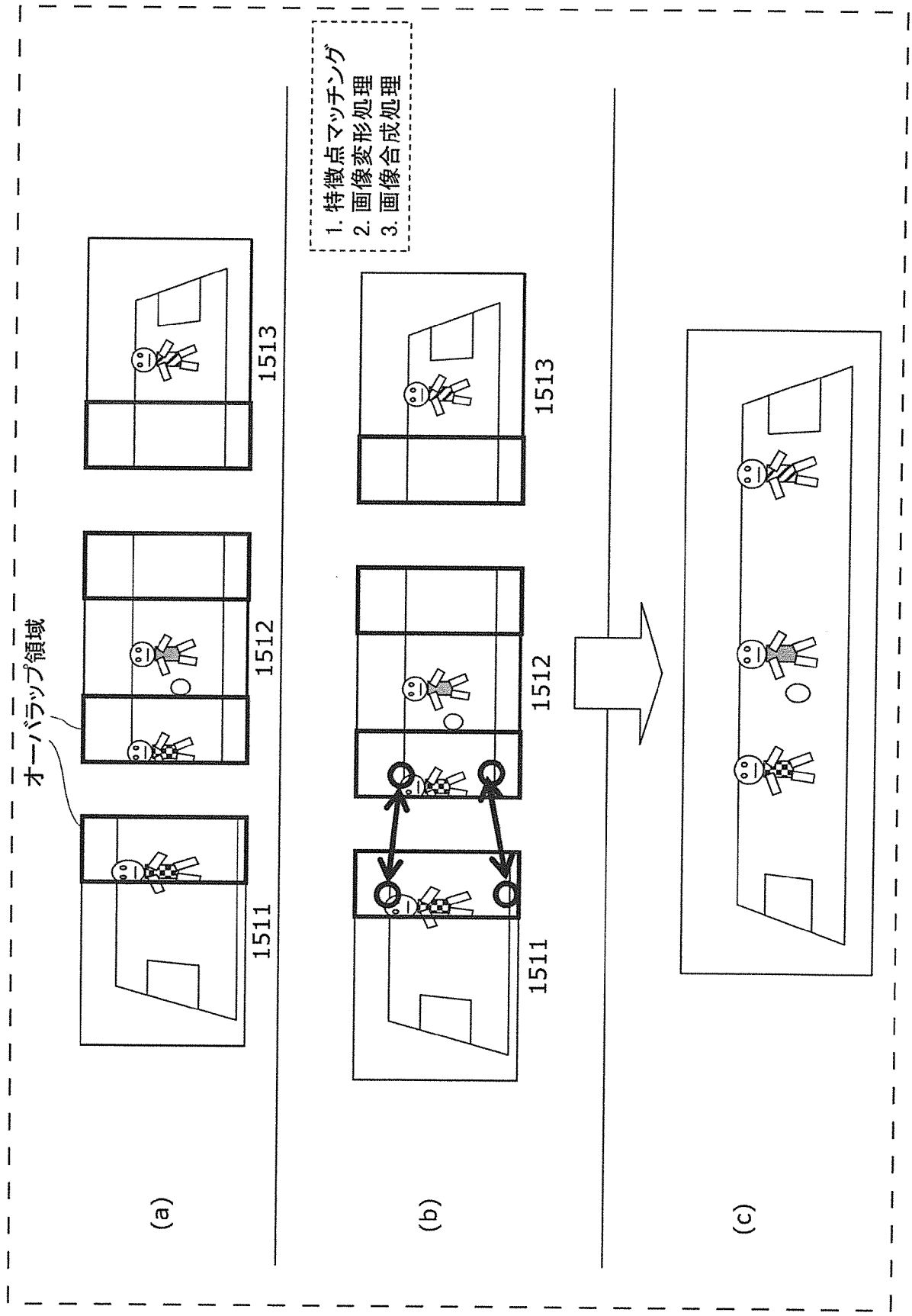
[図14]


1400
1410
撮影システム
1401
撮影制御部
1402
映像撮影部
1402
映像撮影部
1402
映像撮影部
1403
通信I/F
1420
編集システム(送信装置)
1421
通信I/F
1423
映像生成部
1422
位置特定部
コンテンツDB
シーンDB
1426
映像提供部
通信ストリーム
1424
自動映像編集部
領域算出部
クロッピング部
1424a
1424b
情報取得部
1425
通信I/F
1427
1430
再生システム(受信装置)
1433
入力I/F
1434
アプリケーション実行部
ストリームデコード部
1432
通信I/F
1431

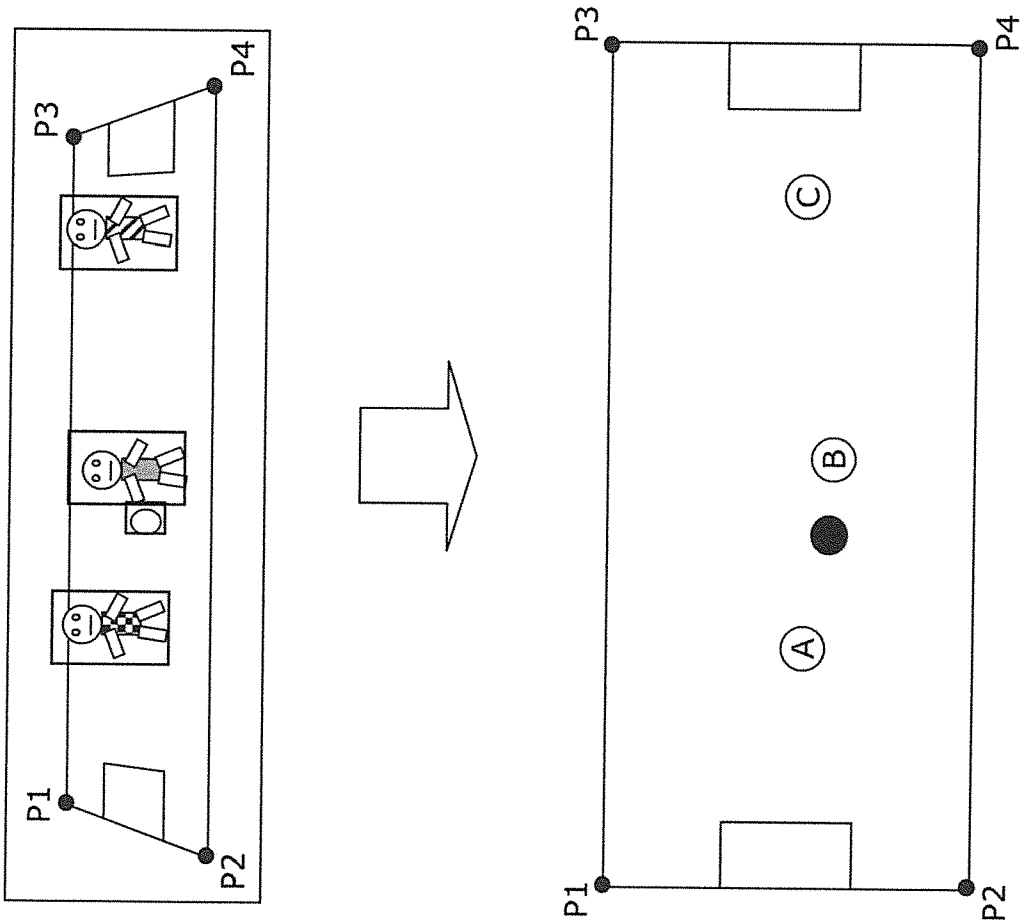
[図15]



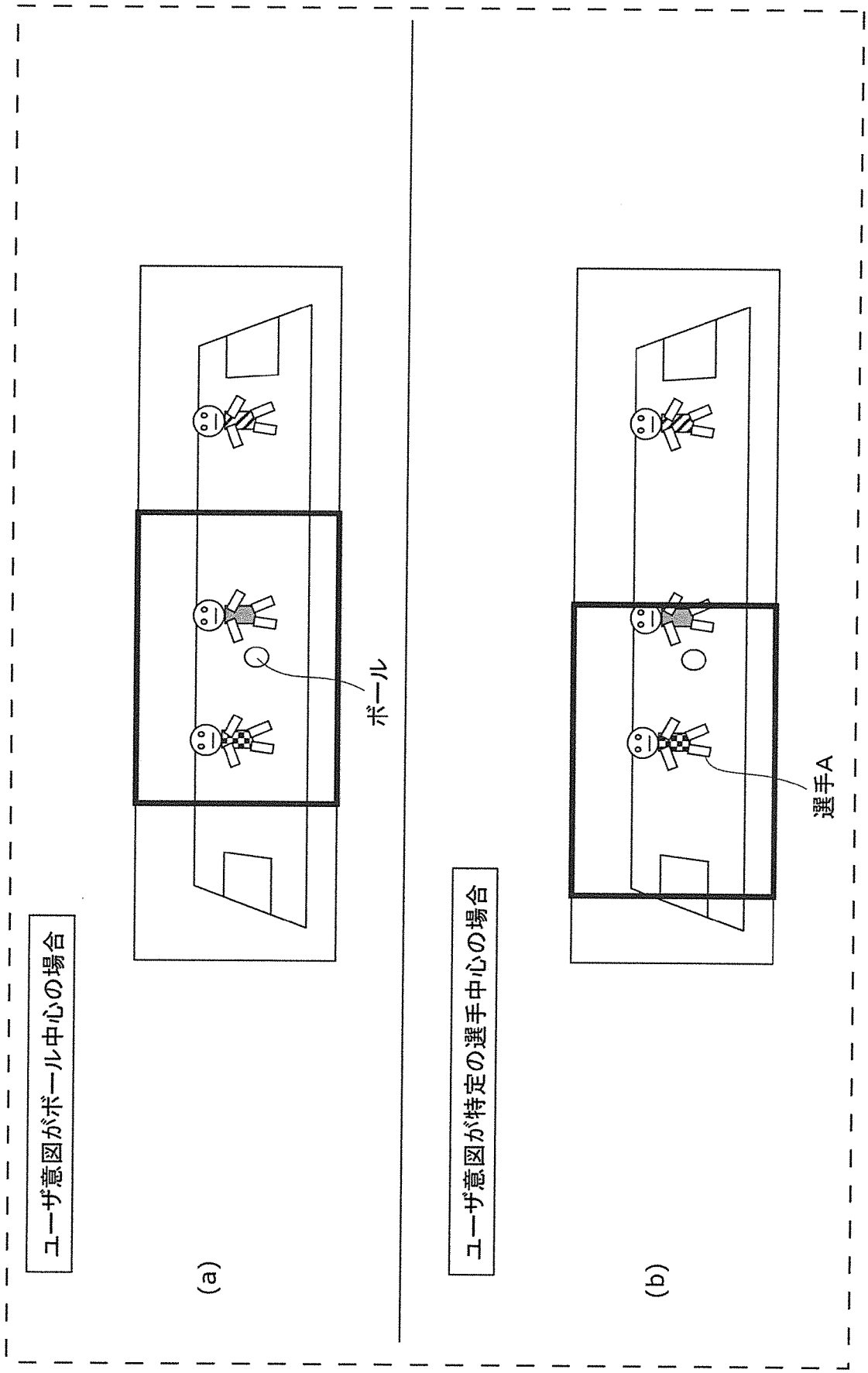
[図16]



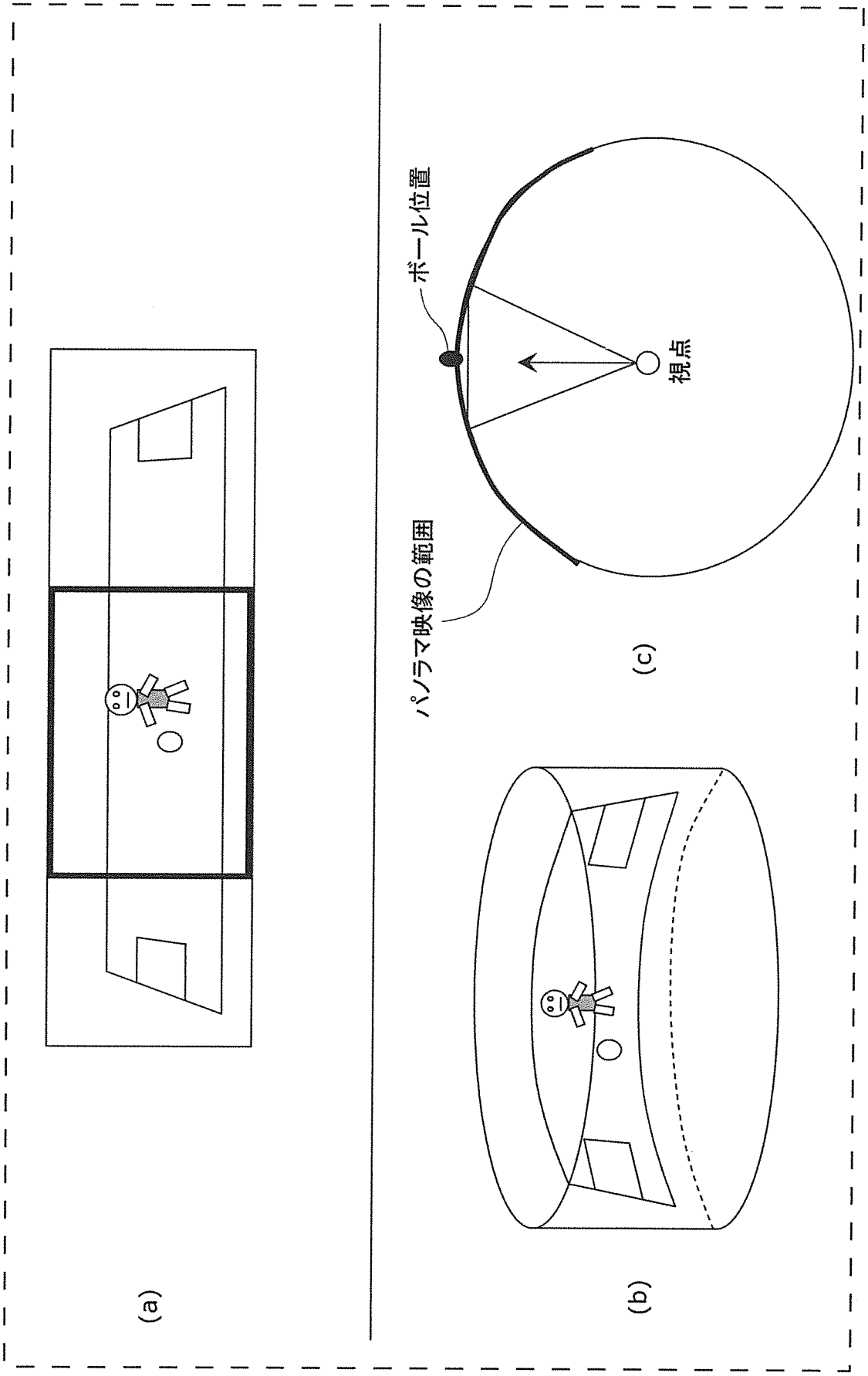
[図17]



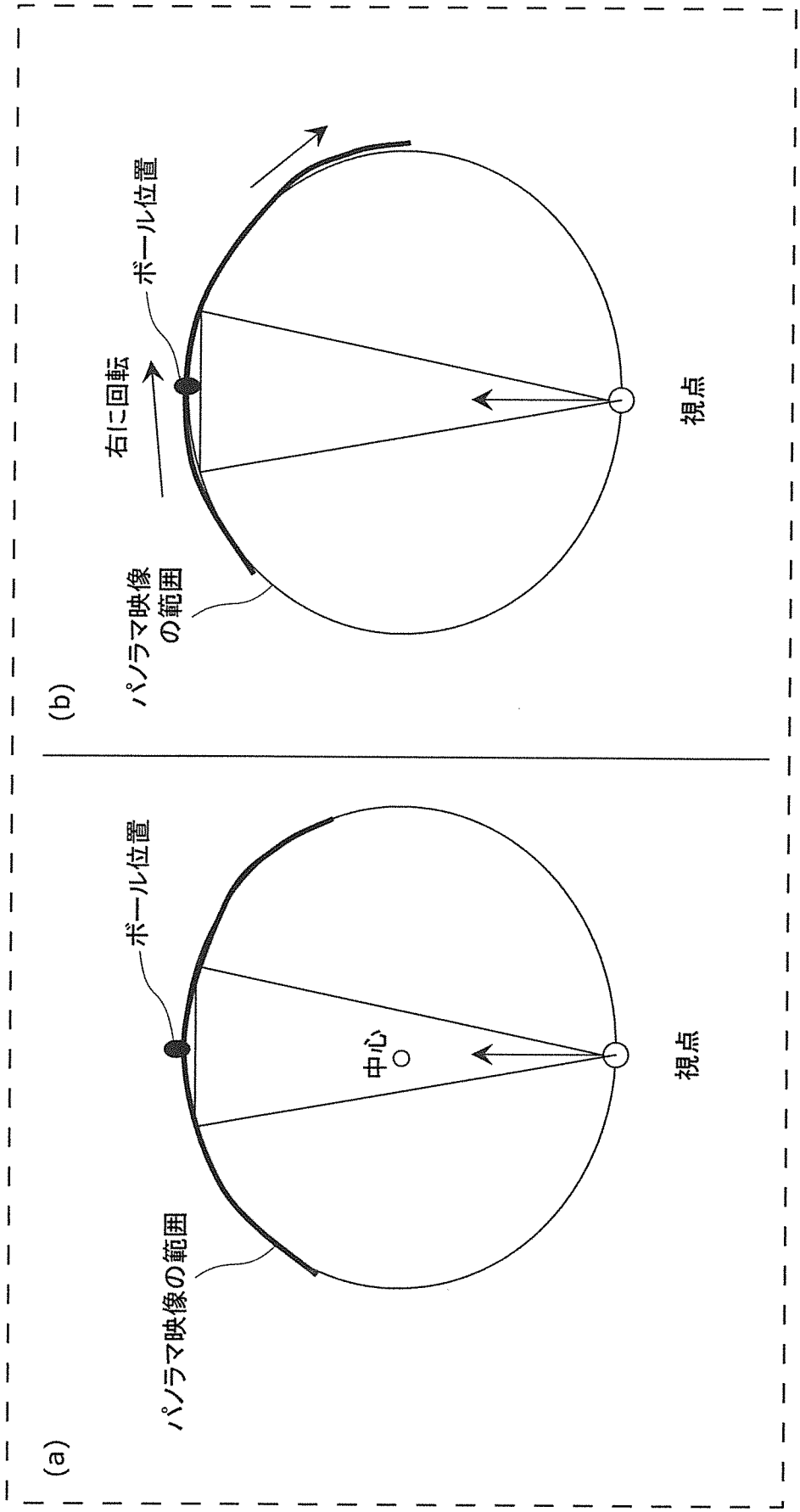
[図18]



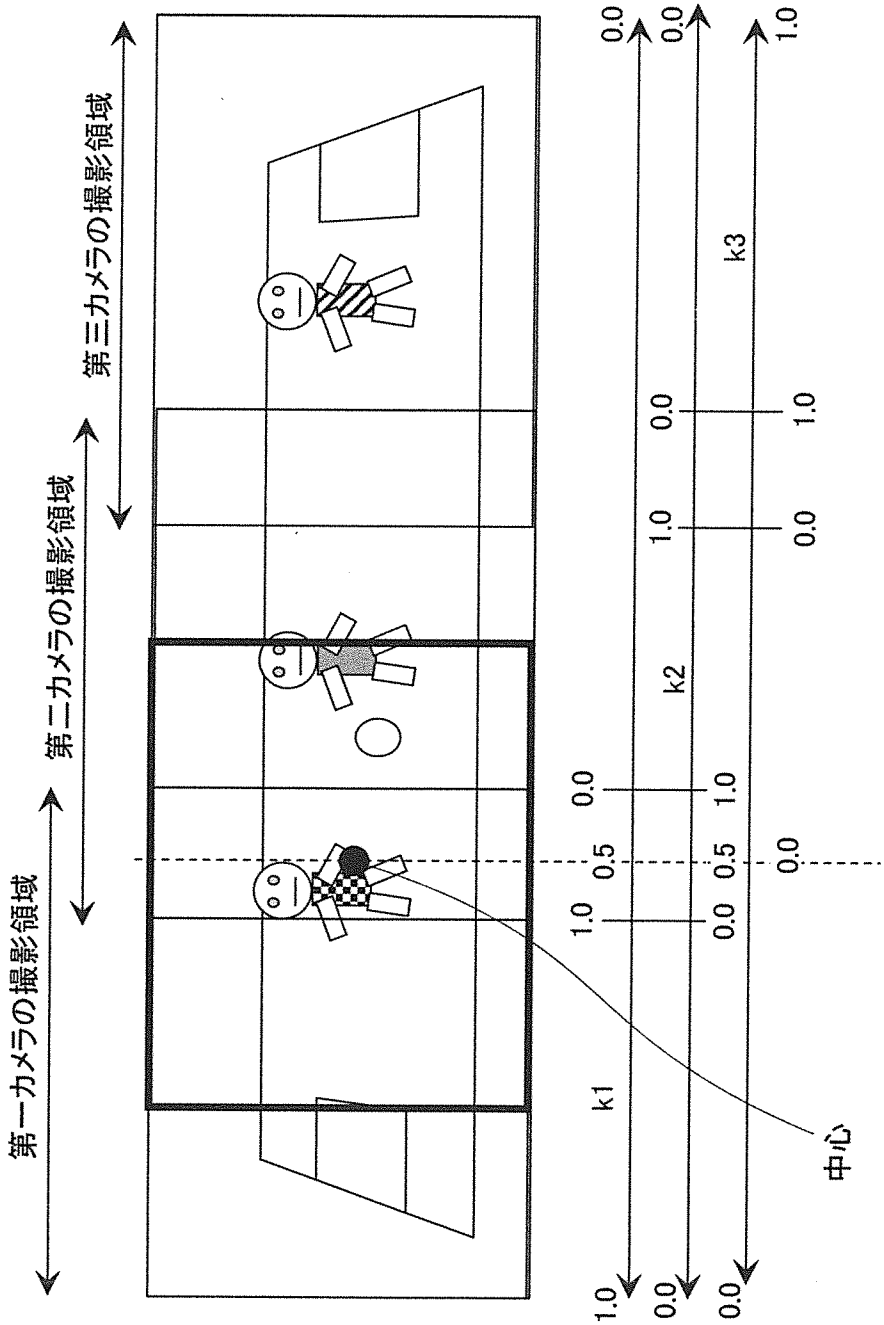
[図19]



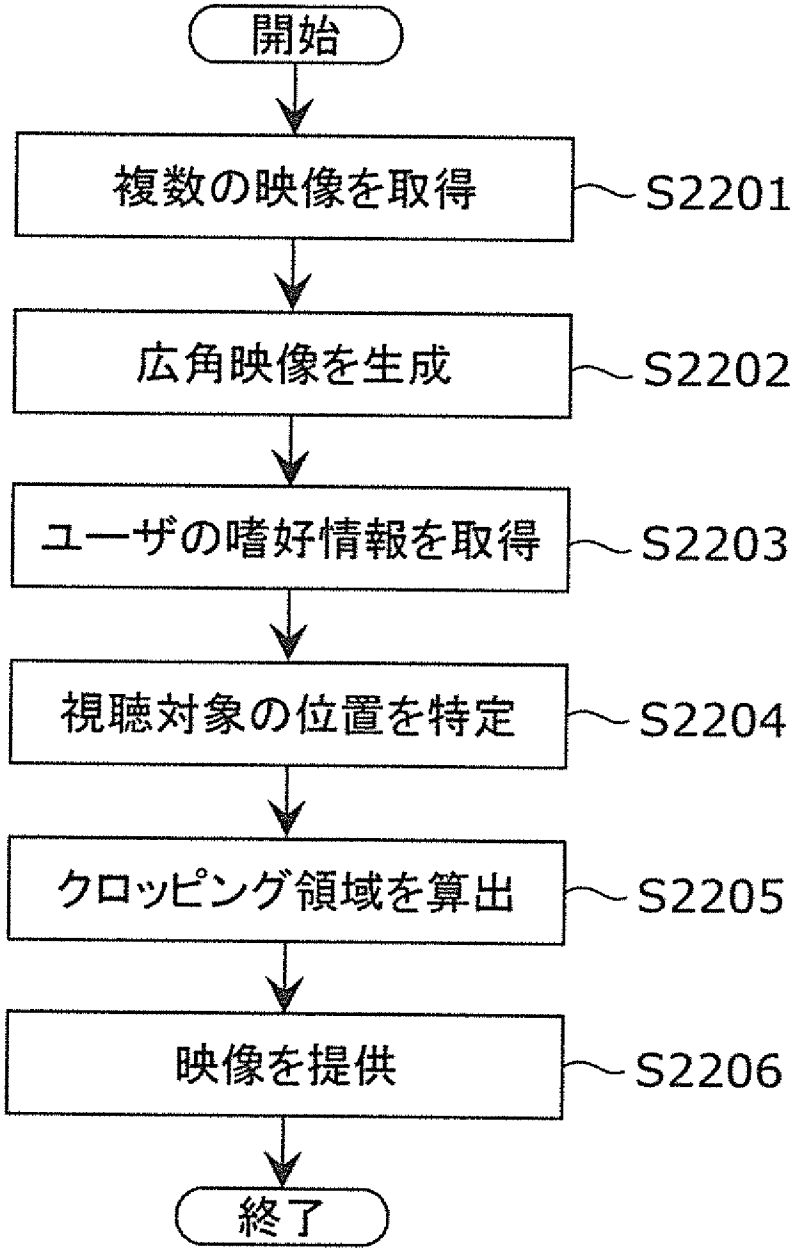
[図20]



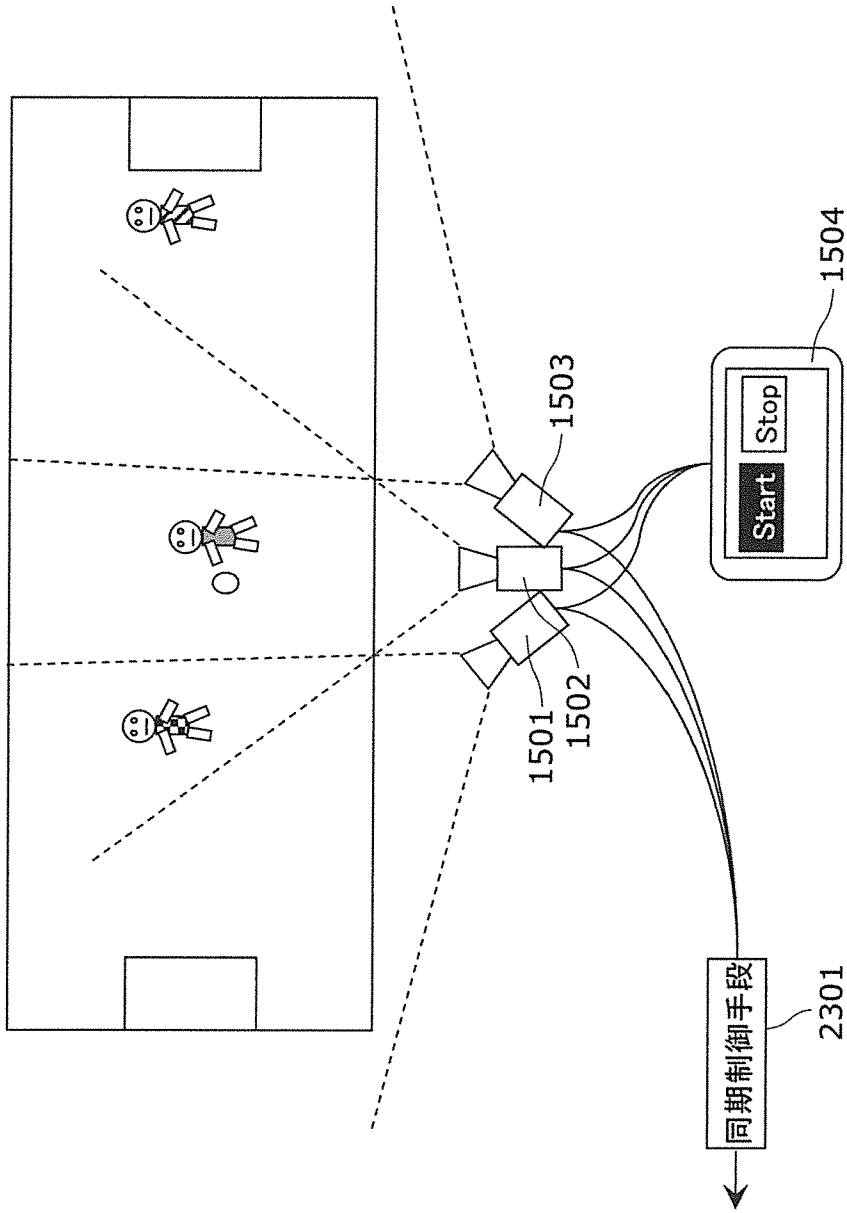
[図21]



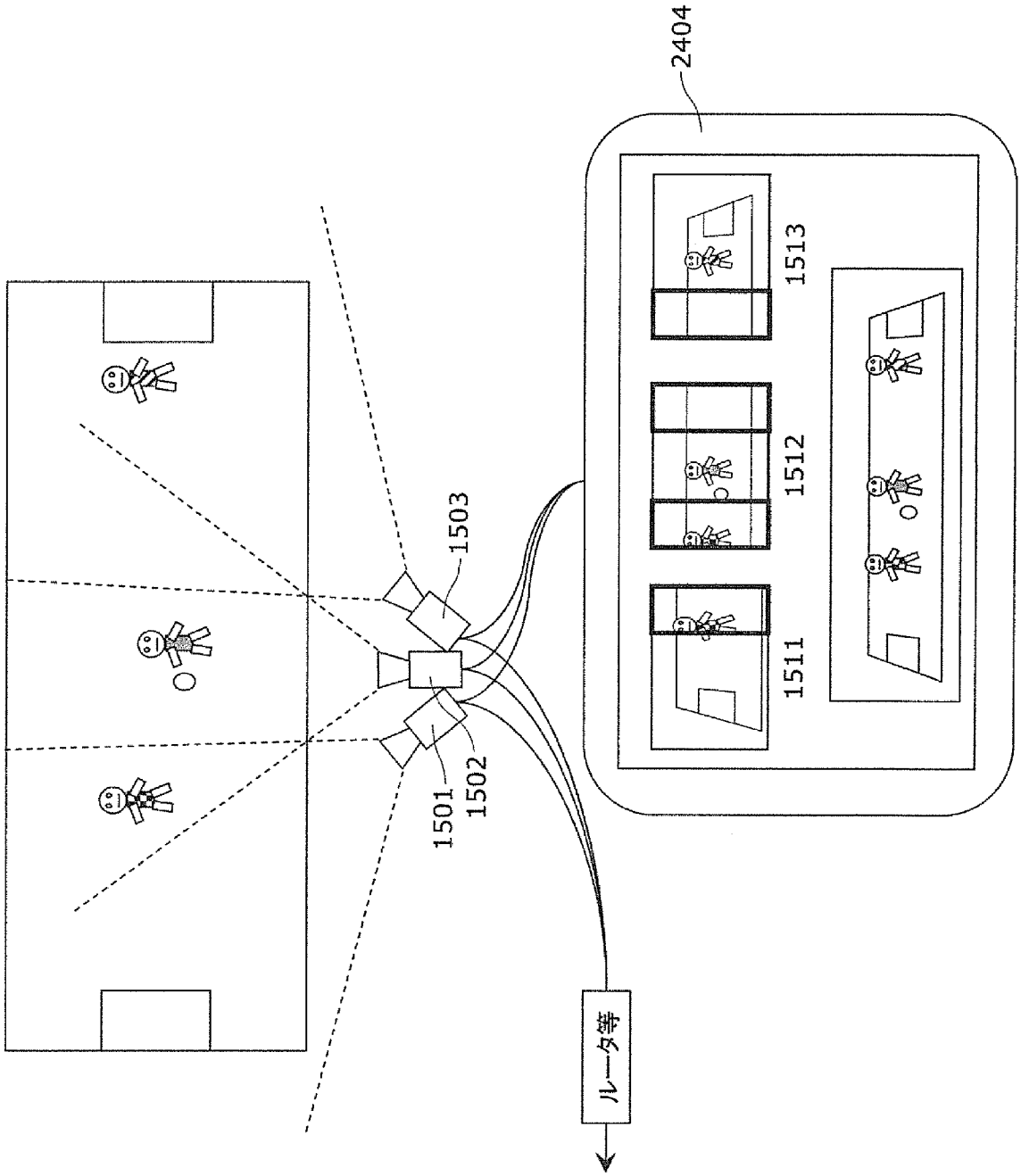
[図22]


開始
複数の映像を取得
S2201
広角映像を生成
S2202
ユーザの嗜好情報を取得
S2203
視聴対象の位置を特定
S2204
クロッピング領域を算出
S2205
映像を提供
S2206
終了

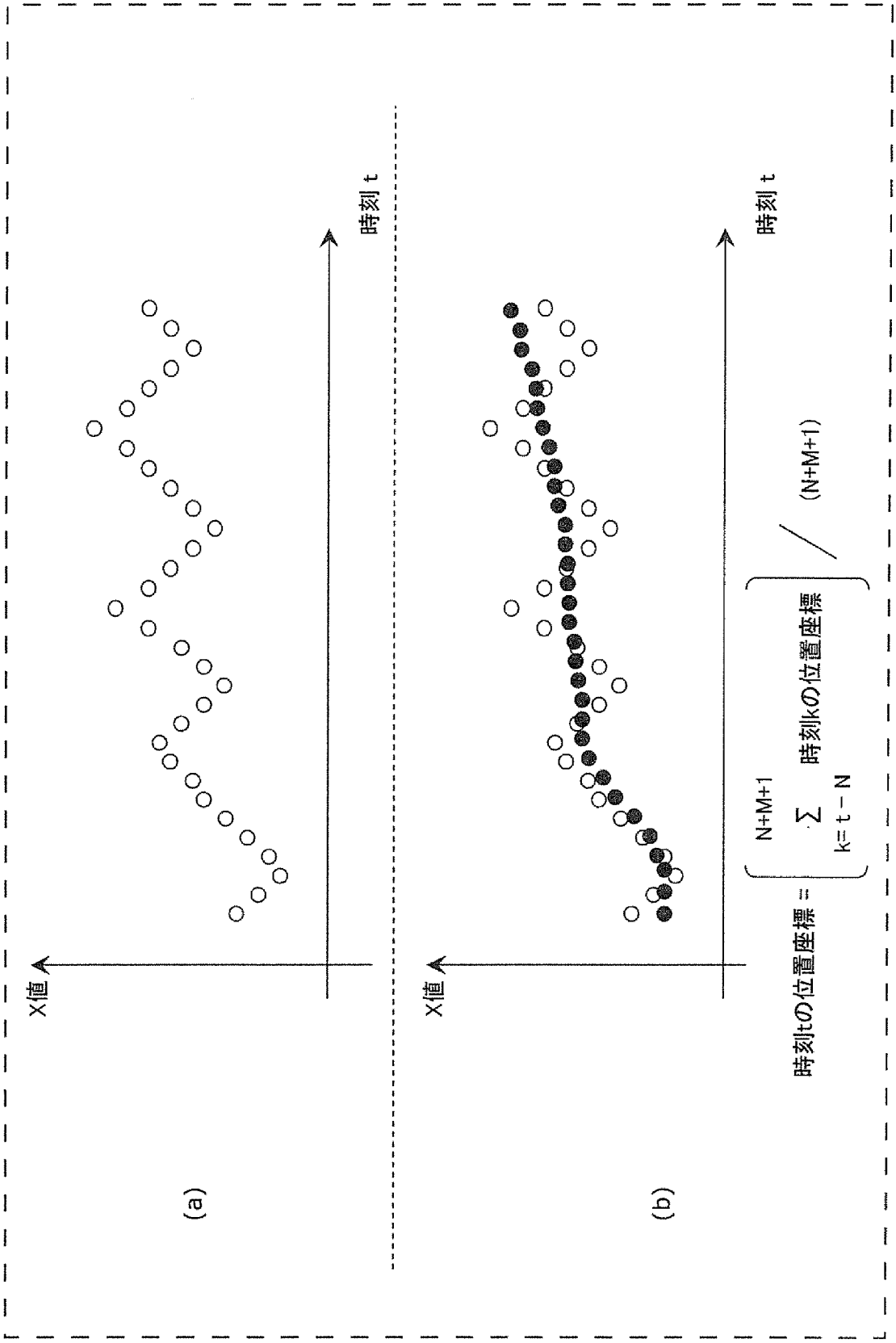
[図23]



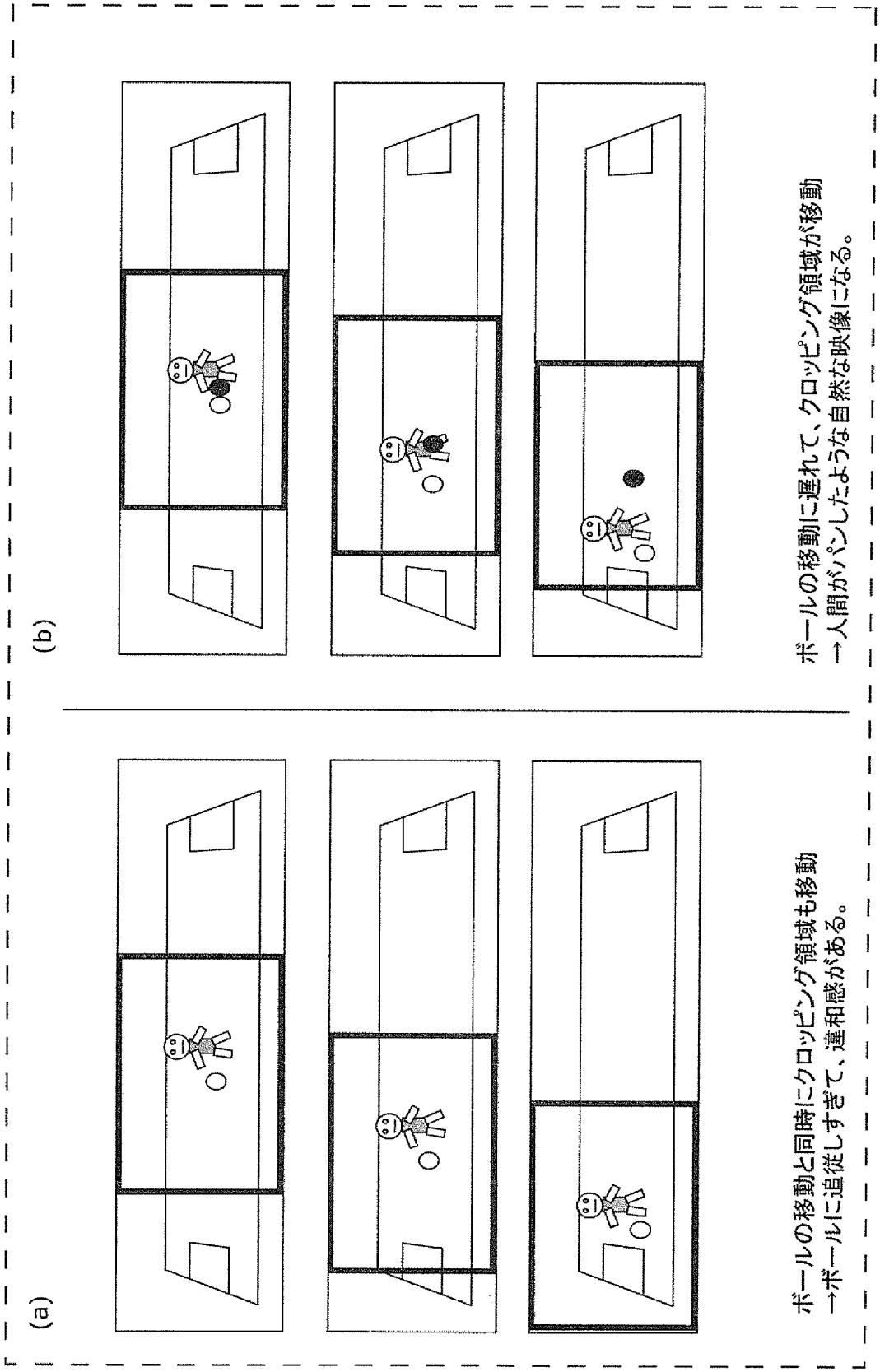
[図24]



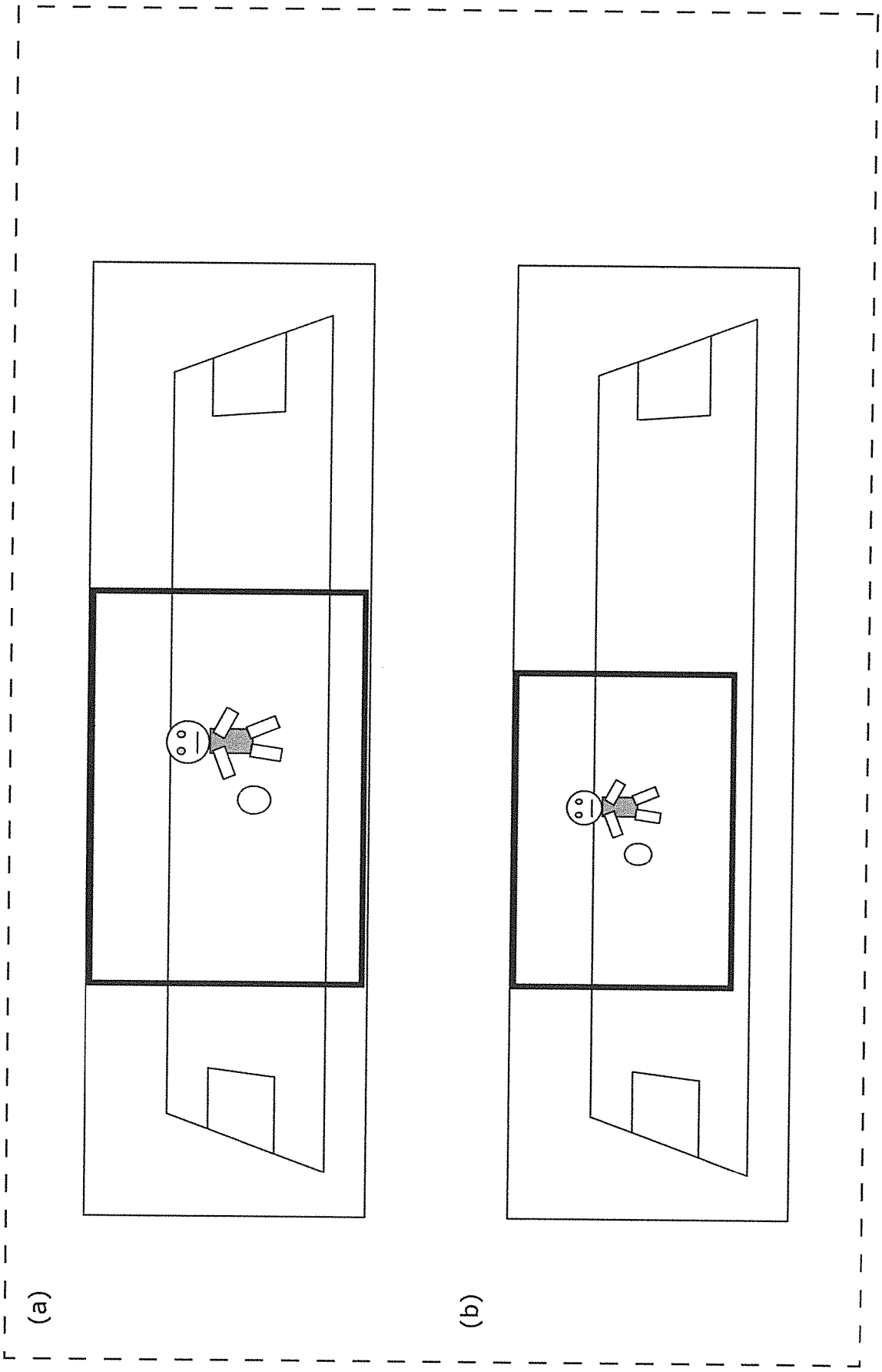
[図25]



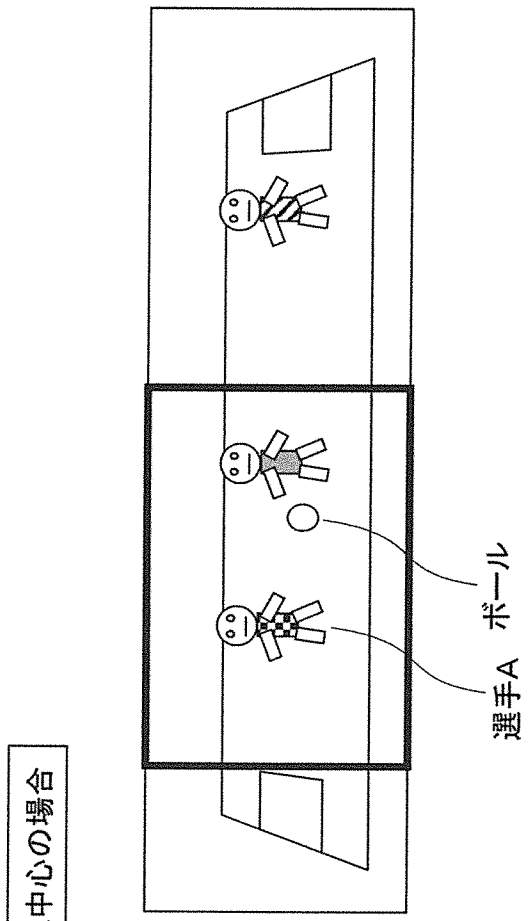
[図26]


(a)
ボールの移動と同時にクロッピング領域も移動
→ボールに追従しすぎて、違和感がある。
(b)
ボールの移動に遅れて、クロッピング領域が移動
→人間がパンしたような自然な映像になる。

[図27]

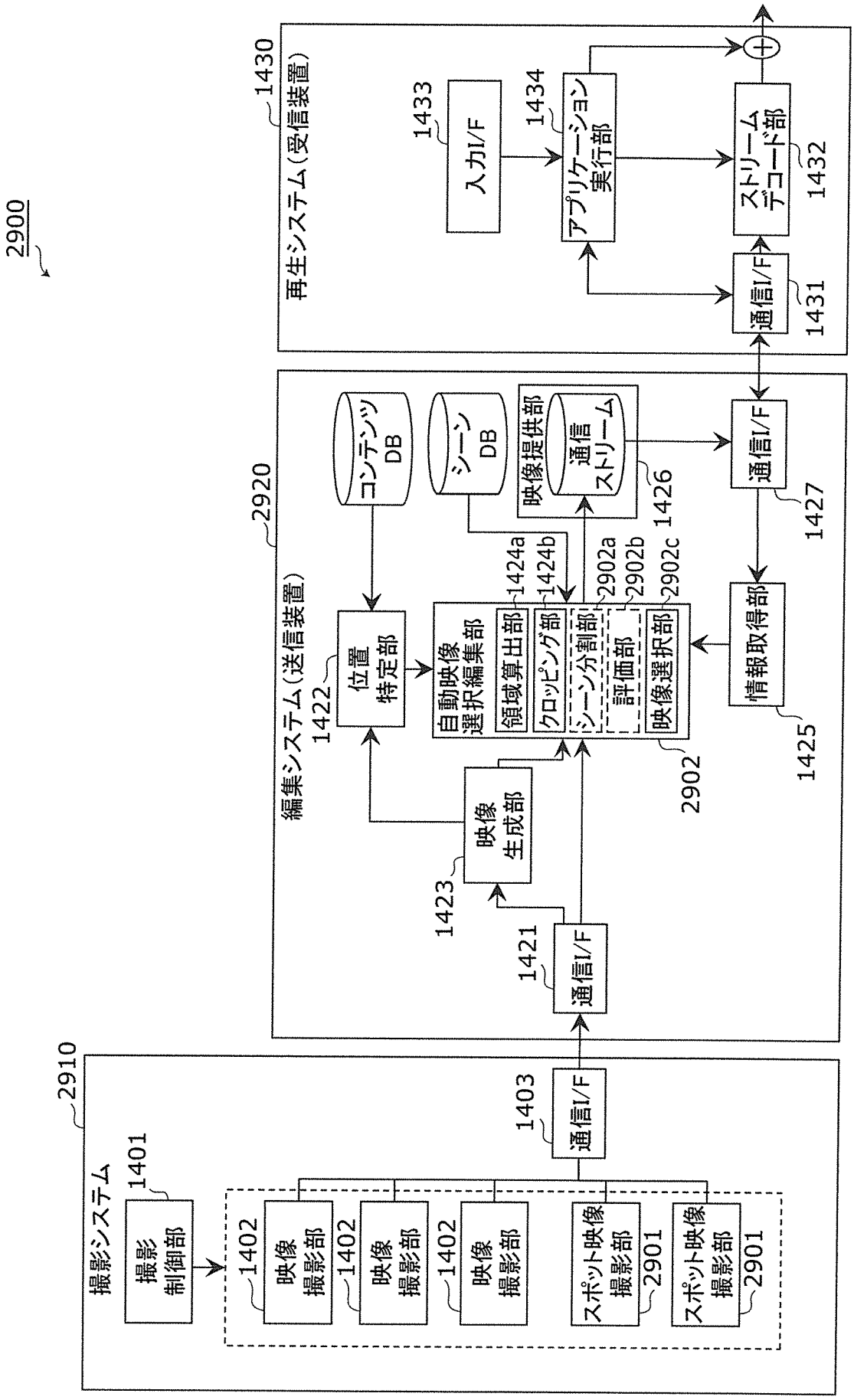


[図28]

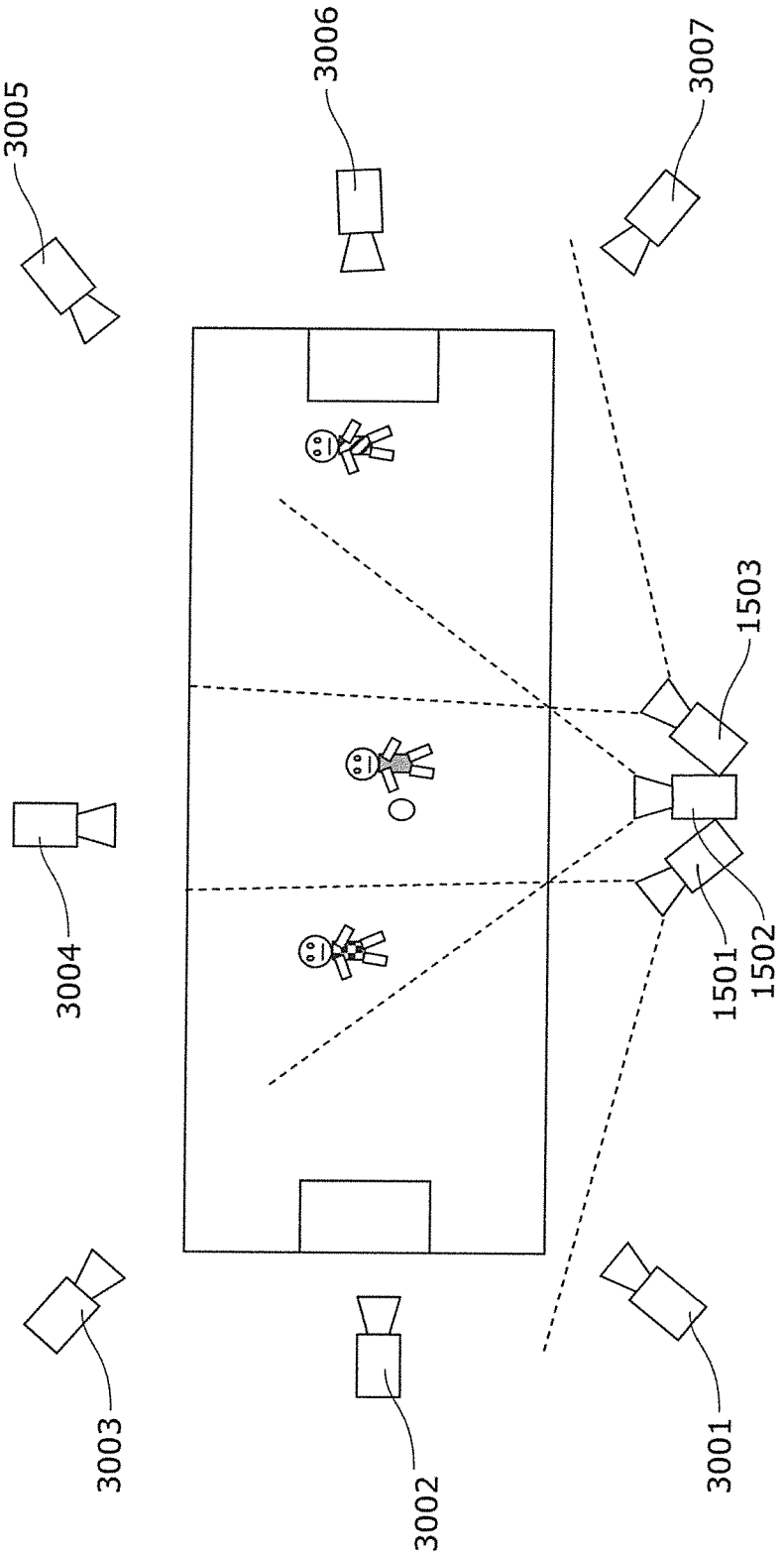


ユーザ意図がボールと選手A中心の場合

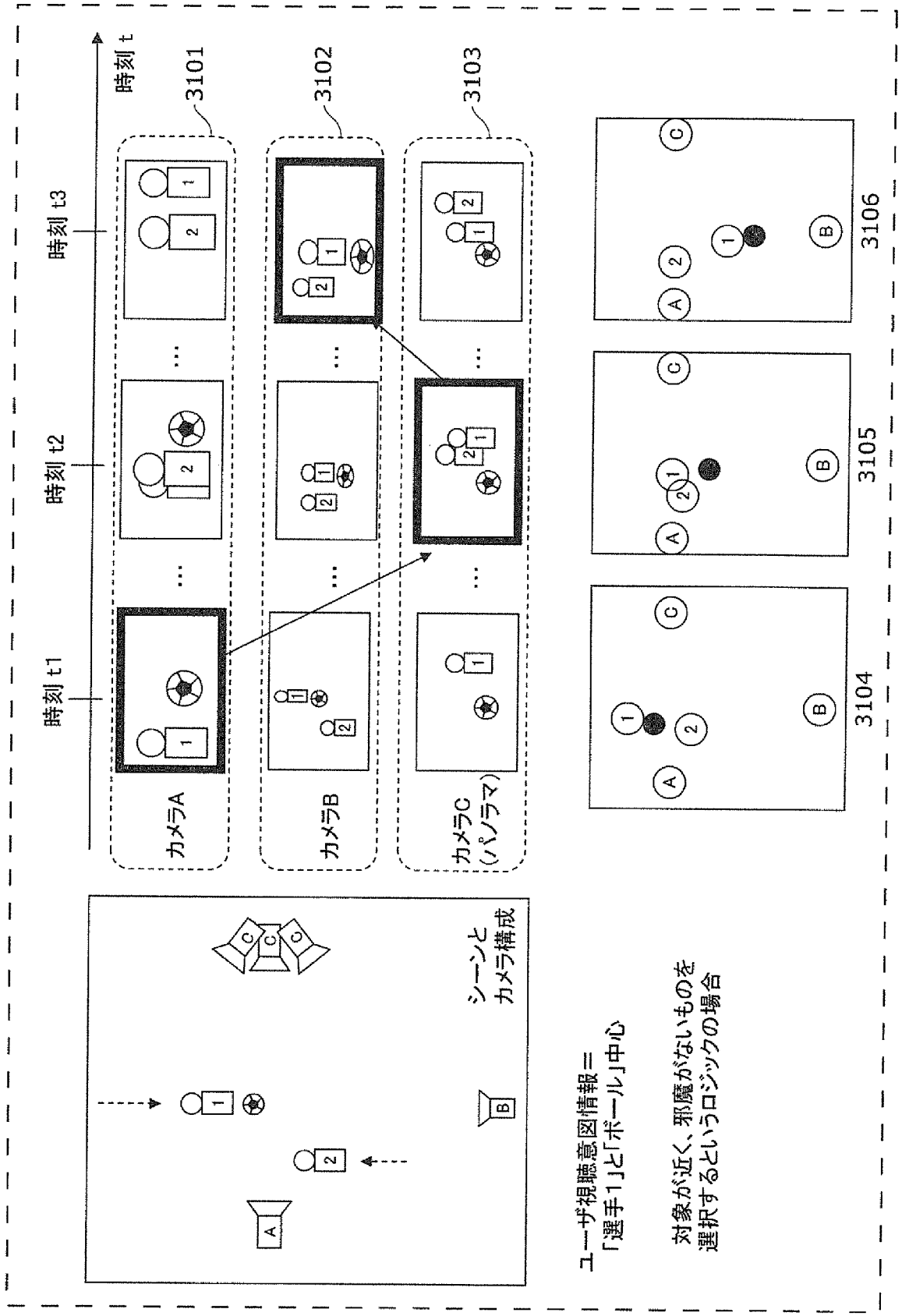
[図29]


2900
再生システム(受信装置)
1430
入力I/F
1433
アプリケーション実行部
1434
ストリームデコード部
1432
通信I/F
1431
編集システム(送信装置)
2920
コンテンツDB
シーンDB
映像提供部
通信ストリーム
1426
通信I/F
1427
位置特定部
1422
自動映像選択編集部
領域算出部
クロッピング部
シーン分割部
評価部
映像選択部
1424a
1424b
2902a
2902b
2902c
2902
情報取得部
1425
映像生成部
1423
通信I/F
1421
撮影システム
2910
撮影制御部
1401
映像撮影部
1402
映像撮影部
1402
映像撮影部
1402
スポット映像撮影部
2901
スポット映像撮影部
2901
通信I/F
1403

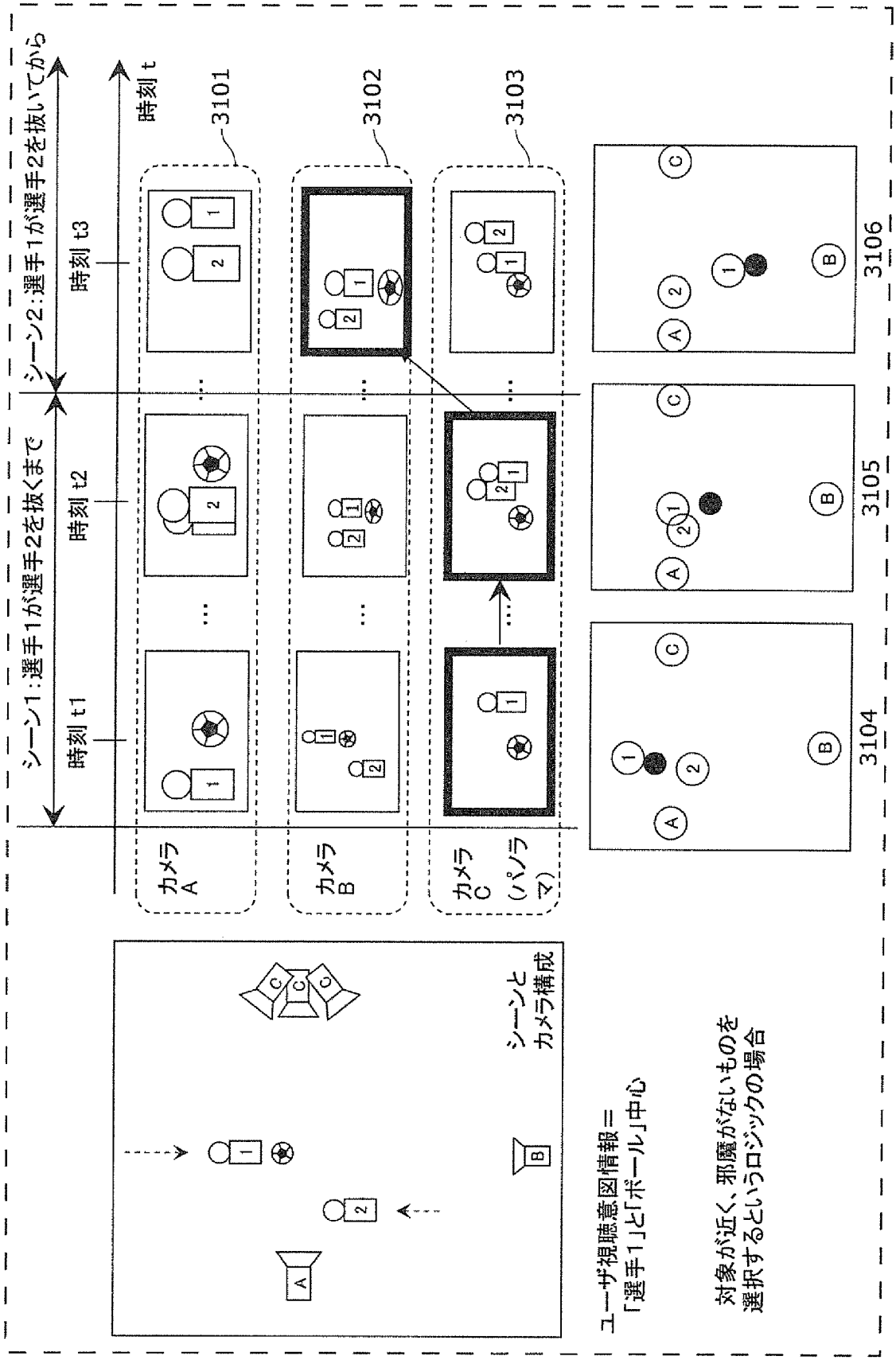
[図30]



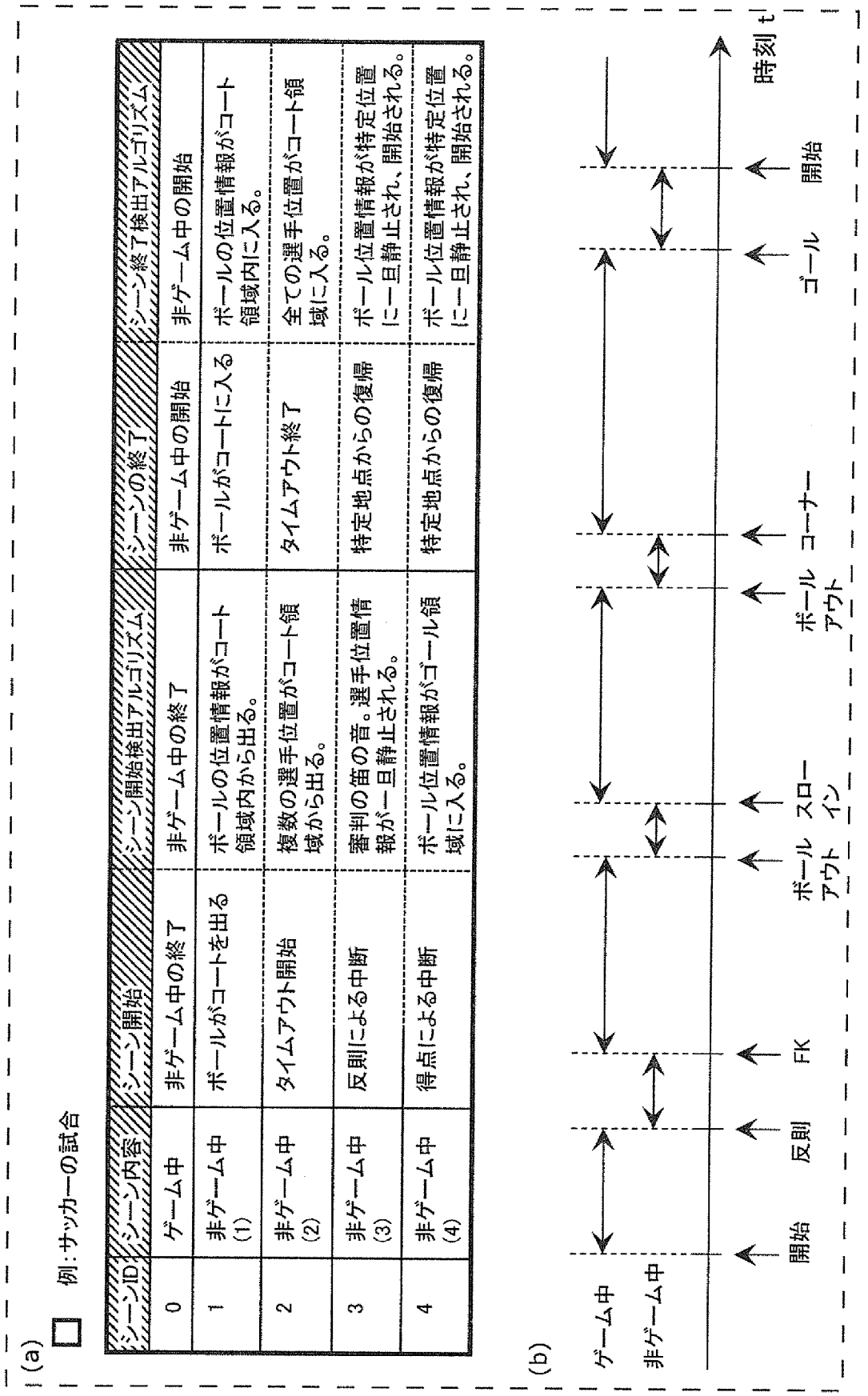
[図31]


時刻 t
時刻 t1
時刻 t2
時刻 t3
3101
3102
3103
3104
3105
3106
カメラA
カメラB
カメラC
(パノラマ)
シーンと
カメラ構成
ユーザ視聴意図情報＝
「選手1」と「ボール」中心
対象が近く、邪魔がないものを
選択するというロジックの場合

[図32]

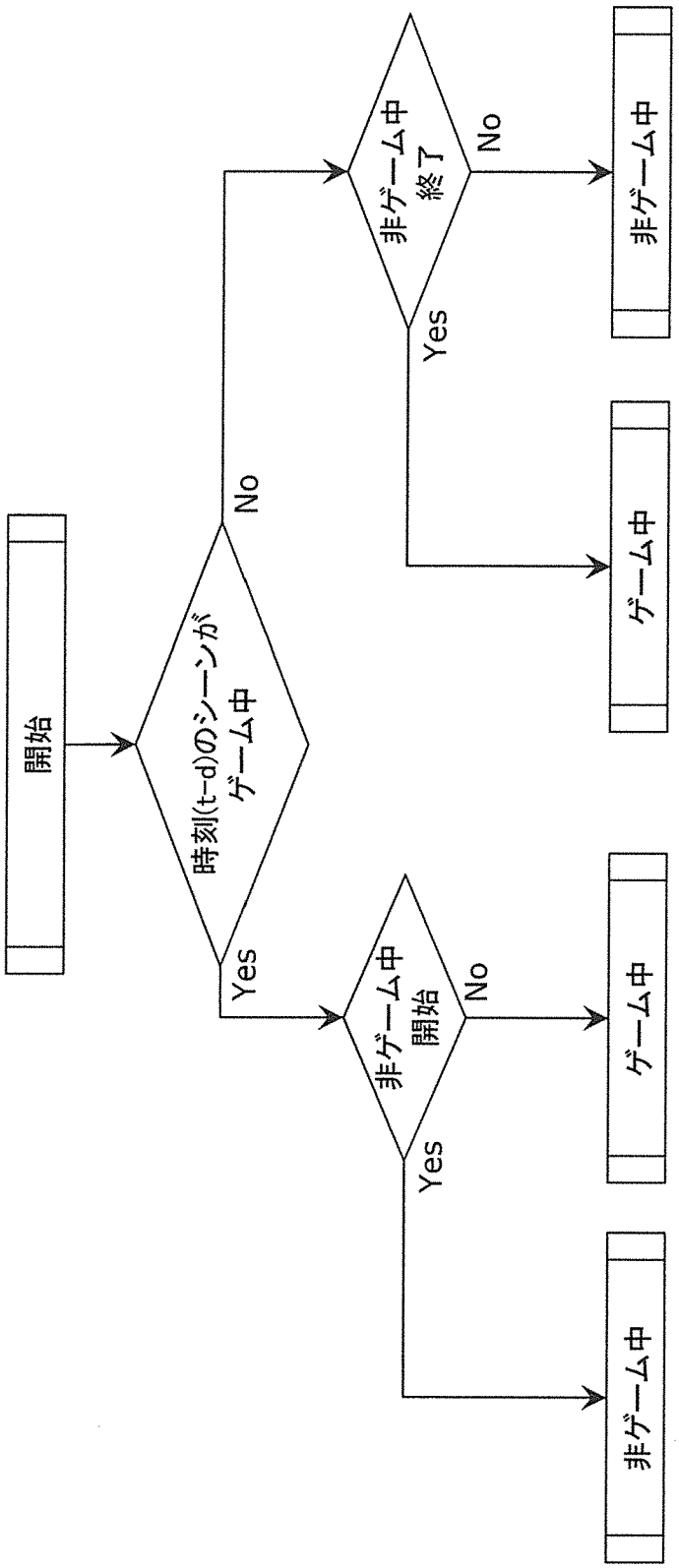

シーン1:選手1が選手2を抜くまで
シーン2:選手1が選手2を抜いてから
時刻 t1
時刻 t2
時刻 t3
時刻 t
3101
3102
3103
3104
3105
3106
カメラA
カメラB
カメラC（パノラマ）
シーンとカメラ構成
ユーザ視聴意図情報＝「選手1」と「ボール」中心
対象が近く、邪魔がないものを選択するというロジックの場合

[図33]

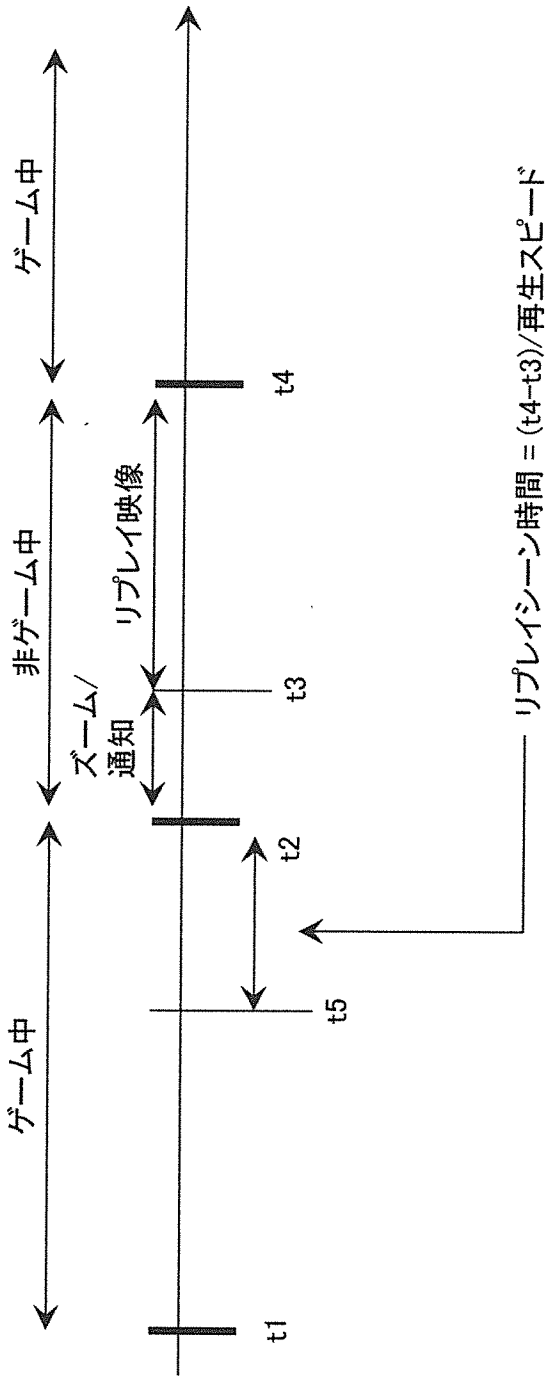


シーンID	シーン内容	シーン開始	シーン開始検出アルゴリズム	シーンの終了	シーン終了検出アルゴリズム
0	ゲーム中	非ゲーム中の終了	非ゲーム中の終了	非ゲーム中の開始	非ゲーム中の開始
1	非ゲーム中 (1)	ボールがコートを出る	ボールの位置情報がコート領域内から出る。	ボールがコートに入る	ボールの位置情報がコート領域内に入る。
2	非ゲーム中 (2)	タイムアウト開始	複数の選手位置がコート領域から出る。	タイムアウト終了	全ての選手位置がコート領域に入る。
3	非ゲーム中 (3)	反則による中断	審判の笛の音。選手位置情報が一旦静止される。	特定地点からの復帰	ボール位置情報が特定位置に一旦静止され、開始される。
4	非ゲーム中 (4)	得点による中断	ボール位置情報がゴール領域に入る。	特定地点からの復帰	ボール位置情報が特定位置に一旦静止され、開始される。

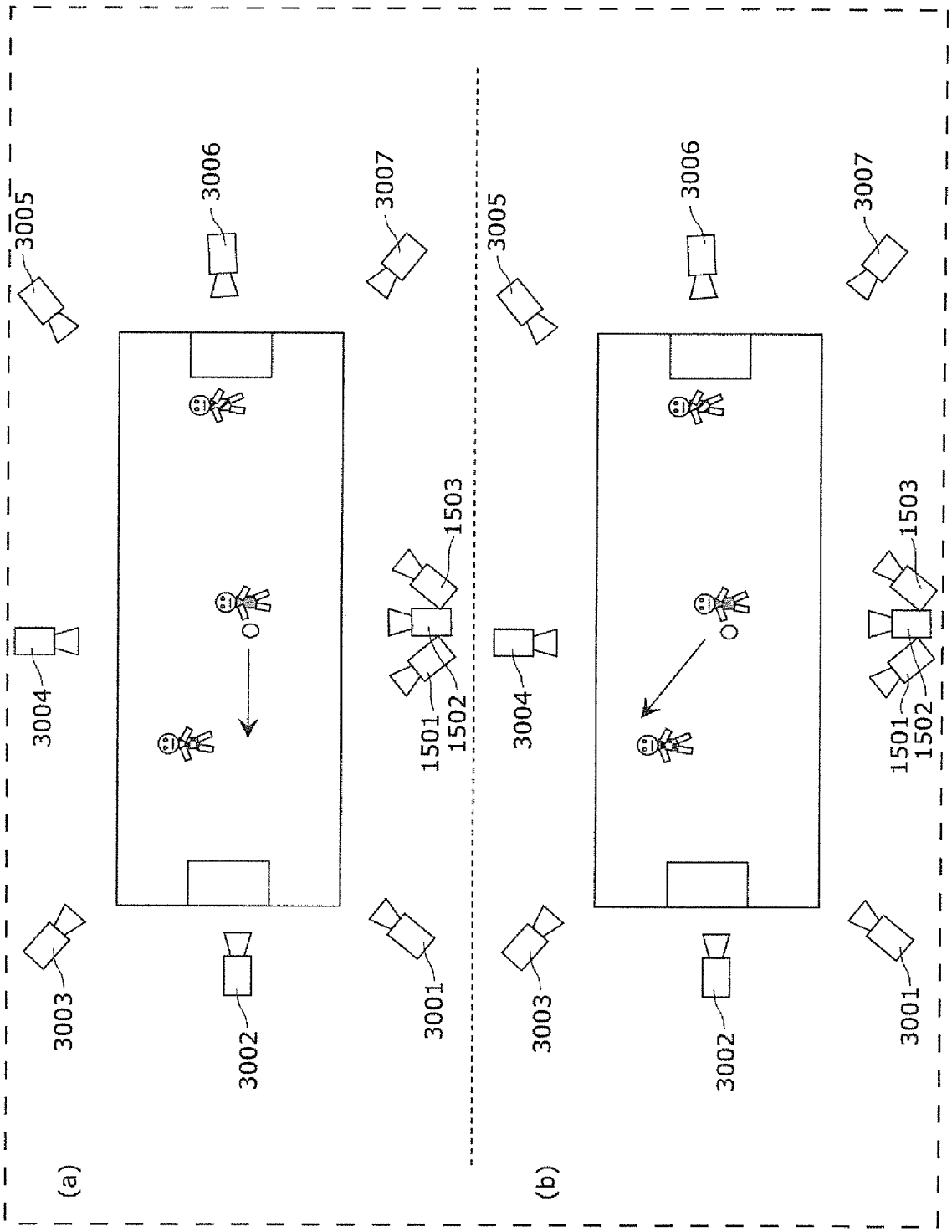
[図34]


開始
時刻(t-d)のシーンがゲーム中
No
Yes
非ゲーム中終了
No
Yes
非ゲーム中
ゲーム中
非ゲーム中開始
No
Yes
ゲーム中
非ゲーム中

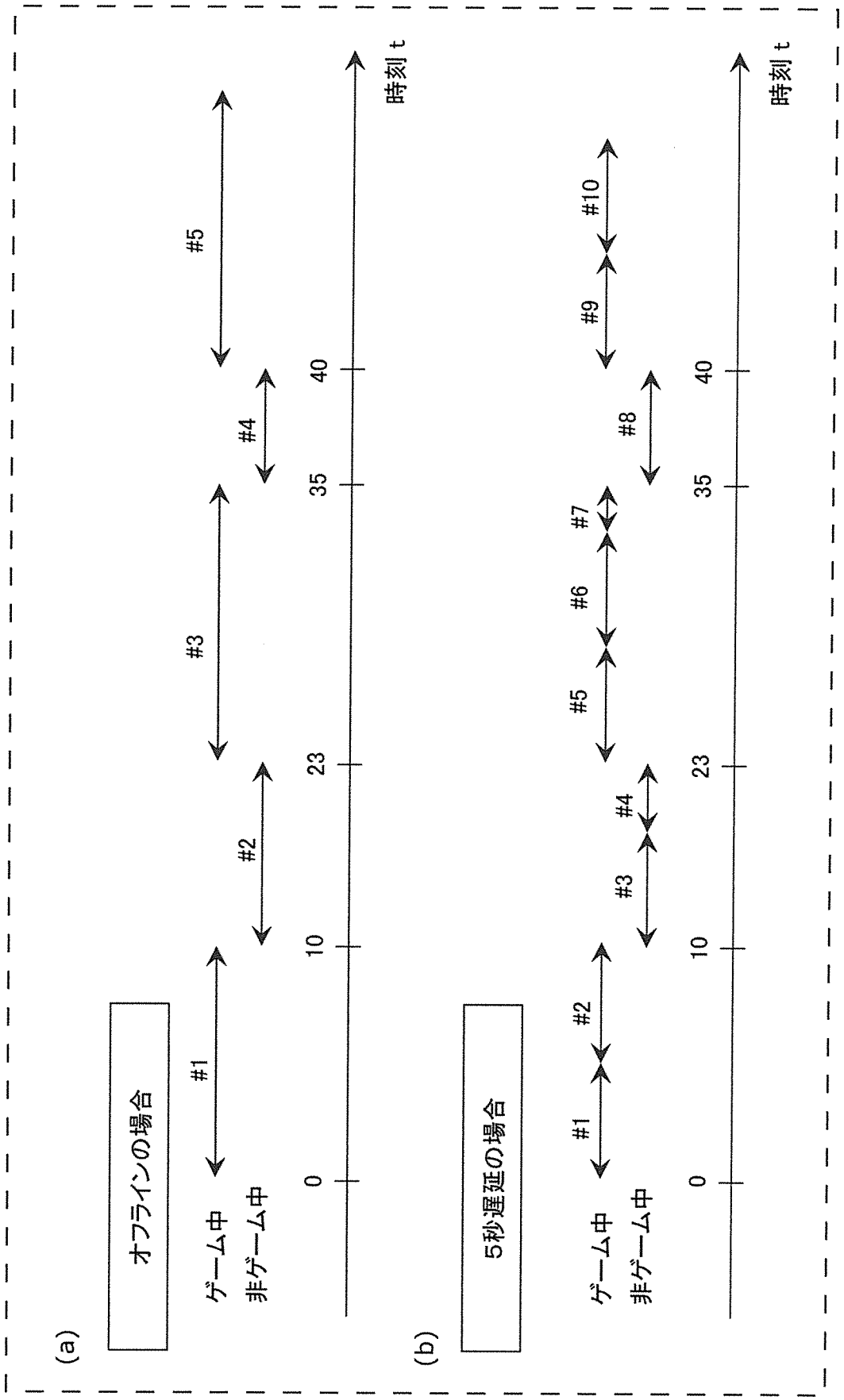
[図35]


ゲーム中
非ゲーム中
ズーム/通知
リプレイ映像
ゲーム中
t1
t2
t3
t4
t5
リプレイシーン時間 = (t4-t3)/再生スピード

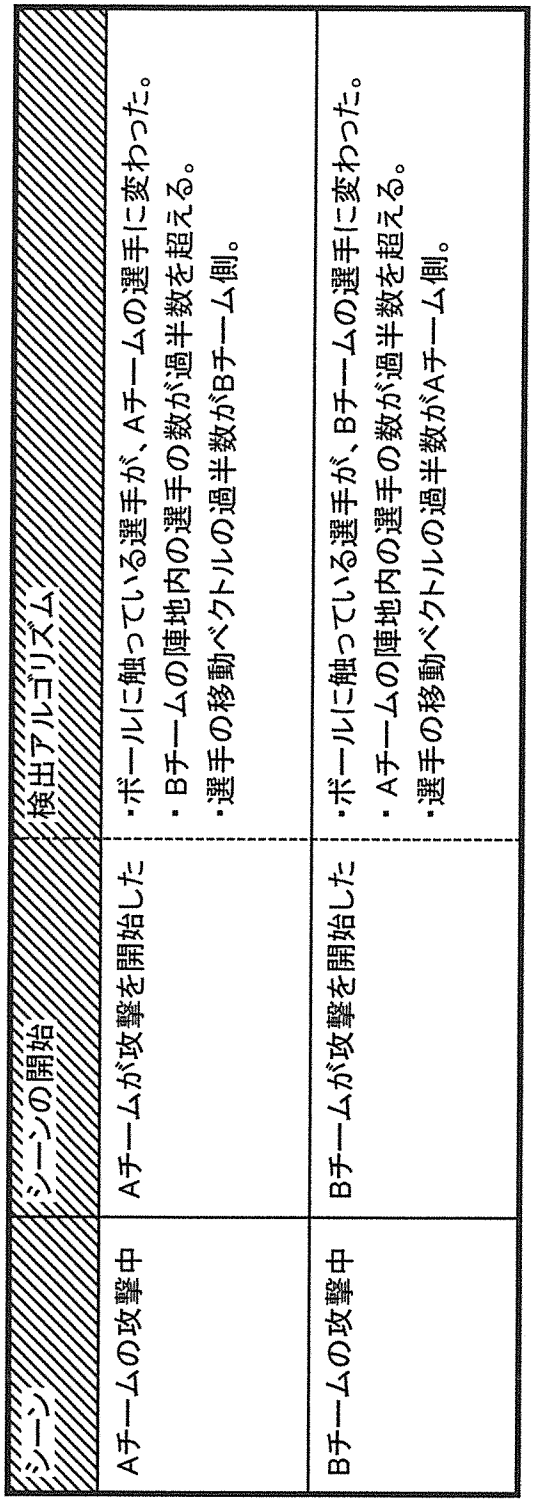
[図36]



[図37]

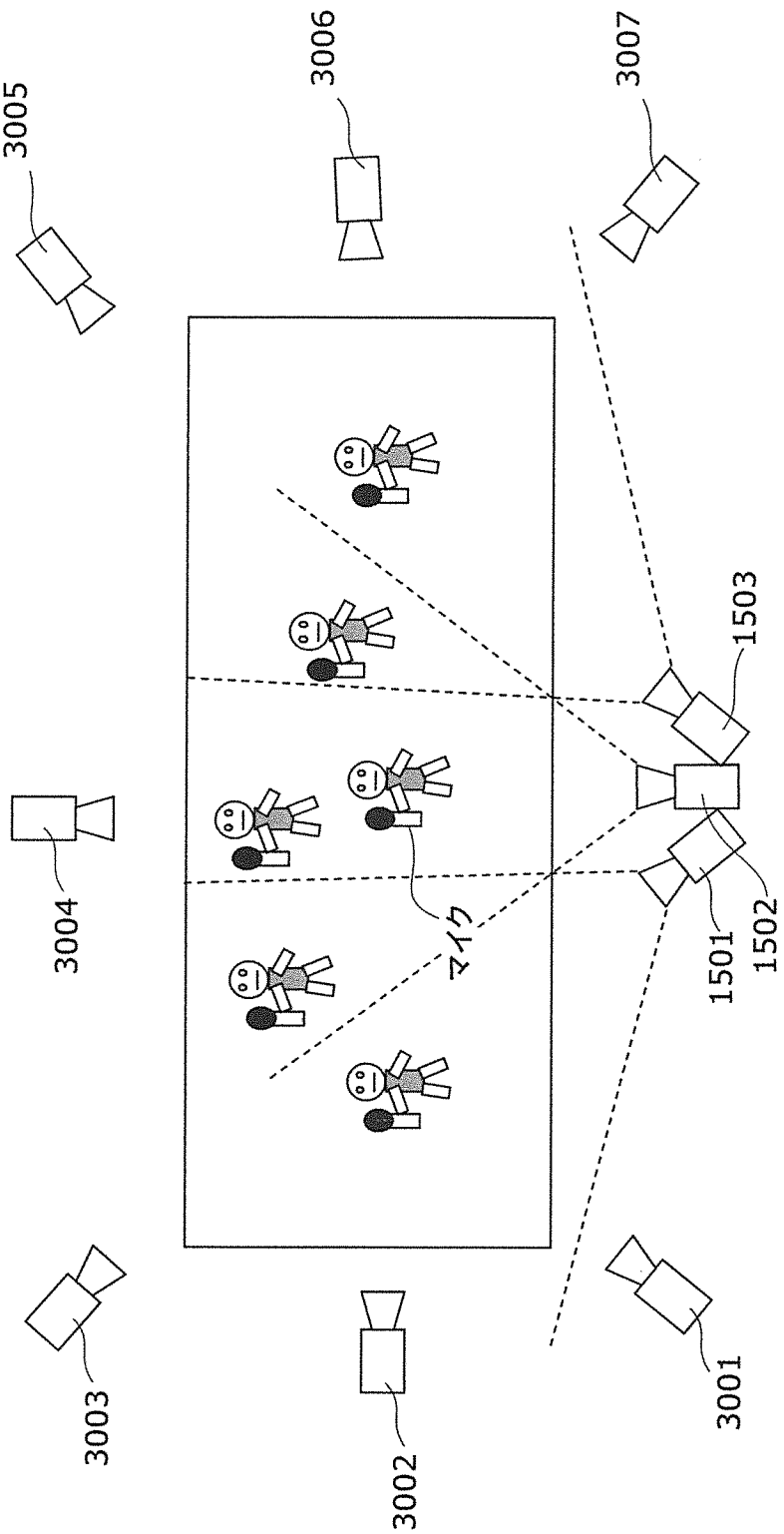

(a)
オフラインの場合
ゲーム中
非ゲーム中
#1
#2
#3
#4
#5
0
10
23
35
40
時刻 t
(b)
5秒遅延の場合
ゲーム中
非ゲーム中
#1
#2
#3
#4
#5
#6
#7
#8
#9
#10
0
10
23
35
40
時刻 t

[図38]

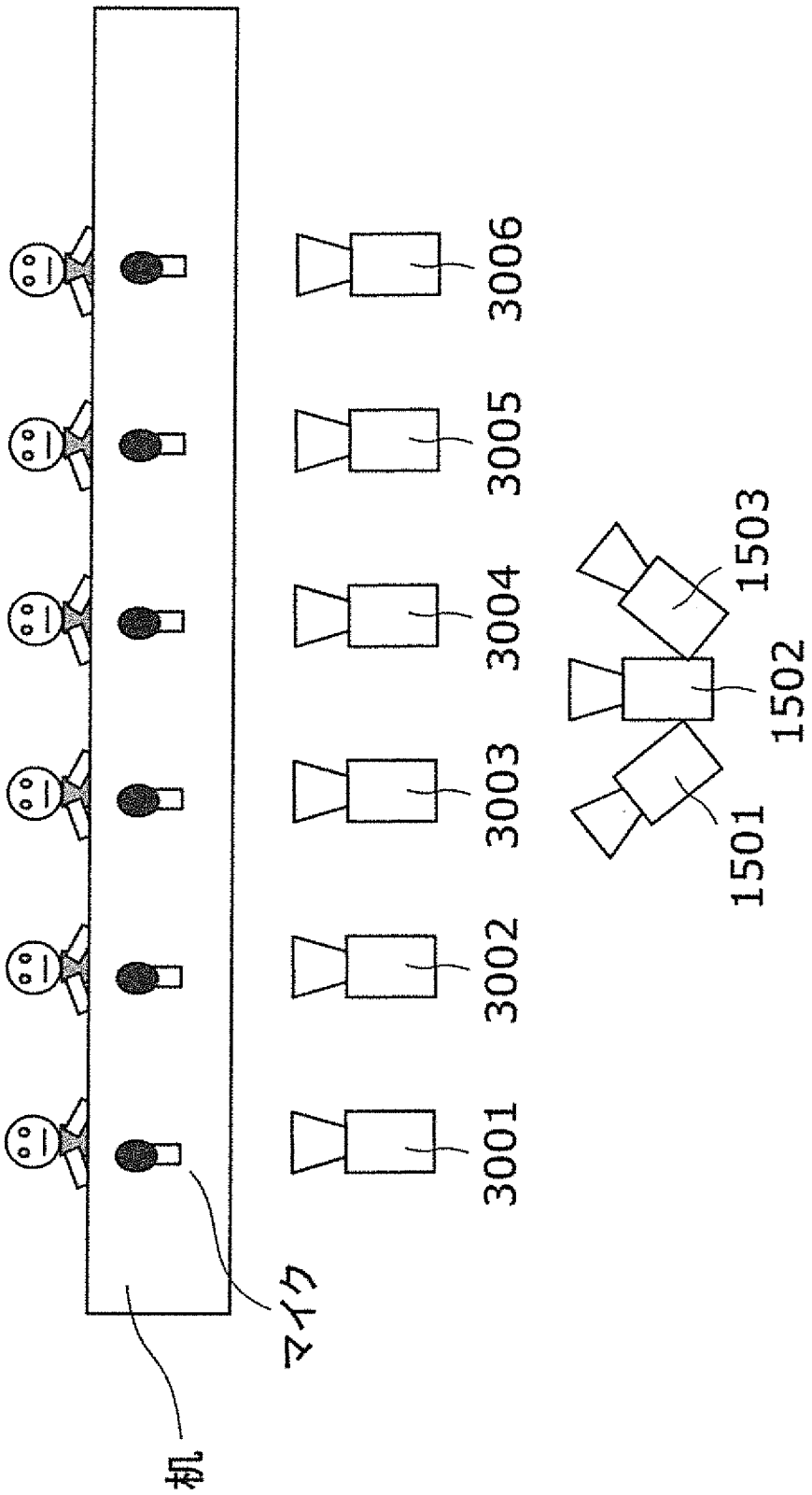


シーン	シーンの開始	検出アルゴリズム
Aチームの攻撃中	Aチームが攻撃を開始した	・ボールに触っている選手が、Aチームの選手に変わった。 ・Bチームの陣地内の選手の数が過半数を超える。 ・選手の移動ベクトルの過半数がBチーム側。
Bチームの攻撃中	Bチームが攻撃を開始した	・ボールに触っている選手が、Bチームの選手に変わった。 ・Aチームの陣地内の選手の数が過半数を超える。 ・選手の移動ベクトルの過半数がAチーム側。

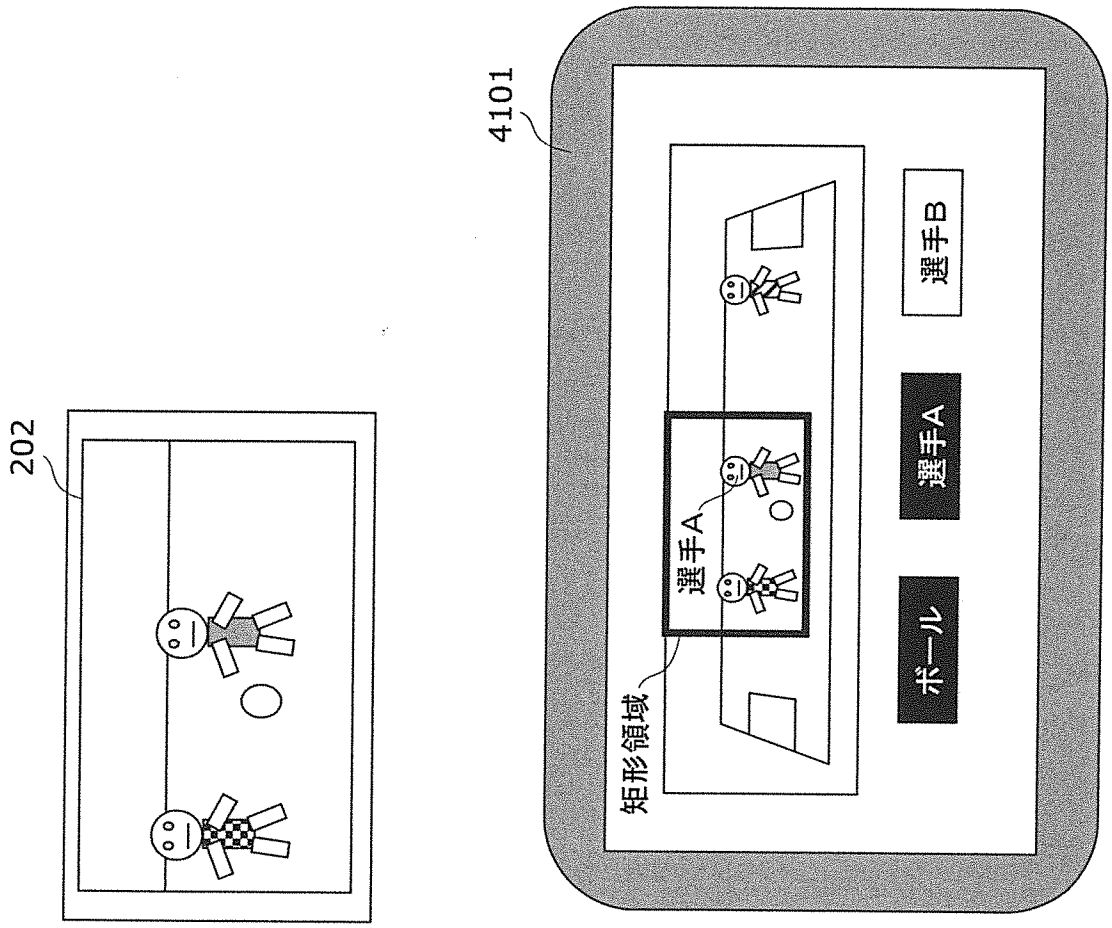
[図39]



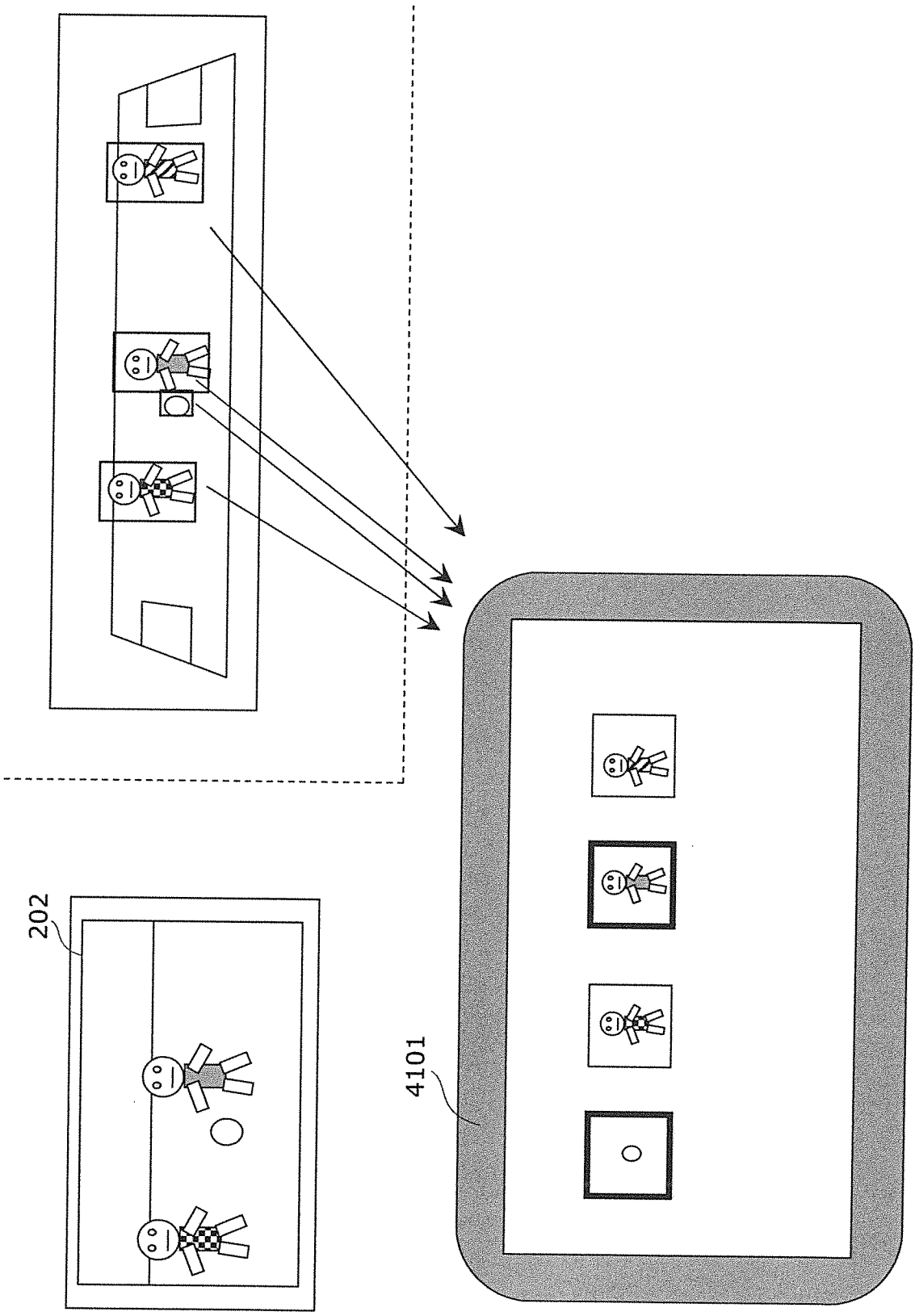
[図40]



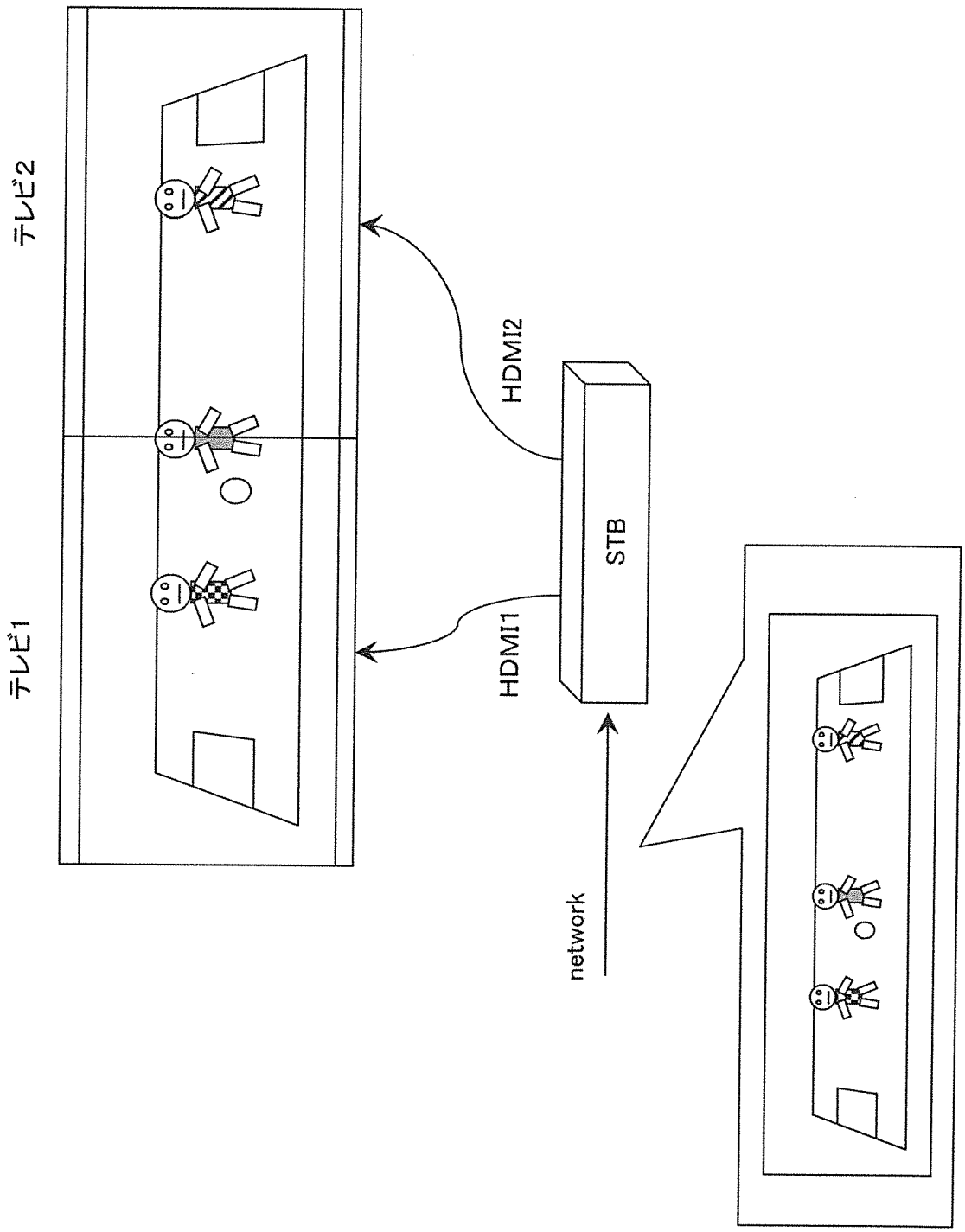
[図41]



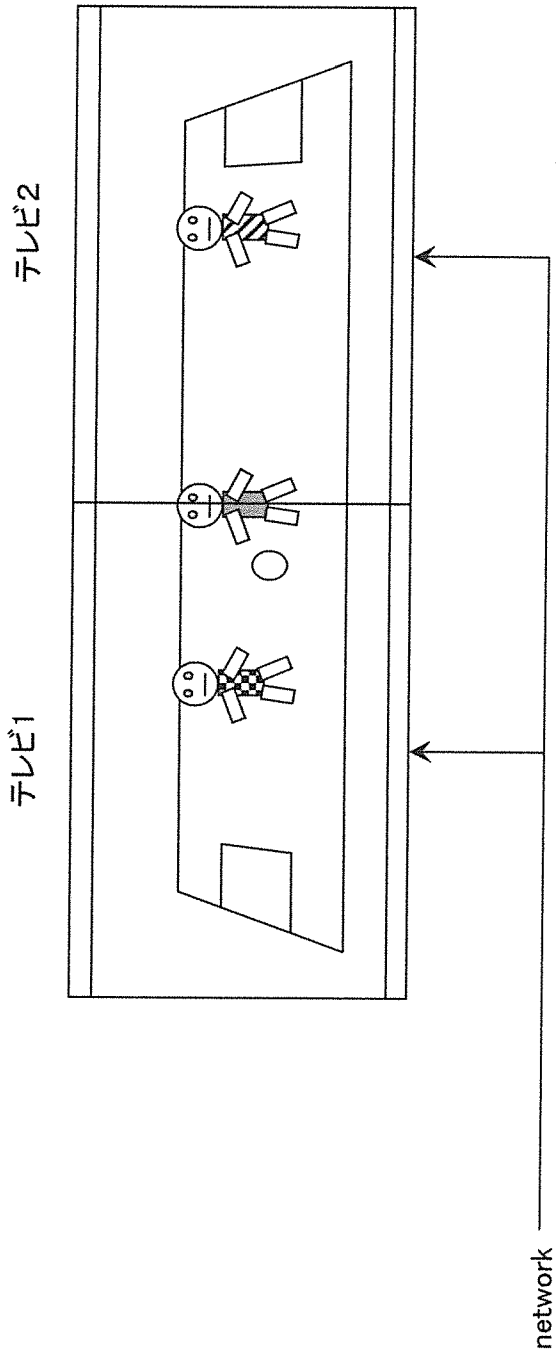
[図42]



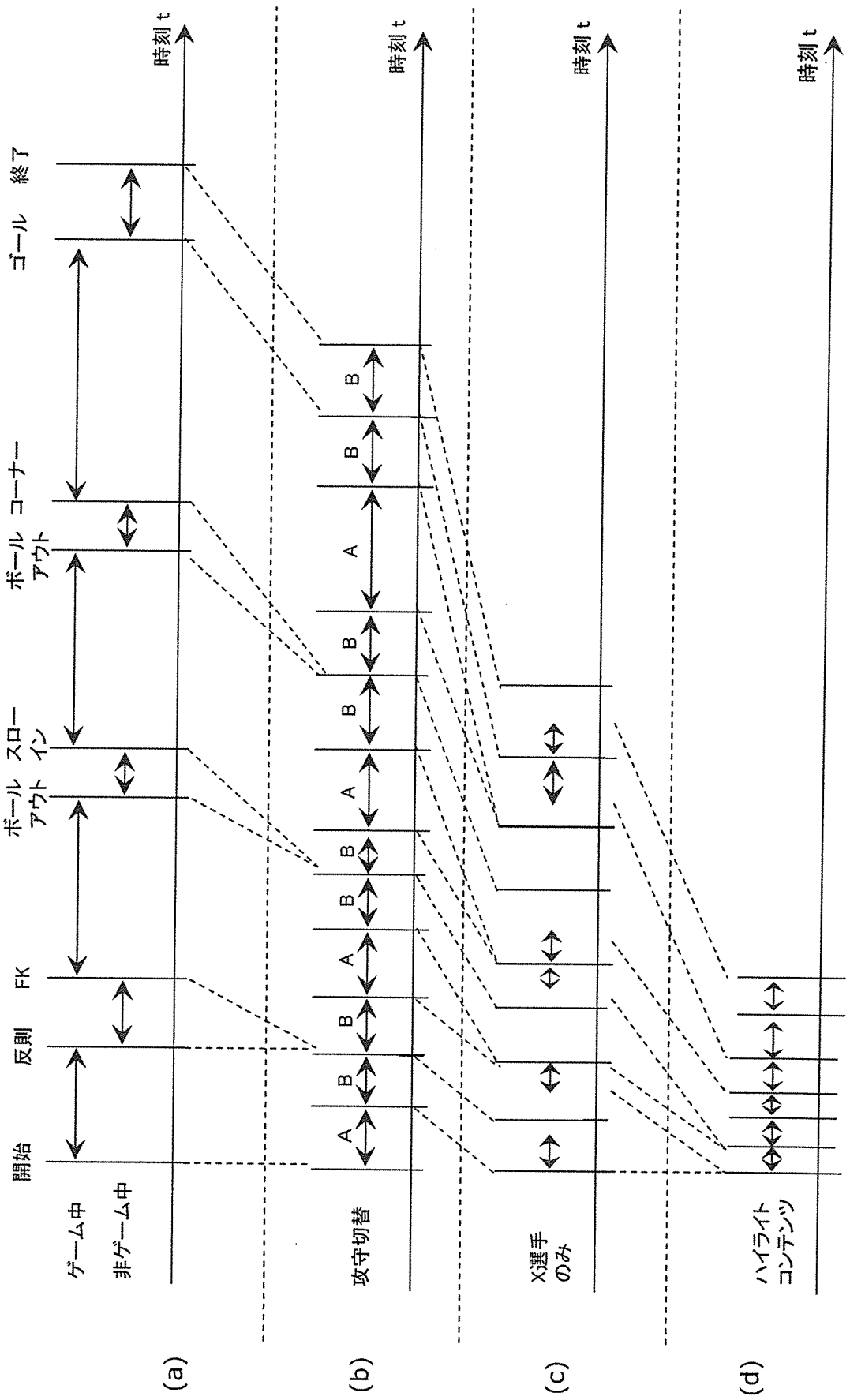
[図43]



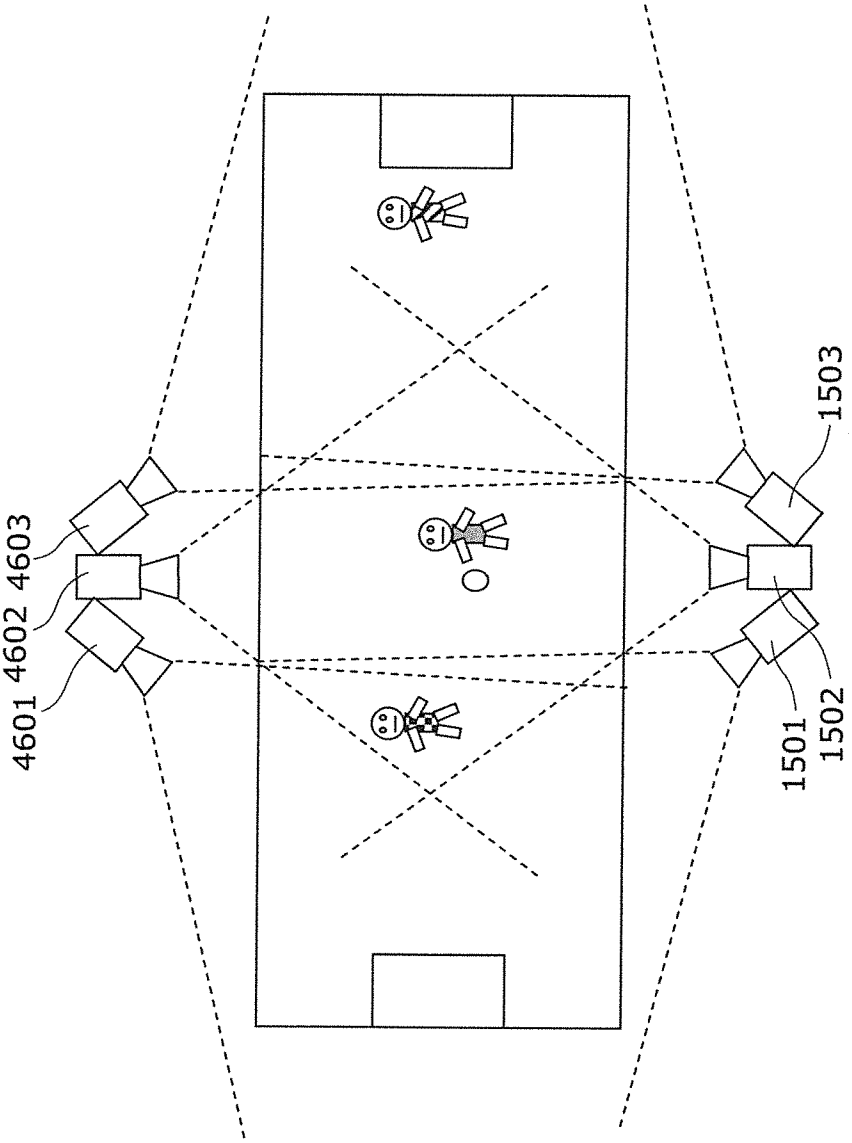
[図44]



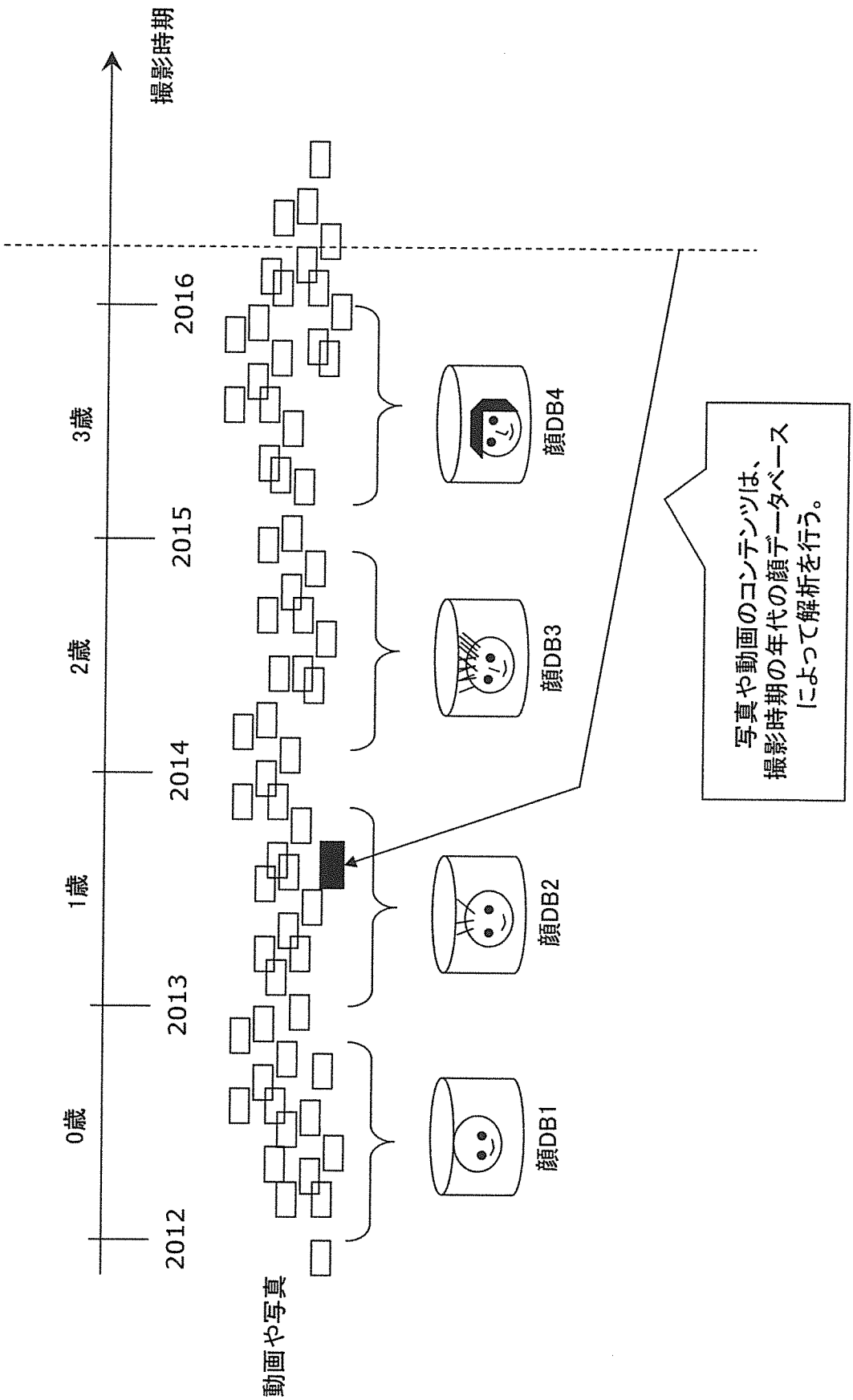
[図45]


(a)
(b)
(c)
(d)
時刻 t
ゲーム中
非ゲーム中
開始
反則
FK
ボールアウト
スローイン
ボールアウト
コーナー
ゴール
終了
攻守切替
X選手のみ
ハイライトコンテンツ
A
B

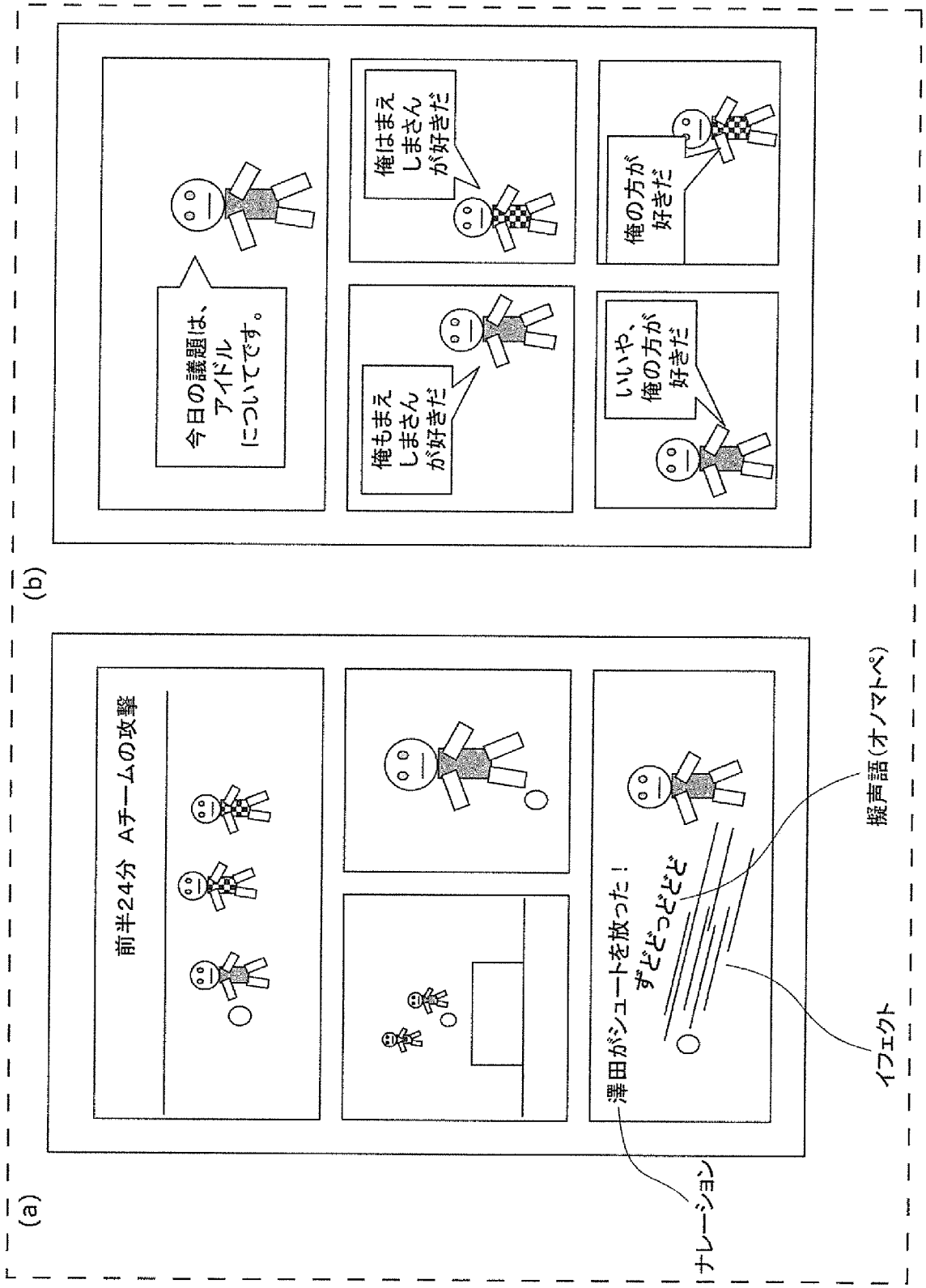
[図46]



[図47]


撮影時期
2012
2013
2014
2015
2016
0歳
1歳
2歳
3歳
動画や写真
顔DB1
顔DB2
顔DB3
顔DB4
写真や動画のコンテンツは、
撮影時期の年代の顔データベース
によって解析を行う。

[図48]


(a)
前半24分 Aチーム の攻撃
澤田がシュートを放った！
ずどどどど
擬声語（オノマトペ）
エフェクト
ナレーション
(b)
今日の議題は、アイドルについてです。
俺はまえしまさんが好きだ
俺の方が好きだ
俺もまえしまさんが好きだ
いいや、俺の方が好きだ

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2013/004742

A. CLASSIFICATION OF SUBJECT MATTER
H04N21/2668(2011.01)i, *H04N5/265*(2006.01)i, *H04N21/2187*(2011.01)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)
H04N21/2668, H04N5/265, H04N21/2187

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)
CSDB(Japanese Patent Office)

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	JP 2005-333552 A (ViewPLUS Inc.), 02 December 2005 (02.12.2005), paragraphs [0001], [0032], [0034], [0038] to [0046], [0048] to [0051] (Family: none)	1,17-19 2,3 4-16
Y	JP 2008-244922 A (Mitsubishi Electric Corp.), 09 October 2008 (09.10.2008), paragraphs [0001], [0011] to [0021] (Family: none)	2,3

☐ 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
02 October, 2013 (02.10.13)	15 October, 2013 (15.10.13)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１３／００４７４２

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　H04N21/2668(2011.01)i, H04N5/265(2006.01)i, H04N21/2187(2011.01)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.　H04N21/2668, H04N5/265, H04N21/2187

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）

ＣＳＤＢ（日本国特許庁）

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	ＪＰ　２００５－３３３５５２　Ａ（株式会社ビュープラス）２００５．１２．０２，段落［０００１］，［００３２］，［００３４］，［００３８］－［００４６］，［００４８］－［００５１］（ファミリーなし）	１，１７－１９
Ｙ		２，３
Ａ		４－１６

☑　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	国際調査報告の発送日
０２．１０．２０１３	１５．１０．２０１３

国際調査機関の名称及びあて先
日本国特許庁（ＩＳＡ／ＪＰ）
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）　５Ｃ　９１８９
久保　光宏
電話番号　０３－３５８１－１１０１　内線　３５３９

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　２００８－２４４９２２　Ａ（三菱電機株式会社）２００８．１０．０９，段落［０００１］，［００１１］－［００２１］（ファミリーなし）	２，３