

(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
2019年12月5日(05.12.2019)



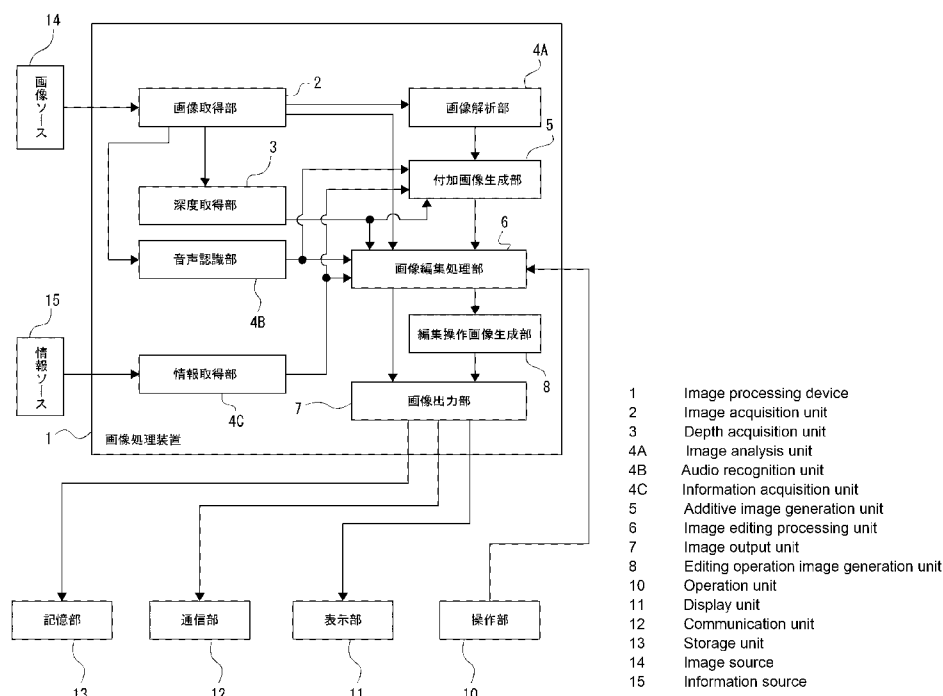
(10) 国際公開番号

WO 2019/230225 A1

- (51) 国際特許分類:
H04N 5/91 (2006.01) G09G 5/36 (2006.01)
G06T 13/80 (2011.01) G09G 5/377 (2006.01)
G09G 5/00 (2006.01)
- (21) 国際出願番号: PCT/JP2019/016198
- (22) 国際出願日: 2019年4月15日(15.04.2019)
- (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 2018-102028 2018年5月29日(29.05.2018) JP
- (71) 出願人: ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒1080075 東京都港区港南1丁目7番1号 Tokyo (JP).
- (72) 発明者: 高橋 正宏 (TAKAHASHI, Masahiro); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP). 柘植 隆弘 (TSUGE, Takahiro); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP). 浅井 聡 (ASAI, Satoshi); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP). 野々山 陽 (NONOYAMA, Yo); 〒1080075 東京都港区港南1丁目7番1号 ソニー株式会社内 Tokyo (JP).
- (74) 代理人: 岩田 雅信, 外 (IWATA, Masanobu et al.); 〒1010032 東京都千代田区岩本町1丁目3番9号 ハクセイビル8階 テクノピア国際特許事務所 Tokyo (JP).

(54) Title: IMAGE PROCESSING DEVICE, IMAGE PROCESSING METHOD, AND PROGRAM

(54) 発明の名称: 画像処理装置、画像処理方法、プログラム



(57) Abstract: In this invention, to achieve image effects, etc., in a moving image that do not feel incongruous, an image processing unit is made to comprise an additive image generation unit for generating an additive image to be added to moving image data and an image editing processing unit for carrying out image editing processing in which depth information for the pixels in the moving image data to have the additive image added thereto is used to add the additive image to the moving image data.

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, 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, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告(条約第21条(3))

(57) 要約: 動画において違和感のない画像エフェクト等を実現する。このため画像処理部は、動画データに付加する付加画像を生成する付加画像生成部と、付加画像を付加する対象とした動画データにおける画素の深度情報を用いて、動画データに付加画像を付加する画像編集処理を行う画像編集処理部とを備えるようにする。

明 細 書

発明の名称：画像処理装置、画像処理方法、プログラム

技術分野

[0001] 本技術は動画に対する画像編集を行うための画像処理装置、画像処理方法、プログラムに関する。

背景技術

[0002] 各種の画像編集を行う画像処理技術が知られている。

特許文献1には、複数の画像を配置して合成する際に位置関係を考慮してレイアウトを決める処理を行うことが記載されている。

先行技術文献

特許文献

[0003] 特許文献1：特開2010-79570号公報

発明の概要

発明が解決しようとする課題

[0004] 映画やドラマなどのコンテンツとしての動画において、背景を変更したり画像エフェクトを加えるのは一般的であるが、これら処理は特殊な撮像手法であったり編集ソフトを駆使して実現することが通常である。そしてこのような動画編集作業は、複雑で時間を要するだけでなく作業者にも高度なスキルを必要とする。

またスポーツ映像などにも画像エフェクトは有効である。そして画像エフェクトをリアルタイムに実現する上ではモーションキャプチャを応用することが主流だが、マーカーを使用するなど高度な処理が必要であり容易ではない。

そこで本開示では、奥行情報を活用することで簡易且つ有効な動画編集を実現することを目的とする。

課題を解決するための手段

[0005] 本技術に係る画像処理装置は、動画データに付加する付加画像を生成する

付加画像生成部と、付加画像を付加する対象とした動画データにおける画素の深度情報を用いて、動画データに付加画像を付加する画像編集処理を行う画像編集処理部と、を備える。

この場合、動画としての複数フレームの画像データとして、画素について被写体までの距離（奥行き）の値となる深度情報を有する画像データを処理対象とする。なお、深度情報は、画像データを構成する1画素毎に付加されていても良いし、カラー画素としての画素ブロック毎（例えばR画素、G画素、B画素の3画素の単位）などでもよいに付加されていても良い。或いは4画素、9画素など、画素ブロック毎に深度情報が付加されていても良い。

画像処理装置は、このような画素（画素ブロック）毎に、深度情報を用いて付加画像を動画に付加する画像編集を行う。

[0006] 上記した本技術に係る画像処理装置においては、前記画像編集処理部は、動画データの被写体の深度情報と前記付加画像に設定された深度情報とを用いて被写体と付加画像の前後関係を反映した画像編集を行うことが考えられる。

即ち付加画像について設定された深度情報と被写体（動画データの各フレームの画素）の深度情報により、付加画像と被写体の前後関係が規定される。

[0007] 上記した本技術に係る画像処理装置においては、前記付加画像は、動画データにおける当該付加画像の挿入位置における深度情報に応じてサイズ調整されて動画データに合成されることが考えられる。

即ち付加画像生成部もしくは画像編集処理部は、付加画像を深度情報に基づいてサイズ調整する。これにより付加画像自体が挿入位置の遠近に応じたサイズ（遠近法に沿った状態）で表現されるようにする。

[0008] 上記した本技術に係る画像処理装置においては、前記付加画像は、動画データにおける動体検出結果を用いて生成されるエフェクト画像であることが考えられる。

例えば動画内で動きのある被写体の動きの追尾や動きの軌跡などを表現す

るエフェクト画像を付加画像として生成し、動画に合成する。

[0009] 上記した本技術に係る画像処理装置においては、前記エフェクト画像には動体に応じた深度情報が設定され、前記画像編集処理部は、動画データの被写体の深度情報と前記エフェクト画像に設定された深度情報とを用いて被写体と付加画像の前後関係を反映した画像編集を行うことが考えられる。

エフェクト画像について、動体の動きに応じた深度情報を設定し、エフェクト画像が動体の深度に対応した画像となるようにする。

[0010] 上記した本技術に係る画像処理装置においては、前記エフェクト画像は、動体に応じた深度情報が設定され、深度情報に応じてサイズ調整されて動画データに合成されることが考えられる。

即ち付加画像生成部もしくは画像編集処理部は、エフェクト画像を深度情報に基づいてサイズ調整する。これによりエフェクト画像自体が挿入位置（つまり各フレームにおける動体の位置）の遠近に応じたサイズ（遠近法に沿った状態）で表現されるようにする。

[0011] 上記した本技術に係る画像処理装置においては、前記付加画像は、動画データにおける動体の軌跡を表すエフェクト画像であることが考えられる。

例えば動画内で動きのある被写体の動き軌跡を表現するエフェクト画像を付加画像として生成し、動画に合成する。

[0012] 上記した本技術に係る画像処理装置においては、前記付加画像は、動画データにおける動体の動作態様又は動作種別を表すエフェクト画像であることが考えられる。

動画内で動体の動作として各種の態様や種別があるが、それらを表現するエフェクト画像を付加画像として生成し、動画に合成する。

[0013] 上記した本技術に係る画像処理装置においては、前記付加画像は、動画データにおける音声検出結果を用いて生成されるテロップ画像であることが考えられる。

動画内の人物の発話などを検出し、そのテロップ画像を付加画像として生成し、動画に合成する。

[0014] 上記した本技術に係る画像処理装置においては、前記テロップ画像は、動画データからの音声認識に基づいてテキストデータを表示する画像であることが考えられる。

動画内で発話される音声を認識しテキストデータを取得する。そしてそのテキストデータをテロップ表示する。

[0015] 上記した本技術に係る画像処理装置においては、前記テロップ画像は、動画データから認識される音声の音量に応じて異なる態様の画像とされることが考えられる。

例えばテロップ画像は、認識した音声の音量、声量に応じてサイズ、フォントなどが異なるなど、異なる表示態様とする。

[0016] 上記した本技術に係る画像処理装置においては、前記テロップ画像は、動画データにおける発話者の感情情報に応じて異なる態様の画像とされることが考えられる。

例えば認識した音声から発話者の感情を推定する。そして感情（怒り、喜び、驚き）などにより、サイズ、フォントなどが異なるなど、異なる表示態様とする。

[0017] 上記した本技術に係る画像処理装置においては、前記付加画像は、取得情報を用いて生成される情報提示画像であることが考えられる。

動画に関して情報ソースから情報を取得し、その情報を提示する画像を付加画像として生成する。そして深度情報を用いて動画に合成する。

[0018] 上記した本技術に係る画像処理装置においては、動画データに付加した前記付加画像を編集するためのユーザインタフェース画像として、動画の時間軸と深度軸が表現される編集画像を生成する編集操作画像生成部を備えることが考えられる。

付加画像の合成状態を編集したい場合に、ユーザが編集に用いる編集操作画像を生成し、表示されるようにする。この場合に編集画像は、動画の時間軸と深度軸が表現されるものとする。

[0019] 上記した本技術に係る画像処理装置においては、前記編集操作画像は、一

方の軸が時間軸で他方の軸が深度軸とされる表示領域を有することが考えられる。

例えば1つの領域として、例えば横方向が時間軸、縦方向が深度軸となるような領域を設けて編集用の画面が形成されるようにする。

[0020] 上記した本技術に係る画像処理装置においては、前記編集操作画像は、時間軸方向の情報と、ある時点の画像が同時に表示される画像であることが考えられる。

例えば時間軸方向の情報がタイムラインとして表示されつつ、その一部の時点の画像が同時に表示されるようにする。

[0021] 上記した本技術に係る画像処理装置においては、前記編集操作画像は、時間軸方向の情報を示す画像上で、ある時点を指定することで、当該時点の画像が時間軸方向の情報と同時に表示される画像であることが考えられる。

例えば時間軸方向の情報がタイムラインとして表示されつつ、或る時点を指定することで、タイムラインの情報が提示されたまま、当該時点の画像が表示されるようにする。

[0022] 上記した本技術に係る画像処理装置においては、前記編集操作画像は、時間軸方向のある時点の画像を立体画像として表示する画像であることが考えられる。

例えば時間軸方向の情報がタイムラインとして表示されつつ、タイムライン上で指定した或る時点の画像が立体画像として表示されるようにする。

[0023] 本技術に係る画像処理方法は、動画データに付加する付加画像を生成する手順と、付加画像を付加する対象とした動画データにおける画素の深度情報を用いて、動画データに付加画像を付加する画像編集処理を行う手順とを情報処理装置が実行する画像処理方法である。

本技術に係るプログラムは、これらの各手順に相当する処理ステップを情報処理装置に実行させるプログラムである。

これらの方法又はプログラムにより、多様な画像編集を容易に実現できる画像処理装置を実現できる。

発明の効果

[0024] 本技術によれば深度情報を用いることで、動画に対する画像合成や画像エフェクトとしての付加画像が、動画内において自然に加えられる状態となり、品質の高い編集動画を容易に生成できる。

なお、ここに記載された効果は必ずしも限定されるものではなく、本開示中に記載されたいずれかの効果であってもよい。

図面の簡単な説明

[0025] [図1]本技術の実施の形態の画像処理装置のブロック図である。

[図2]実施の形態の処理対象の画像の深度情報の説明図である。

[図3]実施の形態の画像処理装置を実現する情報処理装置のブロック図である。

[図4]第1の実施の形態の動体に対するエフェクト画像の説明図である。

[図5]第1の実施の形態の動体に対するエフェクト画像の説明図である。

[図6]第1の実施の形態の動体に対するエフェクト処理例のフローチャートである。

[図7]第1の実施の形態のエフェクト画像設定処理のフローチャートである。

[図8]第1の実施の形態の動体に対する他のエフェクト処理例のフローチャートである。

[図9]第1の実施の形態のエフェクト画像設定処理のフローチャートである。

[図10]第2の実施の形態のテロップ画像の説明図である。

[図11]第2の実施の形態の音声認識に基づく編集処理例のフローチャートである。

[図12]第2の実施の形態のテロップ画像生成処理のフローチャートである。

[図13]第3の実施の形態の情報提示画像の説明図である。

[図14]第3の実施の形態の情報提示画像生成処理のフローチャートである。

[図15]第4の実施の形態のタッチパネルを想定した操作インタフェースの説明図である。

[図16]第5の実施の形態の三面図によるプレビューの説明図である。

[図17]第6の実施の形態の操作インターフェースの説明図である。

[図18]第7の実施の形態のタイムライン表示の説明図である。

[図19]第7の実施の形態のプレビュー表示の説明図である。

[図20]第7の実施の形態の縮小表示の説明図である。

[図21]第7の実施の形態の複数プレビュー表示の説明図である。

[図22]第7の実施の形態の正面画表示の説明図である。

[図23]第7の実施の形態の立体画表示の説明図である。

[図24]第7の実施の形態の操作インターフェースで編集するエフェクト画像例の説明図である。

[図25]第7の実施の形態の編集インターフェースの画面遷移の説明図である。

[図26]第7の実施の形態の画像例に対応するタイムライン表示の説明図である。

[図27]第7の実施の形態の画像例に対応するプレビュー表示の説明図である。

[図28]第7の実施の形態の画像例に対応する縮小表示の説明図である。

[図29]第7の実施の形態の画像例に対応する複数プレビュー表示の説明図である。

[図30]第7の実施の形態の画像例に対応する正面画表示の説明図である。

[図31]第7の実施の形態の画像例に対応する立体画表示の説明図である。

[図32]第7の実施の形態の操作インターフェース表示処理のフローチャートである。

発明を実施するための形態

[0026] 以下、実施の形態を次の順序で説明する。

<1. 画像処理装置の構成>

<2. 第1の実施の形態：動体に対する画像エフェクト>

<3. 第2の実施の形態：音声認識に基づくテロップ>

<4. 第3の実施の形態：情報提示画像の付加>

<5. 第4の実施の形態の：編集インターフェース>

< 6. 第 5 の実施の形態の：編集インタフェース>

< 7. 第 6 の実施の形態の：編集インタフェース>

< 8. 第 7 の実施の形態の：編集インタフェース>

< 9. まとめ及び変形例>

[0027] < 1. 画像処理装置の構成>

図 1 は本開示の画像処理装置 1 の機能構成と画像処理装置 1 の周辺構成を示している。画像処理装置 1 は、画像取得部 2、深度取得部 3、画像解析部 4 A、音声認識部 4 B、情報取得部 4 C、付加画像生成部 5、画像編集処理部 6、画像出力部 7、編集操作画像生成部 8 を有している。

また画像処理装置 1 の周辺構成の一例として、操作部 10、表示部 11、通信部 12、記憶部 13、画像ソース 14、情報ソース 15 を示している。

操作部 10、表示部 11、通信部 12、記憶部 13、画像ソース 14、情報ソース 15 のそれぞれは画像処理装置 1 と一体の機器内に設けられてもよいし、別体の機器とされて画像処理装置 1 と有線又は無線通信により接続されるものであっても良い。

[0028] まず画像処理装置 1 の周辺構成について説明する。

操作部 10 は画像編集のためのユーザの各種操作を検知する。この操作部 10 は、実際のキー、スイッチ等の操作子の操作を検知する構成でもよいし、コンピュータ装置におけるマウスやキーボードであったり、音声入力、ジェスチャ入力、非接触入力等の操作を検知する構成でもよい。また操作部 10 は、例えばタブレット、スマートフォン等の情報処理装置において画面上或いはパッド上でのタッチ操作、タップ操作等を検知する構成などでもよい。

[0029] 表示部 11 は、ユーザ（画像処理装置 1 の使用者等）に対して各種表示を行う表示部であり、例えば画像処理装置 1 を含む装置の筐体に設けられた LCD（Liquid Crystal Display）や有機 EL（Electro-Luminescence）ディスプレイ等のディスプレイデバイスとされる。或いは表示部 11 は画像処理装置 1 を有する装置とは別体の表示デバイスとされてもよい。

この表示部 11 ではユーザインタフェースのための各種表示を行う。例えば画像処理装置 1 が編集処理により生成した画像の表示、編集前の動画の表示、上記の操作部 10 として機能する操作アイコン等の操作子画像の表示、その他必要な表示を行う。

[0030] 通信部 12 は他の機器との間での有線又は無線接続による通信を行う。例えば W I F I (Wireless Fidelity: 登録商標) やブルートゥース (登録商標) 等の無線通信規格等の通信方式による通信を行う。或いは通信部 12 は、外部の表示装置、記録装置、再生装置等の間で画像データ (静止画ファイルや動画ファイル) の通信を行うようにしたり、ネットワーク通信部として、例えばインターネット、ホームネットワーク、L A N (Local Area Network) 等の各種のネットワークによる通信を行い、ネットワーク上のサーバ、端末等との間で各種データ送受信を行うものでもよい。

この通信部 12 は、例えば画像処理装置 1 が編集処理により生成した画像を外部機器に送信する。また通信部 12 が、編集処理に用いる画像データを外部から受信するようにしてもよい。

[0031] 記憶部 13 は例えば不揮発性メモリからなり、元の動画データや動画データを編集 (特に付加画像の付加) した編集画像データを記憶する。記憶部 5 の実際の形態は多様に考えられる。例えば記憶部 13 は、装置筐体に内蔵されるフラッシュメモリ等の固体メモリでもよいし、装置に着脱できるメモリカード (例えば可搬型のフラッシュメモリ) と該メモリカードに対して記録再生アクセスを行うカード記録再生部による形態でもよい。また H D D (Hard Disk Drive)、光ディスク及びディスクドライブなどとして実現されることもある。

この記憶部 13 は画像処理装置 1 が生成した編集画像データを記憶する。また画像処理装置 1 が編集処理に用いる画像データを記憶部 13 から読み出すようにしてもよい。

なお、記憶部 13 には、画像処理装置 1 としての処理をコンピュータ装置に実行させるためのプログラムが記憶されてもよい。

[0032] 画像ソース 14 は、画像処理装置 1 が編集処理に用いる画像データの供給元という意味で示している。処理対象の動画としての画像データが画像ソース 14 から画像処理装置 1 に供給される。

例えば画像ソース 14 は動画撮像を行う撮像装置であったり、記憶媒体から動画を読み出して出力する再生装置であったりする。或いは通信部 12 や記憶部 13 が画像ソース 14 に相当する場合もある。

画像ソース 14 によって供給される画像データは、例えばスポーツ等のリアルタイム放送としての撮像動作であってもよいし、過去に撮像又は生成された動画であってもよい。例えば映画、ドラマ等として制作された動画、一般ユーザが制作・投稿した動画、アニメーション動画、コンピュータグラフィックス等で生成された動画など、あらゆる動画が想定される。即ちあらゆる動画が画像処理装置 1 の編集処理対象として考えられる。

[0033] 情報ソース 15 は、画像処理装置 1 に対して何らかの情報の供給元という意味で示している。情報とは、例えば編集対象の動画に対して関連する情報、編集により動画上に付加する情報などが想定される。例えばスポーツ中継の動画における、選手やチームの情報、タイムの情報、技や得点などの情報などがある。

この情報ソース 15 は例えば情報サーバとしてネットワークを介して情報を提供する情報処理装置であってもよいし、各種のセンサ装置であってもよい。

[0034] 続いて画像処理装置 1 の機能構成を説明する。

画像処理装置 1 における画像取得部 2 は編集処理の対象とする画像データを画像ソース 14 から取得する。例えばリアルタイム放送する動画であったり、動画コンテンツとして画像ソース 14 が提供可能な動画の中でユーザが選択した動画などが処理対象となる。

[0035] 深度取得部 3 は、処理対象とした画像について画素の深度情報を検出する。本実施の形態では、画像処理装置 1 が編集処理の対象とする動画としての画像データには、各フレームの画素毎に深度情報 DP が付加されている。こ

の深度情報 DP は、画像の撮像時の被写体の奥行き位置の情報である。

例えば 1 フレームの画像データが図 2 のように、0 行～ n 行、0 列～ m 列の画素（ピクセル）配置で構成され、 $\{(m+1) \times (n+1)\}$ 個の画素 $PX0-0$ 、 $PX0-1 \cdots PXn-m$ で形成されているとする。

各画素 $PX0-0$ 、 $PX0-1 \cdots PXn-m$ には、深度情報 $DP0-0$ 、 $DP0-1 \cdots DPn-m$ が付加されている。それぞれ、その画素の被写体の撮像時の奥行き情報である。深度取得部 3 は、処理対象とした画像から深度情報 $DP0-0$ 、 $DP0-1 \cdots DPn-m$ を検出する。

なお、この図 2 に示す各画素 $PX0-0 \sim PXn-m$ は、それぞれが R 画素、G 画素、B 画素のような色別の 1 つの画素と考えても良いし、R 画素、G 画素、B 画素をまとめた 1 つのカラー画素と考えてもよい。つまり深度情報 DP ($DP0-0 \sim DPn-m$) は、個々の画素単位で付されても良いし、カラー画素ユニット単位で付加されていても良い。さらに、図示しないが、複数の画素（又は複数のカラー画素）をまとめた画素ブロック毎に深度情報 DP が付加されていても良い。

撮像画像に被写体の深度情報が付加されるようにする技術は既知であるため詳細は省略する。

なお、深度取得部 3 は、処理対象の画像データについて他の機器やセンサで検出された深度情報や、あらかじめ登録された深度情報などを受信取得してもよい。また深度取得部 3 はユーザが任意に設定した深度情報を検出するものでもよい。

深度取得部 3 が検出した画素の深度情報は画像編集処理部 6 や付加画像生成部 5 の処理で用いられる。

[0036] 図 1 の画像解析部 4 A は、画像取得部 2 が取得した画像データの解析を行い、画像内の被写体の認識、動体の認識、背景の認識、動体の動作種別や動作態様の判定、人物の識別、人物の表情、感情の推定、動体の動きの推定など、付加画像生成部 5 での付加画像生成に必要な解析処理を行う。付加画像とは画像編集処理によって動画に付加する画像のことである。

[0037] 音声認識部 4 B は、画像取得部 2 が取得した処理対象の動画としての画像データにおける音声データを入力し、音声認識・解析を行う。そして音声内容に関する情報を付加情報生成のために付加画像生成部 5 に出力する。また音声認識部 4 B は検出した発話のタイミングや音声認識に基づく付加画像表示を行う期間（フレーム期間）等の情報を画像編集処理部 6 に出力する。

後述の第 2 の実施の形態の処理で説明するが、例えば音声認識部 4 は例えば動画とともに収録された音声内で、被写体人物が発話した音声を判別、解析して、発話内容をテキストデータとすることなどを行う。もちろん、人の発話でない周囲の音に応じた擬音データを生成することも考えられる。例えば風の音を認識して「ヒュー」などと擬音のテキストデータを生成するなどである。

[0038] 情報取得部 4 C は情報ソース 1 5 から情報を取得する。取得する情報とは付加画像生成部 5 での付加画像生成に用いる情報である。

例えば情報取得部 4 C は、スポーツ中継の動画に対しては、上述した選手やチームの情報、タイム情報、試合のスコア情報などを取得し、付加画像生成部 5 や画像編集処理部 6 に供給する。

[0039] 付加画像生成部 5 は、画像編集処理により動画に付加する付加画像を生成する。実施の形態では、付加画像として、エフェクト画像、テロップ画像、情報提示画像を生成する例を述べる。

例えば第 1 の実施の形態の場合、付加画像生成部 5 は、画像解析部 4 A による動体認識に応じたエフェクト画像を生成する。

第 2 の実施の形態の場合、付加画像生成部 5 は、音声認識部 4 B による音声認識結果に応じたテロップ画像を生成する。

第 3 の実施の形態の場合、付加画像生成部 5 は、情報取得部 4 C によって取得した情報を示す情報提示画像を生成する。

[0040] 画像編集処理部 6 は、画像取得部 2 が取得した画像データに対する編集処理を行う。本実施の形態の場合、特に画像編集処理部 6 は、元の動画上に、付加画像生成部 5 が生成した付加画像（エフェクト画像、テロップ画像、情

報提示画像等）を付加する画像編集を行う。またこの画像編集に際して画像編集処理部 6 は、深度取得部 3 からの各フレームの画素の深度情報 $DP_0 \sim DP_{n-m}$ を用いる。

[0041] 画像出力部 7 は、画像編集処理部 6 が作成した編集画像としての画像データを出力する。すなわち画像出力部 7 は、編集処理により作成された編集後の動画を表示部 11 に出力して表示させる。

また画像出力部 7 は、編集後の画像データを通信部 12 に出力して外部機器に送信させることもできる。

また画像出力部 7 は、編集後の画像データを記憶部 13 に出力して記憶媒体に記憶させることもできる。

[0042] 編集操作画像生成部 8 は、ユーザの手動編集を可能とするインタフェース画像を生成する。

本実施の形態では、画像編集処理部 6 は、画像取得部 2 が取得した動画について、ユーザの付加画像の選択などに応じて、リアルタイムで自動的に編集と行い、編集結果の動画が画像出力部 7 から出力される。

但し、付加画像の合成は必ずしもリアルタイムに限らないし、付加画像の合成編集内容もユーザが調整したり変更してもよい。

そこで編集操作用の画像が用意される。編集操作画像生成部 8 は、このユーザの操作のための画像を生成して、例えば画像出力部 7 を介して表示部 11 で表示されるようにしている。

編集操作画像上でのユーザの操作は、操作部 10 の機能（例えば画面上のタッチ操作）として画像編集処理部 6 に認識されて、編集処理に反映される。また編集操作画像生成部 8 はユーザ操作に応じて編集操作画像の切替等を行う。

[0043] 例えば以上の図 1 のような機能構成を備えることで、実施の形態の画像処理装置 1 は、後述する各種の画像編集を行い、編集画像を生成することができる。

このような画像処理装置 1 は、例えば図 3 のようなハードウェア構成の情

報処理装置１００で実現される。

[0044] 図３に示すように情報処理装置１００は、ＣＰＵ（Central Processing Unit）１５１、ＲＯＭ（Read Only Memory）１５２、ＲＡＭ（Random Access Memory）１５３を有して構成される。

ＣＰＵ１５１は、ＲＯＭ１５２に記憶されているプログラム、または記憶部１５９からＲＡＭ１５３にロードされたプログラムに従って各種の処理を実行する。ＲＡＭ１５３にはまた、ＣＰＵ１５１が各種の処理を実行する上において必要なデータなども適宜記憶される。

ＣＰＵ１５１、ＲＯＭ１５２、およびＲＡＭ１５３は、バス１５４を介して相互に接続されている。このバス１５４にはまた、入出力インタフェース１５５も接続されている。

[0045] 入出力インタフェース１５５には、液晶パネル或いは有機ＥＬパネルなどよりなるディスプレイ１５６、キーボード、マウスなどよりなる入力部１５７、スピーカ１５８、ＨＤＤなどより構成される記憶部１５９、通信部１６０などが接続可能である。

[0046] 図１の表示部１１が情報処理装置１００と一体機器の場合、ディスプレイ１５６は表示部１１を意味する。もちろん表示部１１は情報処理装置１００とは別体機器とされてもよく、その場合、ディスプレイ１５６は入出力インタフェース１５５に接続される別体機器とされればよい。

[0047] 図３の入力部１５７は、情報処理装置１００を使用するユーザが用いる入力デバイスを意味するが、例えば図１の操作部１０としての機能も含む。

通信部１６０は、インターネットを含むネットワークを介しての通信処理や、周辺各部の機器との間の通信を行うもので、例えば図１の通信部１２としての機能も含む。

[0048] 入出力インタフェース１５５にはまた、必要に応じてドライブ１６１が接続され、メモリカード１６２が装着され、メモリカード１６２から読み出されたコンピュータプログラムが、必要に応じて記憶部１５９にインストールされたり、ＣＰＵ１５１で処理したデータが記憶される。もちろんドライブ

161は、磁気ディスク、光ディスク、光磁気ディスク等のリムーバブル記憶媒体に対する記録再生ドライブとされてもよい。

ドライブ161及びメモリカード162は図1の記憶部13としての機能も持つ。

[0049] この図3のようなハードウェア構成において実施の形態の画像処理装置1としての処理、即ち図1の画像取得部2、深度取得部3、画像解析部4A、音声認識部4B、情報取得部4C、付加画像生成部5、画像編集処理部6、画像出力部7、編集操作画像生成部8としての処理を行うことができる。

即ちこれらの処理はCPU151で起動されるソフトウェアにより実現される。そのソフトウェアを構成するプログラムは、ネットワークからダウンロードされたり、リムーバブル記憶媒体から読み出されたりして図3の情報処理装置100にインストールされる。或いはそのプログラムが記憶部159としてのHDD等に予め記憶されていてもよい。

そしてCPU151において当該プログラムが起動されることで、詳しくは後述するように各種の画像編集処理を行うことができるようにされる。

[0050] なお、実施の形態の画像処理装置1は、図3のようなハードウェア構成の情報処理装置（コンピュータ装置）100が単一で構成されることに限らず、複数のコンピュータ装置がシステム化されて構成されてもよい。複数のコンピュータ装置は、LAN等によりシステム化されていてもよいし、インターネット等を利用したVPN（Virtual Private Network）等により遠隔地に配置されたものでもよい。複数のコンピュータ装置には、クラウドコンピューティングサービスによって利用可能なコンピュータ装置が含まれてもよい。

またこの図3の情報処理装置100は、据え置き型、ノート型等のパーソナルコンピュータ、タブレット端末やスマートフォン等の携帯端末として実現できる。さらには情報処理装置100としての機能を有するテレビジョン装置、モニタ装置、画像編集機器、撮像装置等の電子機器でも、本実施の形態の画像処理装置1を搭載することができる。

[0051] < 2. 第 1 の実施の形態：動体に対する画像エフェクト >

画像処理装置 1 による動画に対して付加画像を付加する編集処理の第 1 の実施の形態として、動体についてのエフェクト画像を付加する処理について説明する。

動体についてのエフェクト画像とは、例えば、

- ・ 動画に表れる動体についての動きを強調する画像
- ・ 動画に表れる動体の動きの軌跡を示す画像
- ・ 動画に表れる動体の動きの種別を表現する画像
- ・ 動画に表れる動体の動きの態様を表現する画像
- ・ 動画に表れる動体の動きの予測を表示する画像
- ・ 動画に表れる複数の動体の別を表現する画像
- ・ これらの複合的な画像

などが考えられる。もちろん以上は例示で、エフェクト画像がこれらに限られるものではない。

また動体とは、動画上で動きの或る被写体オブジェクトであり、人、動物、物、これらの一部（例えば人の手など）などが想定される。

[0052] このような動体についてのエフェクト画像を加えた画像の例を図 4，図 5 に示す。

図 4 A から図 4 F はフィギュアスケートの動画のシーンを示している。

図 4 A は演技開始前の画像である。動体 80 としてスケート選手を対象とする。

図 4 B は演技開始前に、選手の動き出しを予測させるエフェクト画像 E F を付加した画像である。

図 4 C は演技中において、選手の身体の移動の軌跡を示すとともに動作を強調するようなエフェクト画像 E F を付加した画像である。

図 4 D は演技中において、スピンを表現するエフェクト画像 E F を付加した画像である。スピンを表す画像として渦巻き状のラインを表示しているが、この場合、選手の後ろ側の部分（選手より奥側となる部分）は、選手の体

に遮られて見えない状態となっている。

図4 Eは、例えば演技終了のポーズに至る動きを示すエフェクト画像E Fを付加した画像である。エフェクト画像E Fは直前からの身体の位置に応じた画像となっており、それは奥行き方向の位置（深度）に応じた大きさとなっている。

図4 Fはエフェクト画像E Fを付加していない画像である。

[0053] 例えばスケート競技の動画において、このようなエフェクト画像E Fを付加する編集を行うことで、より迫力のある動画、美観が向上する動画、選手の動きがわかりやすくなる動画などを実現できる。

[0054] 図5 Aはサッカーのフリーキックのシーンを示している。

この場合、攻撃側選手が蹴ったボールを動体80とし、そのボールの軌跡を表現するエフェクト画像E Fが付加されている。

このエフェクト画像E Fは、各時点での他の選手等との前後関係が反映されている、例えばボールの軌跡において、守備側選手よりも後方となっている部分は、選手の頭部に隠れて見えていない状態となっている。各時点（各フレーム）での各被写体（ボールや選手）の深度情報を用いることで、このようにボールの軌跡を示すエフェクト画像E Fを、実際の被写体との間での正しい前後関係で表示されるようにしている。

またボールの位置に従って、軌跡を表示するエフェクト画像E Fの幅が小さくなっている。これは、深度情報に基づいて、撮像位置からの各時点でのボールの距離（遠近）を表現するものとなる。

またボールの軌跡のエフェクト画像E Fは、色或いは濃さなどが徐々に変化するようにされる。これは例えばボールの回転数の解析結果に応じた色や濃さとすることで、ボールの動作態様を示すものとしている。

[0055] 図5 Bは剣道の試合のシーンを示している。

この場合、両選手の竹刀を動体80とし、竹刀の動きを表現し、かつ強調するエフェクト画像E Fを付加している。

各エフェクト画像E Fは、右側の選手の竹刀と左側の選手の竹刀とで、色

や表示態様を変化させる。これにより、両者の竹刀の動きを、視聴者が、より明確にわかるようにしている。

各竹刀の軌跡を示すエフェクト画像E Fは、各時点（各フレーム）の竹刀や選手の深度情報により前後関係を反映した状態となるようにする。これにより撮像装置の存在する位置からみた前後関係がエフェクト画像E Fにも反映される。

[0056] 以上のようなエフェクト画像E Fを付加する画像処理装置1の処理例を説明する。

なお、以下説明する各実施の形態の処理例は、図1に示した各機能により画像処理装置1において実行される。

[0057] 図6の処理例は、例えば画像ソースに記録された動画コンテンツを取得して、動画編集を行う場合の例としている。なお、リアルタイム中継の際などに適用できる、入力された動画に対してリアルタイムで編集処理を行う例は後に図8で述べる。

[0058] 図6のステップS101で画像処理装置1は、処理対象の動画について被写体の識別を行う。即ち画像内の情報をデプスマップで展開し、被写体である人や物の位置関係を把握する。ここでいうデプスマップとは、例えば図2のような画素毎の深度情報のマップである。デプスマップにより、各被写体の前後関係が把握できる。ここでいう「前後」の「前」とは、撮像装置に近い側、「後」は撮像装置から遠い側（奥側）となる。

なお、各被写体が写っている画素範囲は、輝度情報、色情報、エッジ検出等の解析から判定できるが、各画素の深度情報を用いることで、各被写体が写されている画素範囲をより精度良く判定できる。1つの被写体が写っている画素範囲では、深度の値が、大きくは異ならないためである。

[0059] ステップS102で画像処理装置1は、動体の状態を把握する。即ち画像データに含まれる被写体について、動体、静体の認識を行い、例えば背景と人物などのように各被写体の内容を把握する。

[0060] ステップS103で画像処理装置1は、エフェクトの選定を行う。即ち動

体 80 に対して付加するエフェクト画像 EF としてどのようなものを付加するかを選択する。例えば回転を表現するエフェクト画像 EF、軌跡を表現するエフェクト画像 EF などを選択する。またそのエフェクト画像 EF の色や形、効果継続時間などの詳細を決定する。

以上のエフェクトの選定は、画像処理装置 1 が自動設定してもよいし、ユーザ入力に応じて設定してもよい。或いはある程度は自動設定されて、一部をユーザが設定項目を修正できるようなものとしてもよい。

[0061] ステップ S 104 で画像処理装置 1 は、表示位置の決定を行う。これは画像内の動体 80 に対するエフェクト画像 EF の表示位置（動体 80 を基準とした相対的な位置）を決定する処理となる。各フレーム単位でのエフェクト画像 EF の表示位置は、ここで決めた相対位置に応じてステップ S 107 で決定される。

ステップ S 105 で画像処理装置 1 は、エフェクト対象のフレームの特定を行う。例えば動画内で画像効果を開始するフレームを特定する。或いはフレーム範囲（開始フレームと終了フレーム）を特定しても良い。

これは、ユーザが動画のシーンのうちでフレームを指定入力し、それを画像処理装置 1 が開始フレーム、終了フレームとすることが考えられる。

或いは、画像処理装置 1 が画像データ内から、エフェクト画像 EF を付加する動体 80 を検出し、その動体 80 が表れているシーンを自動選択し、そのシーンの開始フレーム、終了フレームを特定するものとしてもよい。

さらには、画像処理装置 1 が、動体 80 が表れており、かつ動体 80 が動いているフレーム区間を自動的に特定するものとしてもよい。

[0062] 以上のステップ S 105 までで、動画のどの区間でどのようなエフェクト画像 EF の合成が行われるかが決定される。ここまでの処理は、動画に関しては主に図 1 の画像取得部 2、深度取得部 3、画像解析部 4 A の機能により実行される。ユーザ指定に関する設定は操作部 10 からの入力に応じた画像編集処理部 6 の機能による処理となる。

以降、画像処理装置 1 は、動画を構成する各フレーム毎に、ステップ S 1

06～S116の処理を行っていく。

- [0063] ステップS106で画像処理装置1は、処理対象のフレーム画像データを取得（特定）する。例えば最初は、ステップS105で特定したエフェクト対象の区間の最初のフレームの画像データを処理対象とする。

ステップS107で画像処理装置1は付加画像生成部5の機能により、取得したフレーム画像データに対する付加画像となるエフェクト画像EFを設定する。

- [0064] このエフェクト画像EFの設定処理を図7に例示する。

まずステップS201で画像処理装置1はエフェクトとして付加する画像の選択又は生成を行う。これはステップS103で選定したエフェクトの選定に応じた画像を、当該フレームに付加するエフェクト画像EFとして用意する処理である。用意する画像は、エフェクトの種別毎等について予め用意された画像のうちから当該フレームに用いる画像を選択することとしても良いし、当該フレームの画像内容に応じて生成してもよい。

- [0065] ステップS202で画像処理装置1は、エフェクトとして付加するものとして用意した画像について、画像サイズ、形状、色等の画像態様を設定する。例えば画像における動体80の当該フレームでのサイズや動作態様に応じてこれらを設定する。

これらの画像態様が設定された状態で、当該フレームに付加されるエフェクト画像EFが決定されることになる。

なお、1フレームの画像に付加されるエフェクト画像EFは、画素が連続した1つの画像に限らず、複数の画像の場合も当然ある。

- [0066] ステップS203で画像処理装置1は、エフェクト画像EF自体を構成する各画素の深度情報DPeを設定する。例えば、当該フレームでの動体80の深度に基づいて深度情報DPeを設定する。

ステップS204で画像処理装置1は、エフェクト画像EFについての当該フレーム内での挿入領域を設定する。つまり画面上のどこにエフェクト画像EFを付加するかを決定する処理である。これは例えば当該フレームでの

動体 80 の位置や、前のフレームでのエフェクト画像 E F の位置などに基づいて設定する。

なお、エフェクト画像 E F が複数の場合は、それぞれのエフェクト画像 E F に対して挿入領域が設定される。

[0067] 以上のようにエフェクト画像 E F の設定を行ったら、画像処理装置 1 は画像編集処理部 6 の機能により図 6 のステップ S 108 以降の処理を行う。

画像処理装置 1 はステップ S 108 以降で、当該フレームの画像データ上で、設定した挿入領域に含まれる各画素 P X を、順次エフェクト画像 E F の各画素について設定した深度情報 D P e と比較していく。

[0068] まずステップ S 108 で画像処理装置 1 は、フレーム画像データにおいて挿入領域とされた領域内の 1 つの画素 P X を特定し、ステップ S 109 でその画素 P X の深度情報 D P を特定する。

なお、ここでは、設定された挿入領域が 1 つであれば、当然、その挿入領域内の画素の 1 つを処理対象として特定する。

設定された挿入領域が複数であれば、例えば最初の挿入領域内の画素の 1 つを処理対象として特定する。

[0069] ステップ S 110 で画像処理装置 1 は、対象としている画素 P X の深度情報 D P と、その画素 P X の位置に挿入するとされた、エフェクト画像 E F における画素の深度情報 D P e を比較する。

$D P < D P e$ であれば、フレーム画像データにおける当該画素 P X は、エフェクト画像 E F よりも手前側の被写体の画素である。そこで画像処理装置 1 はステップ S 112 に進んで、当該画素 P X を編集後の画像データ上でそのまま表示する画素に設定する。

$D P < D P e$ でなければ、フレーム画像データにおける当該画素 P X は、エフェクト画像 E F の深度以降の奥側の被写体の画素である。そこでこの場合、画像処理装置 1 はステップ S 111 に進んで、画像データ上で、当該画像 P X の画素データ値をエフェクト画像 E F（付加画像）における、その画素位置に対応する画素データ値に置換する。つまりエフェクト画像 E F の画

素データを表示する画素とする。

[0070] ステップS 1 1 3で画像処理装置1は、現在処理対象としている挿入領域内の全ての画素について、以上の処理を終えたか否かを確認する。終わってなければステップS 1 0 8に戻り、次に現在処理している挿入領域内で処理対象とする画素P Xを特定して、同様にステップS 1 0 9～S 1 1 2の処理を行う。

ステップS 1 1 3で、1つの挿入領域の全画素についてステップS 1 0 9～S 1 1 2の処理を終了したと判断したら、画像処理装置1はステップS 1 1 4で、他に設定された挿入領域があるか否かを確認する。

図7のステップS 2 0 4で設定された挿入領域が1つである場合は、以上で当該フレームについての処理を終える。

複数の挿入領域が設定されている場合であって、未処理の挿入領域がある場合は、ステップS 1 1 5で次の挿入領域を選択し、その挿入領域についてステップS 1 0 8～S 1 1 3の処理を行う。

[0071] 画像処理装置1はステップS 1 1 4で全ての挿入領域について処理が終了したことを確認した時点で、現在のフレームについての処理を終え、ステップS 1 1 6に進む。

エフェクトを行う区間としての全フレームについての処理を完了してなければ、画像処理装置1はステップS 1 0 6に戻り、処理対象として次のフレームのフレーム画像データを取得し、そのフレームについて同様にステップS 1 0 7～S 1 1 5の処理を行う。

エフェクト画像E Fを付加する全てのフレームについて以上の処理を終えることで、図6の画像処理を終える。

[0072] この図6の処理を終えた時点で、動画上で例えば図4 B～図4 Eで説明したようにエフェクト画像E Fが付加された編集後の画像データが生成されていることになる。

このように生成された編集画像データは、画像出力部7が出力することで、表示部11に表示されたり、通信部12により外部機器に転送されたり、

記憶部 13 で記憶媒体に記憶される。

[0073] 以上の図 6 の処理では、深度情報 DP、DP_eを用いてエフェクト画像 EF を元のフレーム画像データに合成したり、深度情報 DP、DP_eに応じてサイズ等の表示態様を設定しているため、元の動画のシーンの動体 80 に対して違和感のないエフェクト画像 EF を実現できる。これはグラフィカルな効果を被写体空間に自然になじませるような画像表現ができることにもなる。

[0074] ところで、以上の図 6 の処理は、予め撮像等により生成された動画に対して行う処理例として述べたが、例えば放送等において、撮像した動画に対して、ほぼリアルタイムでエフェクト画像 EF を付加したい場合もある。そのようなリアルタイム性をもった処理例を図 8 に示す。

例えば放送スタッフとしてのオペレータが、ある時点からエフェクト開始を指示する操作を行うことで、画像処理装置 1 が、その時点以降にリアルタイムでエフェクト画像 EF を付加していく処理である。

なお、図 8 の場合、図 6 のステップ S103、S104 で説明したエフェクトの選定や表示位置の決定は予め行われており、そのエフェクトの開始／終了タイミングがオペレータによって指示（或いは何らかの自動制御による指示）されるものとする。例えば図 5A や図 5B のような動体 80 の軌跡を表現するエフェクト画像 EF を、動体の移動軌跡上の位置に表示することなどが決定されているとする。

[0075] ステップ S151 で画像処理装置 1 はエフェクト開始の指示を認識したら、ステップ S152 以降に進む。

その後は、ステップ S155 でエフェクト終了の指示を検知するまで、画像処理装置 1 は、撮像されて供給されてくる各フレームについて、ステップ S152～S154 及び S107～S115 の処理を行う。

[0076] ステップ S152 で画像処理装置 1 は、フレーム画像データを取得する。つまり画像ソース 14（例えば撮像装置）からリアルタイムに入力されてくる 1 フレームを取得する。

なおこのことから、１フレーム期間内に、ステップＳ１５２～Ｓ１５４及びＳ１０７～Ｓ１１５の処理が行われるようにすることが想定される。

[0077] ステップＳ１５３で画像処理装置１は、フレーム画像データにおける被写体の識別を行う。即ち画像内の情報をデプスマップで展開し、被写体である人や物の位置関係を把握する。

ステップＳ１５４で画像処理装置１は、動体の状態を把握する。即ちフレーム画像データに含まれる被写体について、動体、静体の認識を行い、例えば背景と人物などのように各オブジェクトの内容を把握する。

[0078] ステップＳ１０７でエフェクト画像ＥＦの設定を行う。

動体の軌跡を表現するエフェクト画像ＥＦの設定の例を図９Ａに示す。

[0079] 図９Ａの処理例の場合、ステップＳ２１０で画像処理装置１は、前フレームに付加したエフェクト画像ＥＦを取得する。もちろんエフェクト開始から最初のフレームであれば、前フレームに付加したエフェクト画像ＥＦは存在しないため、不要である。

ステップＳ２１１で画像処理装置１は、現フレームでの動体８０の位置を取得する。ここでいう動体８０の位置とは現フレームの画面内の位置、つまり動体８０が写されている画素範囲のことである。

[0080] ステップＳ２１２で画像処理装置１は、今回のフレームのエフェクト画像ＥＦとなる軌跡画像を生成する。具体的には、前フレームのエフェクト画像ＥＦ、つまり前フレームの時点までの軌跡画像に、現フレームの動体８０の位置に応じて画像を追加（軌跡を表現する画像を延長）することで、現フレームの軌跡画像（エフェクト画像ＥＦ）を生成する。エフェクト開始からの最初のフレームの場合は、現フレームの動体８０の位置に応じて軌跡画像を生成する。

[0081] ステップＳ２０３Ａで画像処理装置１は、各フレームでの動体８０の位置を反映して、軌跡画像（エフェクト画像ＥＦ）自体を構成する各画素の深度情報ＤＰｅを設定する。

なお、前フレームまでの軌跡部分は、前フレームまでのエフェクト画像を

そのまま利用することで、既に画素位置や深度は設定されているものとして利用することができる。ここでは、今回軌跡として延長した部分の画素について、現フレームの動体80の画素に付与された深度情報DPを用いて、エフェクト画像EFの深度情報DP_eを設定すればよい。フレームによっては動体が他の被写体に隠れている場合もあるが、その場合は、前フレームまでの動体80の軌跡から、現フレームでの動体の位置や深度を推定すればよい。

[0082] ステップS204で画像処理装置1は、エフェクト画像EFについての当該フレーム内での挿入領域を設定する。つまり画面上のどこにエフェクト画像EFを付加するかを決定する処理である。これは例えば当該フレームでの動体80の位置や、前のフレームでのエフェクト画像EFの位置などに基づいて設定する。

[0083] 以上のようにエフェクト画像EFの設定を行ったら、画像処理装置1は図8のステップS108以降に進む。ステップS108～S115は、図6のステップS108～S115と同様である。即ち画像処理装置1は、当該フレームの画像データ上で、設定した挿入領域に含まれる各画素PXを、順次エフェクト画像EFの各画素について設定した深度情報DP_eと比較していき、前後関係を反映して、当該フレームの元の画像データとエフェクト画像EFのいずれを表示画素とするかを決定していく。

[0084] 以上の処理を1フレームについて行い、次のフレームでは、再び図8の処理をステップS152から実行する。

これにより、動画の進行に従って動体の移動軌跡を描くようなエフェクト画像EFが動画に付加される。

[0085] ところで、ステップS107のエフェクト画像EFの設定は図9Bのように行ってもよい。

図9BのステップS210、S211は図9Aと同様であり、前フレームのエフェクト画像EFを取得し、また現フレームの動体80の位置を取得する。

[0086] ステップS222で画像処理装置1は、現フレームの動体80の情報を取

得する。これは動作態様であったり、動作の種別などの情報である。

動作態様とは、動体 80 の速度、挙動、移動方向、移動姿勢、移動時の状態などである。

例えば図 5 A のボールの場合、ボールの移動速度、回転速度、移動方向なども動作態様の一つとなる。野球のピッチャーの投げる球種も動作態様の情報として想定される。

動作の種別は、動体の移動に関する種別であったり、動体自体の種別や動体を動かす人などの種別が想定される。例えばサッカーの A チームと B チームの試合でボールを蹴った人がいずれのチームであるかとか、剣道の選手の別というのは動体を動かす人の種別の一つである。また例えば人や動物を動体 80 としたときに、動体 80 としての男女の別、年齢層の別、動物の別（例えば犬や猫の別、飛んでいる鳥の種別など）などもここでいう動体 80 の種別の例となる。

[0087] ステップ S 2 2 3 で画像処理装置 1 は、取得した動体 80 の情報（態様・種別）に応じて、エフェクト画像 E F の画像態様を設定する。例えば情報に応じて表示するエフェクト画像 E F の色、輝度、濃度、グラデーションなどを設定する。

そしてステップ S 2 1 2 で、画像態様の設定を反映して軌跡画像、即ち現フレームに付加するエフェクト画像 E F を生成する。

ステップ S 2 0 3 A、S 2 0 4 は図 9 A と同様である。

[0088] 例えばこの図 9 B のように、動体 80 の態様や種別の情報を反映させるようにした軌跡画像（エフェクト画像 E F）を生成することで、視聴者がより動作を把握し易かったり、或いは面白みのある動画を提供できることになる。

例えば図 5 A のボールの軌跡においてグラデーション的に色や濃度が変化することで、ボールの回転数を表現する画像とすることができる。

また図 5 B の場合、剣道の選手毎に竹刀の動きの軌跡を違う色で表示することで、各選手の竹刀の動きを把握しやすい動画が実現できる。

図示しないが、サッカーの試合でボールの軌跡をつねに最後にタッチした選手のチームに応じた色などとする事で、チーム毎のボールの支配率を表現するような動画も実現できる。

[0089] 以上、第1の実施の形態としては、動体80に対する画像エフェクトに関して処理例を説明したが、動体80に応じた画像エフェクトは更に多様に考えられる。

動体80の情報としては、動体80が人間の場合、さらに個人識別の情報、表情の情報、表情や音声等から推定される感情の情報などを用いて、エフェクト画像EFの態様を変化させることも考えられる。

また動体80についてのエフェクト画像EFは、軌跡画像に限らず、動体追尾を示す画像、動体の移動予測を示す画像、動体の種別や動作態様のみを示す画像なども想定される。

[0090] <3. 第2の実施の形態：音声認識に基づくテロップ>

第2の実施の形態として、動画に対する付加画像として音声認識に基づくテロップを表示する例を説明する。

例えば音声認識技術を用いることで、動画に対してリアルタイムもしくは非リアルタイムの編集でテロップを入れることができる。発話した人物とテロップの前後関係も厳密に再現することで、画一的でなく、発話者がわかりやすいテロップ表示、シーンに違和感を与えにくいテロップ表示、より表現力の高いテロップ表示等が実現できる。

[0091] 図10Aは、多数の人が登場するシーンで、それぞれの発話内容を、それぞれの発話者81の近辺にテロップ画像TPとして示している例である。

なお、これは説明上の極端な例であり、動画では実際に多数の人が一斉に同時に発話するシーンは多くなく、時間的にずれて発話されることが多いため、図示の状態のように多数のテロップが1フレームに付加されることは少ない。

各テロップ画像TPは、その発話者81の深度に応じて表示されることで、他の被写体との前後関係が反映された状態で画像に合成されることになる。

。

また、テロップ画像ＴＰとは、その表示態様、表示タイミングなどにかかわらず、動画内で発生した音を文字で表す画像のことを言う。

[0092] このようなテロップ画像ＴＰを付加する処理例を図１１に示す。

ステップＳ１７１で画像処理装置１は動画内の音声データから発話検出を行う。

ステップＳ１７２で画像処理装置１は音声認識及び発話者の特定を行う。

ステップＳ１７３で画像処理装置１は、発話内容をテキストデータ化する

。

ステップＳ１７４で画像処理装置１は、検出した発話内容をテロップ画像ＴＰとして表示するフレーム区間を設定する。

即ち画像処理装置１は、音声認識部４Ｂが動画における発話を検出し、その解析を行って、発話者８１や発話内容を特定し、また発話内容をテキストデータ化する。テロップ画像ＴＰを表示するフレーム区間は、例えば発話開始のタイミングのフレームから、発話が終了したタイミングのフレームよりも１～２秒程度後のフレームなどとして設定する。これらの情報を付加画像生成部５と画像編集処理部６が受け取ることになる。

なお発話者８１の特定には、画像解析部４Ａの画像解析により、口が動いている被写体を検出し、その口が動いている区間と発話音声の区間と比較して、略一致した場合に、その被写体を発話者８１とすることなどが考えられる。また、発話音声の周波数解析などにより、男性／女性の推定や年齢層の推定などを行い、画像解析結果と照合して発話者を推定することも考えられる。

動画に登場する人物の声紋情報などがあれば、音声解析により、動画内の発話者を正確に特定できる。

[0093] ステップＳ１７５で画像処理装置１は付加画像生成部５の機能によりテロップ画像ＴＰを生成する。

テロップ画像ＴＰの生成処理を図１２Ａに示す。

ステップS 2 5 0で画像処理装置1は音声解析により生成したテキスト情報を処理対象として取得する。

ステップS 2 5 1で画像処理装置1は、テキスト情報をテロップ画像化する。例えば所定の吹き出し、フォント、色等の設定に基づいてテロップ画像TPを生成する。

[0094] テロップ画像を生成したら、その検出した発話についてテロップ表示を行う区間の各フレームについて、主に画像編集処理部6の機能によりステップS 1 7 6～S 1 7 8及びS 1 0 8～S 1 1 3の処理を行う。

ステップS 1 7 6で画像処理装置1は、テロップ表示を行うフレーム区間における1フレームの画像データを処理対象として取得する。

ステップS 1 7 7で画像処理装置1は、取得したフレームが、それまでのフレームからシーンチェンジしているか否かを判定する。例えば前フレームとの各画素データの差分などからシーンチェンジ判定を行い、もしシーンチェンジと判定した場合は図11の処理を終える。シーンチェンジと判定されなければ画像処理装置1はステップS 1 7 8に進む。

[0095] ステップS 1 7 8で画像処理装置1は、当該フレームでのテロップ画像TPのテロップ画像サイズ及びそのテロップ画像TPの挿入領域を設定する。またテロップ画像TPにおける各画素の深度情報DPeを設定する。

例えばテロップ画像TPの各画素の深度情報DPeは、発話者81としての被写体の頭部が写った画素の深度情報DP、また頭部の深度情報DPから所定のオフセット量を与えた深度などとする可以考虑。

テロップ画像TP内の各画素の深度情報DPeは同じ値でよいが、特に徐々に深度の値を変化させるなどしてもよい。テロップの最初と最後で奥行き位置が変化するようなものとする場合である。

また、テロップ画像TPのサイズは、その深度情報DPeに応じたサイズとすることが考えられる。例えば深度情報DPeの値が小さい（浅い）場合はサイズを大きく、深度情報DPeの値が大きい（深い）場合はサイズを小さくすることで、遠近法に従って、近くの発話者81の場合は大きなテロッ

プ、遠くの発話者 81 の場合は小さなテロップとすることができる。

挿入領域については、例えば発話者 81 の頭部近辺などにおいて、設定したテロップ画像 TP のサイズに応じた領域を設定する。

[0096] 現フレームについて以上の設定を行ったら、ステップ S108～S113 で、設定したテロップ画像 TP を現フレームの画像に合成していく。これは図 6 のステップ S108～S113 と同様である。即ち画像処理装置 1 は、当該フレームの画像データ上で、設定した挿入領域に含まれる各画素 PX を、順次テロップ画像 TP の各画素について設定した深度情報 DP_e と比較していき、前後関係を反映して、当該フレームの元の画像データとテロップ画像 TP のいずれを表示画素とするかを決定していく。

[0097] 以上の処理を 1 フレームについて行い、ステップ S179 でテロップ画像 TP を表示するフレーム区間までの処理が終了したと判定されるまで、再び処理をステップ S176 から次のフレームについて実行する。

これにより、動画内の発話者 81 に対してテロップ画像 TP が表示される図 10A のような動画が実現される。

特にテロップ画像 TP は、深度情報 DP_e と周囲の被写体の深度情報 DP の比較により画素単位で表示有無が設定されることで、手前側の画像を隠してしまうことがなく、画像内の前後関係を反映した状態で表示される。これにより発話者 81 の状況（位置）を適切に表現したテロップ画像 TP となっている。

また発話者 81 に近い位置にテロップ画像 TP が表示されることで、画像（発話者 81 等）と文字を同時に見やすくなり、映画等の字幕として応用しても、より動画の内容をわかりやすくできるものとなる。

[0098] なお非リアルタイムでテロップ付加編集を行う場合は、動画内で検出される発話毎に図 11 の処理が行われるようにすればよい。

リアルタイムでテロップ画像 TP を付加していくには、動画を数秒程度遅延させて出力するようにすることが考えられる。例えば動画データについて数秒のバッファリングを行い、各フレームが数秒の遅延をもって読み出され

るようにする。そのバッファリングの間に、当該バッファリングしている数秒間のフレームの音声認識を行う。

そして音声認識に応じた 1 又は複数のテロップ画像 TP を用意したうえで、バッファリングから読み出した各フレームについて、ステップ S 1 7 6 以降の処理を行っていくようにすればよい。

[0099] また図 1 1 の処理例では、ステップ S 1 7 4 で設定されてフレーム区間としてテロップ画像 TP が表示される。これは、例えば発話タイミングから、発話が終了して 1 ～ 2 秒程度経過するまでの区間などとしたが、このように発話終了後も多少の時間はテロップ画像 TP を表示させることで、テキスト内容を読みやすくするものである。

但し、ステップ S 1 7 7 でシーンチェンジと判定された場合は、テロップ画像 TP の合成を行わないことになるため、そのフレームでは、前フレームまでのテロップは表示されない。これはシーンチェンジがあってもテロップが続くことで不自然な画像となることを防止するものとなる。

つまり、シーンチェンジとならない限り、発話から発話終了から若干後のタイミングまでの期間、テロップ画像 TP の表示が行われるが、シーンチェンジがあったら、その期間内でもテロップ画像 TP の表示は終了されることになる。これによってテロップの読みやすさと、不自然なテロップ表示の防止とを両立する。

[0100] ところで、テロップ画像 TP については、画一的な画像ではなく、より表現力の高い画像とすることが考えられる。

例えば図 1 0 B は発話者 8 1 が笑っている場合、図 1 0 C は発話者 8 1 が怒っている場合、図 1 0 D は発話者 8 1 が驚いている場合を示している。ここで、笑っている状況、怒っている状況、驚いている状況に応じて、テロップ画像 TP の表示態様を変えている。例えば吹き出しの形状や色、テキストのフォントなどを変えることで、発話者の感情等に応じた画像とする。

このようにすることでテロップ画像 TP は、より表現力が高いものとなる。

[0101] そこで例えば図 11 のステップ S 175 におけるテロップ画像 TP の生成処理について、図 12B のような処理を行うことが考えられる。

ステップ S 250 で画像処理装置 1 は音声解析により生成したテキスト情報を処理対象として取得する。

ステップ S 261 で画像処理装置 1 は、発話者の感情推定を行う。例えば発話音声の音量や周波数特性等の解析を行い、平静状態、興奮状態などを推定する。また画像解析により発話者 81 の表情や動作等から感情を推定する。

ステップ S 262 で画像処理装置 1 は、推定した発話者 81 の感情に応じて、テロップ画像 TP としての吹き出しの形状、テキストサイズ、フォント、色等を設定する。

ステップ S 263 で画像処理装置 1 は、ステップ S 262 での設定に基づいてテキスト情報をテロップ画像化する。

[0102] このようにすることで、推定した発話者 81 の感情に応じてテロップ画像 TP の画像態様を設定し、動画に付加することが可能となる。

[0103] また感情推定よりも簡易な処理としては、音量に応じてテロップ画像 TP の画像態様を変化させることが考えられる。例えば図 11 のステップ S 175 におけるテロップ画像 TP の生成処理について、図 12C のような処理を行う。

[0104] ステップ S 250 で画像処理装置 1 は音声解析により生成したテキスト情報を処理対象として取得する。

ステップ S 271 で画像処理装置 1 は、発話音声の音量情報を取得する。

ステップ S 272 で画像処理装置 1 は、取得した音量情報に応じて、テロップ画像 TP としての吹き出しの形状、テキストサイズ、フォント、色等を設定する。

ステップ S 273 で画像処理装置 1 は、ステップ S 272 での設定に基づいてテキスト情報をテロップ画像化する。

このようにすることで、発話音声の音量が表現されたテロップ画像 TP を

動画に付加することが可能となる。

[0105] なお以上の例では動画内の人物の発話をテロップ画像TPとすることを述べたが、人物の発話音声に限らず、動物の鳴き声、周囲の音（環境音）を文字を用いて表すテロップ画像を付加することも考えられる。その場合も、音源の画像内の位置や深度に応じてテロップ画像TPの表示を行うことが適切となる。

[0106] <4. 第3の実施の形態：情報提示画像の付加>

第3の実施の形態として、動画に対する付加画像として情報提示画像ISを表示する例を説明する。

例えば情報ソース15から取得した情報を用いて、動画に対してリアルタイムもしくは非リアルタイムの編集で情報提示画像ISを入れることができる。このときに、画像内の挿入領域に応じて、他の被写体との前後関係を再現したり、情報提示画像自体に奥行きを与えたりすることで、視聴者がみやすい情報提示、シーンに違和感を与えにくい情報提示等が実現できる。

[0107] 図13Aは、陸上競技（短距離走）のタイムの情報を取得し、そのタイムを情報提示画像ISとして表示している例である。

例えばゴールラインに沿ってタイム表示を行うことで、視聴者に対し、わかりやすく、かつ臨場感があり高揚するような動画を提供できる。この場合に、情報提示画像IS自体に深度情報DPeを与え、遠近法に沿って、遠くなる程小さくなる表示とすることで、画面上の違和感もないようにすることができる。

[0108] 図13Bは例えばスケート演技の軌跡をエフェクト画像EFとして表示するとともに、その演技内の情報（例えば各技の名称や得点等）を情報提示画像ISとして表示させた例である。このような表示をリアルタイムで行ったり、演技後に俯瞰状態で表示できるようにすることで、視聴者が演技内容を理解しやすいものとなる。また情報提示画像ISに深度を設定し、深度の関係性をもって他の被写体と合成することで、例えば画面上でスケート選手の手前には来ない状態で情報表示が行われるなどすることができる。つまり本

来の選手の画像の邪魔にならないような情報提示画像 I S を表示できる。

[0109] リアルタイム処理として情報提示画像 I S を表示させる場合の処理例を図 14 に示す。なお図 8 と同様の処理については同じステップ番号を付し、重複説明を避ける。

エフェクト開始後において、画像処理装置 1 は、フレームタイミング毎に、ステップ S 152 ~ S 160 及び S 107A ~ S 115 の処理を行う。

[0110] 画像処理装置 1 はステップ S 152 でフレーム画像データを取得し、ステップ S 153 で被写体の識別を行い、ステップ S 154 で動体の状態を把握する処理を、図 8 で説明したように行ったら、ステップ S 160 で外部情報を取得する。例えば図 13A の場合の短距離走のタイム（当該フレームのタイミングにおけるタイム）の情報を取得する。

[0111] そしてステップ S 107A で画像処理装置 1 は情報提示画像 I S を設定する。

情報提示画像 I S の設定（生成）については、概略図 7 と同様に行えば良い（図 7 の「エフェクト画像 EF」を「情報提示画像 I S」に読み替える）。

まずステップ S 201 で情報提示画像 I S としての画像を生成する。これは、例えば取得したタイムの情報を示す画像を生成すればよい。

[0112] ステップ S 202 で画像処理装置 1 は、情報提示画像 I S の画像サイズ、形状、色等の表示態様を設定する。ここで、画像サイズや形状は、動画の他の被写体に合わせることが考えられる。例えば図 13A のゴールラインに沿って表示する場合には、その遠近を考慮して形状やサイズを設定する。その意味では、フレーム画像データへの挿入領域の設定（S 204）も、この時点で同時に行うようにすることが考えられる。

ステップ S 203 で画像処理装置 1 は、情報提示画像 I S の各画素の深度情報 D P e を設定する。例えばこのときに、合成する被写体の遠近法に従って徐々に変化させるようにする。即ち情報提示画像 I S における各画素の深度情報 D P e より 1 つ手前となる値を、挿入領域の各画素の深度情報 D P と

する。

[0113] このように情報提示画像 I S を生成したら、画像処理装置 1 はステップ S 1 0 8 ~ S 1 1 5 の処理を図 8 で説明したように行う。

即ち画像処理装置 1 は、当該フレームの画像データ上で、設定した挿入領域に含まれる各画素 P X を、順次情報提示画像 I S の各画素について設定した深度情報 D P e と比較していき、前後関係を反映して、当該フレームの元の画像データと情報提示画像 I S のいずれを表示画素とするかを決定していく。

[0114] 以上の処理を 1 フレームについて行い、ステップ S 1 5 5 でエフェクト終了とされるまでは、再び図 1 4 の処理をステップ S 1 5 2 から実行する。

これにより、動画とともに例えばタイム等の情報を示す情報提示画像 I S が表示される。しかも情報提示画像 I S は、表示される位置に応じた遠近法に沿ったサイズや形状とされて、違和感がない。

また情報提示画像 I S を構成する画素は、元の画素の深度情報 D P に沿った深度情報 D P e とされる。例えば図 1 3 A のような動画である場合、その後、ランナーがタイム表示の部分を駆け抜けていくことになるが、その際もランナーの位置と情報提示画像 I S の各部の前後関係が反映された状態となる。これは、あたかもタイムを示す数値の物体が実際にゴールラインに沿って置かれているような画像表現となり、視覚効果の高い情報提示を実現できる。

[0115] なお、非リアルタイムでの合成処理としても情報提示画像 I S の付加が可能であることは言うまでもない。

また提示する情報は、多様に考えられる。例えばスポーツの動画であれば、競技のスコア、解説のためのテキスト、選手情報などが想定される。

撮像しているカメラに設けたセンサ、競技場に設けたセンサなどの検出データに基づく情報を提示することもできる。

[0116] < 5. 第 4 の実施の形態の：編集インタフェース >

以上の実施の形態では、動画に対する付加画像（エフェクト画像 E F、テ

ロップ画像ＴＰ、情報提示画像ＩＳ）を自動的に付加する例を述べたが、以下では、ユーザが付加画像の編集（例えば付加画像の表示の修正や付加画像自体の追加など）を手動で行うためのインタフェースについて説明する。

[0117] 図１５により第４の実施の形態としての編集インタフェースの例を示す。

図１５Ａ、図１５Ｂは、それぞれ例えばスマートフォン、タブレット、パーソナルコンピュータ等の情報処理装置の表示画面を示している。

この表示画面上には、編集操作のための画像モニタ領域６０、操作領域６１が設けられる。

[0118] 画像モニタ領域６０には編集対象の動画が表示される。動画は通常表示、一時停止表示、変速再生表示などが可能とされる。図では被写体８２，８３としての人物と、エフェクト画像ＥＦが表示されている例としている。

ここで、エフェクト画像ＥＦは半透明の壁状とされた仮想的な物体としている。そして被写体８２はエフェクト画像ＥＦより手前側に、被写体８３はエフェクト画像ＥＦより奥側に居る状態の画像となっている。

[0119] 操作領域６１には、台形状のデプスマップ画像６１ｃが表示される。このデプスマップ画像６１ｃは前端辺が画像の左右を表し、側辺は深度軸６１ｂに沿ったものとして奥行き方向を表している。即ちデプスマップ画像６１ｃとしての台形は遠近法で奥行きを表す画像とされている。

そして、そのデプスマップ画像６１ｃ内に、操作アイコン画像６１ａが表示される。

[0120] 操作アイコン画像６１ａは、画像モニタ領域６０に表示されている画像内でのエフェクト画像ＥＦの奥行き位置及び水平位置に応じた状態でデプスマップ画像６１ｃ内に表示される。

そしてユーザは、タッチパネル操作により操作アイコン画像６１ａのドラッグ等を行うことで、デプスマップ画像６１ｃ内での操作アイコン画像６１ａの位置を移動させることができる。

例えば図１５Ｂは、操作アイコン画像６１ａを手前側に移動させた状態である。

このような操作により、実際の動画内でのエフェクト画像E Fの位置（深度方向及び水平方向の位置）を変化させることができる。図15Bにおける画像モニタ領域60では、エフェクト画像E Fが、右手前側に移動され、被写体82もエフェクト画像E Fよりも奥側になった状態が示されている。

[0121] 例えば以上のようにすることで、エフェクト画像E F等の画面内の位置、深度をタッチパネルを用いた直観的操作により変更することができる。

なお、エフェクト画像E Fに限らず、テロップ画像T Pや情報提示画像I Sについても同様の操作で表示位置（深度方向及び水平方向の位置）を調整することができる。

[0122] なお、ここでは壁状のエフェクト画像E Fを示したが、画面全体の不透明の壁状のエフェクト画像E Fを用いれば、特定の被写体のみを表示させ、他の被写体を壁で隠すといった画像効果が可能である。例えば被写体82だけを表示させ、他をエフェクト画像E Fで隠す画像である。

その場合、当該操作により、壁を前後させて、隠す領域を調整することもできる。

[0123] <6. 第5の実施の形態の：編集インタフェース>

第5の実施の形態の編集インタフェース例を図16に示す。これは三面図によるプレビューを行うようにした例である。

[0124] 図16は情報処理装置において表示される編集画面であるが、この画面は主に5つの領域に分割されて表示が行われている。

即ちフロントビュー領域62、トップビュー領域63、サイドビュー領域64、設定領域65、タイムライン領域66である。

[0125] フロントビュー領域62は通常の動画モニタ画面とされ、動画内容が再生状態、一時停止状態、変速再生状態などで表示される。ここでは先の図15と同様に、被写体82、83及び半透明の壁状のエフェクト画像E Fが表示される例を示している。

[0126] トップビュー領域63は、上側から見た状態のデプスマップとして画像を示している。例えばトップビュー領域63の上下方向に深度軸63aが表示

され、画像の深度が表現される。そして画像内の被写体 8 2, 8 3 及びエフェクト画像 E F が、上方からみた状態として、それらに設定された深度の位置に示される。

[0127] サイドビュー領域 6 4 は、側方から見た状態のデプスマップとして画像を示している。例えばサイドビュー領域 6 4 の左右方向に深度軸 6 4 a が表示され、画像の深度が表現される。そして画像内の被写体 8 2, 8 3 及びエフェクト画像 E F が、側方からみた状態として、それらに設定された深度の位置に示される。

[0128] 設定領域 6 5 には、エフェクト画像 E F の画像及び設定情報が表示される。設定情報とはサイズ、角度、位置、色、深度などの設定についての情報である。設定領域 6 5 の各設定値はユーザ操作により変更可能である。

[0129] タイムライン領域 6 6 は、時間軸 6 6 a に沿って動画のタイムラインを表している。特にタイムライン上において、被写体 8 2, 8 3 やエフェクト画像 E F の出現ポイント、終了ポイントやプレビューされている現在の位置が表示される。

例えば「human1」「human2」は被写体 8 2, 8 3 を、「mol1」はエフェクト画像 E F を表し、○で登場タイミング、●で終了（画面からいなくなる）タイミングを示している。またバー 6 6 b により、フロントビュー領域 6 2、トップビュー領域 6 3、サイドビュー領域 6 4 に示される動画内のプレビューされている時点（フレーム）が示される。

[0130] このような編集インタフェースによれば、ユーザは、フロントビュー領域 6 2 のプレビュー画像をみながら、トップビュー領域 6 3 やサイドビュー領域 6 4 により前後関係を確認することができる。そして、これらをみながら、設定領域 6 5 での操作によりエフェクト画像 E F の設定を任意に変更することができる。

エフェクト画像 E F の深度の設定に関しては、トップビュー領域 6 3 やサイドビュー領域 6 4 におけるエフェクト画像 E F のタッチ操作、ドラッグ操作で変更可能としてもよい。

[0131] またタイムライン領域66における操作により、エフェクト画像EFの登場タイミング、終了タイミングも調整可能である。つまり動画内でのエフェクト画像EFの登場区間（エフェクト実行フレーム区間）を、タイムライン領域66での登場タイミングや終了タイミングのマーク（例えば「○」「●」）をドラッグ操作等により移動させることで変更可能とするとよい。

これらにより、ユーザは、実際の動画の中で、エフェクト画像EFのサイズ、色、深度や、エフェクト画像EFの登場区間を容易に調整できる。

[0132] <7. 第6の実施の形態の：編集インタフェース>

第6の実施の形態としての編集インタフェースの例を図17で説明する。これは深度軸と時間軸を表したタイムラインを表示する例である。

図17に示す編集インタフェース画面では、画像モニタ領域60、レイヤー表示領域67、二軸マップ領域68が設けられている。

[0133] 画像モニタ領域60として、動画のプレビュー表示が行われる。即ち動画内容が再生状態、一時停止状態、変速再生状態などで表示される。ここでは被写体としての5人の人物（被写体人物OB1, OB2, OB3, OB4, OB5）と、3つのエフェクト画像EF（EF1, EF2, EF3）が表示されている例としている。

[0134] レイヤー表示領域67では、被写体人物OB1, OB2, OB3, OB4, OB5の各レイヤーLOB1, LOB2, LOB3, LOB4, LOB5と、エフェクト画像EF1, EF2, EF3のレイヤーが斜視的に示されている。これによって被写体人物OB1～OB5とエフェクト画像EF1, EF2, EF3の前後関係が明確にわかるようにされる。

[0135] 二軸マップ領域68は、横方向が時間軸、縦方向が深度軸とされる領域である。この二軸マップ領域68では、被写体人物OB1, OB2, OB3, OB4, OB5の深度に対して、エフェクト画像EF1, EF2, EF3の深度が各時間において示されている。

[0136] このようなインタフェース画面において、例えば二軸マップ領域68上で、エフェクト画像EFの深度を時間軸で示すラインを任意に変更できるよう

にする。これにより、動画内でのエフェクト画像EFの深度変化を容易に編集できることになる。

また、各エフェクト画像EF1、EF2、EF3の深度変化と被写体人物OB1、OB2、OB3、OB4、OB5の深度位置関係が明確であり、画像効果もわかりやすいものとなる。

[0137] <8. 第7の実施の形態の：編集インターフェース>

第7の実施の形態としての編集インターフェースの例を図18から図32を用いて説明する。これも時間軸と深度軸を有する二軸マップを表示する例であるが、さらに二軸マップを活用しながら多様なプレビューを実現する例である。

[0138] 図18から図23は、編集操作画像70としての各種の表示態様を示している。

本例では、編集操作画像70として図18の二軸マップ領域71を表示させながら、多様な画像態様の切替が行われる。

[0139] まず図18は二軸マップ領域71としてタイムライン表示を行っている状態である。

二軸マップ領域71は、左右方向が時間軸とされてタイムラインとしての機能を持つ。さらに上下方向が深度軸とされることで、時間と深度の二軸のマップとなる。

[0140] なお編集操作画像70には、ツールアイコン72、画面切替アイコン73、現在時間表示74、動画操作アイコン75、拡大縮小操作子76、プレビュー操作子77等も表示される。

ツールアイコン72としてはタイムライン操作、エフェクト画像EFの配置などに用いる操作子が用意されている。

画面切替アイコン73としては、後述する俯瞰、正面、斜視などのプレビュー画面の切替のための操作子が用意されている。

現在時間表示74として、タイムラインでの現在時間（プレビューされている画像（フレーム）についての動画内での時刻（タイムコード）が表示さ

れる。

動画操作アイコン 75 として、タイムラインによる再生／停止／コマ送り／早送りなどの操作のための操作子が用意される。

拡大縮小操作子 76 により、画像の拡大や縮小の操作が可能とされる。

[0141] プレビュー操作子 77 は、二軸マップ領域 71 の時間軸方向のタイミングを指定して、そのタイミングでのプレビュー表示を指示する操作子である。

例えば再生タイミング（再生時間位置）がプレビュー操作子 77 により示され、プレビュー操作子 77 を操作すると図 19 のように、その時点でのプレビュー画像 78 の表示が行われる。

また、プレビュー操作子 77 の時間方向の位置を任意に指定して、その時点でのプレビューを要求する操作ができるようにしてもよい。

図 19 のようにプレビュー画像 78 が表示された状態では、閉じ操作子 79 が表示され、この操作によりプレビュー画像 78 の表示を終了させる操作ができるようにされる。

[0142] 以下では各画面内容について説明するが、ここでは、例えば図 15，図 16 で示したような被写体 82，83 とエフェクト画像 EF が画像内に存在するとして説明する。

図 18 に示すように、二軸マップ領域 71 には、エフェクト画像 EF や被写体 82，83 の深度の時間軸方向の遷移が、深度表示ライン HL 1，HL 2，HL 3 として示される。

例えば深度表示ライン HL 1 は、エフェクト画像 EF の深度の時間軸での変化を示すものとされる。深度表示ライン HL 2、HL 3 は、被写体 82，83 についての深度の時間軸での変化を示すものとされる。

このように深度表示ライン HL 1，HL 2，HL 3 が示されることで、ユーザは動画内での時間進行に伴う被写体 82，83 とエフェクト画像 EF の前後位置関係の変化が明確に把握できることとなる。

[0143] この二軸マップ領域 68 におけるあるタイミングにおけるプレビュー操作子 77 が操作されることで、図 19 のようにプレビュー表示が行われる。

この場合、プレビュー画像 78 の表示は画像内容である被写体 82, 83 とエフェクト画像 EF を上方からみた状態とされる。そして二軸マップ領域 68 の指定されたタイミングを左右に広げて割り込むように行われる。

つまり二軸マップ領域 68 における縦方向の深度軸はプレビュー画像 78 についても適用され、被写体 82, 83 とエフェクト画像 EF の深度関係が表現される。

時間軸については、破線部分（プレビュー画像 78 の部分）は時間の進行が停止された状態となる。換言すれば時間軸が寸断されてプレビュー画像 78 が挿入される。

時間軸が寸断されるため、プレビュー画像 78 の左右では、被写体 82, 83 とエフェクト画像 EF の深度変化はない。このためプレビュー画像 78 における被写体 82, 83 とエフェクト画像 EF は、それぞれ左右の深度表示ライン HL 1, HL 2, HL 3 をつなぐ破線上に位置することになる。

このようなプレビュー画像 78 により、ユーザは任意のタイミングでの被写体 82, 83 とエフェクト画像 EF の深度関係が明確に把握できる。

[0144] プレビュー画像 78 に対して操作を行うことで、エフェクト画像 EF の深度位置を変化させることができる。例えばユーザはエフェクト画像 EF を画面上下（即ち深度軸方向）に移動させるように操作を行うことで、エフェクト画像 EF の深度を調整する。これにより当該プレビュー画像 78 のフレームにおいてエフェクト画像 EF の深度が変更される。この場合に、画面上で左右に連続する深度表示ライン HL 1 も連動して変化するようにしてもよい。

また エフェクト画像 EF の設定（サイズ、位置、種別等）を変化させることができるようにしてもよい。

[0145] 図 20 は縮小表示した状態を示している。拡大縮小操作子 76 を操作することで、画面の拡大縮小を行うことができ、例えば縮小することで、ユーザはより長い時間範囲のタイムラインを見ることができる。

特に図 19 のようにプレビュー画像 78 を表示されると、タイムラインと

して深度表示ラインHL 1, HL 2, HL 3が表示される時間範囲が短くなる。その場合、図20のように縮小表示させることで、長い時間範囲でプレビュー画像78の前後の時点の深度関係の変化を確認できるようになる。

なお、ユーザはタイムバー71aを移動させることにより、タイムライン上の各タイミングを指定できる。

[0146] 図21は複数のタイミングにおけるプレビュー画像78、78Aを表示させた状態である。ユーザは例えばタイムバー71aにより時間軸上での任意のタイミングを指定してプレビュー操作を行うことで、追加のプレビュー画像78Aを表示させることができる。つまり複数のタイミングでのプレビュー画像78, 78Aを同時に確認できる。

これにより時間軸上での被写体82, 83とエフェクト画像EFの前後関係を確認したり、エフェクト画像EFの深度を複数の時間位置で調整していくことが容易となる。

[0147] 図22は正面画表示を行っている状態である。図19のようなプレビュー画像78は、二軸マップ領域68の深度軸に合わせて表示されるため、実際の動画内容ではない。ユーザは画面切替アイコン73の操作により、プレビュー画像78の態様を切り替えることができる。図22の正面プレビュー画像78Fは、動画の元々のフレーム画像の表示となる。この場合、深度軸は画面奥行き方向となるため、タイムライン表示は一時的に消去される。但し、図19のように左右にタイムライン表示を残しながら、プレビュー画像78を正面プレビュー画像78Fに切り替えるようにしてもよい。

[0148] 図23は立体プレビュー画像78Rとして立体画表示を行っている状態である。動画のフレーム画像に対して奥行き方向の表現を加えた画像となる。ユーザは画面切替アイコン73の操作により、この立体プレビュー画像78Rにも切り替えることができる。

この場合、時間軸方向と深度方向が斜視的に表現可能となるため、立体プレビュー画像78Rとともに、タイムライン表示（深度表示ラインHL 1, HL 2, HL 3の表示）も行われている。

これにより画像の前後関係のイメージや、被写体 8 2, 8 3 とエフェクト画像 E F の深度の時間方向の遷移がわかりやすいものとなる。

[0149] 以上のように画面態様を切替ながら編集が可能とされる。以下では、図 2 4 の動画イメージを用いて、図 2 5 から図 3 1 により編集インタフェースの利用例を説明する。

[0150] 図 2 4 A, 図 2 4 B, 図 2 4 C, 図 2 4 D, 図 2 4 E, 図 2 4 F は、動画を構成する各フレームの画像例を示している。これは被写体 8 2 としての人物の周囲に、エフェクト画像 E F としての光源 E F 1 が回っている動画であるとする。エフェクト画像 E F としては、光源 E F 1 に少し遅れて回っているテキスト E F 3 と、光源 E F 1 の軌跡 E F 2 が表示されているとする。

[0151] このような動画においてエフェクト画像の編集を行う。編集のための画面は上述のように各種表示状態の間で遷移可能である。図 2 5 は表示状態の遷移を示している。

図 2 5 A はタイムライン表示（図 1 8 参照）である。

タイムライン表示からは図 2 5 B のプレビュー表示（図 1 9 参照）に遷移できる。

図 2 5 B のプレビュー表示からは、図 2 5 A のタイムライン表示、図 2 5 C の縮小表示（図 2 0 参照）、図 2 5 D の複数プレビュー表示（図 2 1 参照）、図 2 5 E の正面画表示（図 2 2 参照）、図 2 5 F の立体画表示（図 2 3 参照）に遷移できる。これらのそれぞれからプレビュー表示に戻ることもできる。

縮小表示と、複数プレビュー表示は相互に遷移可能である。

立体画表示と、複数プレビュー表示は相互に遷移可能である。

立体画表示と、正面画表示は相互に遷移可能である。

もちろん以上は一例であり、遷移経路は多様に考えられる。

[0152] 図 2 4 の動画の場合、タイムライン表示は図 2 6 のようになる。ここでは深度表示ライン H L 1, H L 2, H L 3 が表示される。深度表示ライン H L 1 は光源 E F 1 の深度の遷移を表している。光源 E F 1 は被写体 8 2 の周囲

を回る移動を行っているため、時間軸方向には正弦波的な遷移となる。深度表示ラインHL 2は被写体8 2の深度の遷移を表している。動画内で被写体8 2は立ち止まって移動していないことから深度表示ラインHL 2は或る深度を保った直線となっている。

深度表示ラインHL 3はテキストEF 3の深度の遷移を表している。テキストEF 3は光源EF 1の後について被写体8 2の周囲を回る移動を行っているため、これも時間軸方向には、深度表示ラインHL 1より少し遅れた正弦波的な遷移となる。

[0153] 或るタイミングでのプレビュー画像7 8が表示されると図2 7のようになる。

プレビュー画像7 8では被写体8 2の周囲に光源EF 1、テキストEF 3、光源の軌跡EF 2が示される。この画面上で、光源EF 1やテキストEF 3の深度を調整することができる。また光源EF 1、軌跡EF 2、テキストEF 3の設定（サイズ、色、形状その他）を変更可能としても良い。

[0154] 縮小表示が行われると図2 8のようになる。これによりプレビュー画像7 8の前後の時間における深度表示ラインHL 1、HL 2、HL 3により、光源EF 1、被写体8 2、テキストEF 3について長い時間範囲での深度遷移が明確に表示される。

[0155] さらに他のある時点のプレビュー画像7 8 Aが表示された状態が図2 9である。このように複数時点のプレビュー画像7 8、7 8 Aを開くことで、タイムラインでの関係を確認しながら、光源EF 1やテキストEF 3の深度や、光源EF 1、軌跡EF 2、テキストEF 3の設定を変更することができる。

[0156] 動画における画像を確認するには図3 0のように正面プレビュー画像7 8 Fを表示させる。これによりユーザは実際の動画のフレーム画像を確認できる。特にユーザは、例えばプレビュー画像7 8で深度等の調整操作を行いながら正面プレビュー画像7 8を確認することで、より望むような調整ができているか否かを確認しやすい。

さらに図 31 の立体プレビュー画像 78R を表示することで、さらに動画内容と奥行きを同時に確認できる。これによりユーザは各種調整操作を行いながら、立体空間でのエフェクト画像 EF の動きを確認でき、編集作業が容易となる。

[0157] 以上のような編集インタフェースでは、二軸マップ領域 71 を基本として編集操作画像 70 としての各種画面を遷移しながら、ユーザに容易な編集環境を提供するものとなる。

この各種編集操作画像 70 は、図 1 の編集操作画像生成部 8 の機能により画像処理装置 1 が生成し、表示部 11 等に表示させる。

画像処理装置 1 は、ユーザの操作に応じた画面遷移を行うために、図 32 の処理を行うことになる。

[0158] 編集操作画像 70 の表示がユーザに指示されると、画像処理装置 1 はステップ S301 から S302 に進み、時間軸と深度軸による二軸マップ領域 71 を備えた編集操作画像、即ち図 18 のタイムライン表示を行う。

その状態で画像処理装置 1 は、ユーザによる操作として、終了操作 (S303)、プレビュー操作 (S304)、拡大縮小操作 (S305)、プレビュー閉じ操作 (S306)、正面画操作 (S307)、立体画操作 (S308) を監視する。

[0159] 終了操作が検知されたら、画像処理装置 1 はステップ S317 で編集操作画像 70 の表示を終了する。

[0160] 例えばプレビュー操作子 77 等によるプレビュー操作を検知したら、画像処理装置 1 はステップ S304 から S310 に進み、プレビュー可能な状態か否かを確認し、可能であればステップ S311 で、指定された時点のフレーム画像によるプレビュー画像 78 の表示を行う。

即ち画像処理装置 1 は、該当時点のフレーム画像における各画素の深度を確認し、各被写体やエフェクト画像 EF の深度をトップビューとして表現した画像を生成し、それを二軸マップ領域 71 の当該タイミングの部分に挿入して表示画像データを生成する。そしてその表示画像データを表示部 11 に

供給して表示させる。

なお、プレビュー不能な場合、例えば有効な画像が存在しない時点に対するプレビュー操作などの場合は、画像処理装置 1 は、検知したプレビュー操作は無効操作としてステップ S 3 0 3 ~ S 3 0 8 の監視ループに戻る。

[0161] 例えば拡大縮小操作子 7 6 による拡大縮小操作を検知したら、画像処理装置 1 はステップ S 3 0 5 から S 3 1 2 に進み、指示された拡大又は縮小が可能か否かを確認し、可能であればステップ S 3 1 3 で拡大表示又は縮小表示を行う。

即ち画像処理装置 1 は、二軸マップ領域 7 1 の縮尺を変更して、その拡大又は縮小した二軸マップ領域 7 1 内でプレビュー画像 7 8 が配置されるような表示画像データを生成する。そしてその表示画像データを表示部 1 1 に供給して表示させる。

なお、拡大又は縮小が不能な場合、例えば最大拡大率又は最小縮小率の状態にあるときなどは、画像処理装置 1 は、検知した拡大縮小操作は無効操作としてステップ S 3 0 3 ~ S 3 0 8 の監視ループに戻る。

[0162] 例えば閉じ操作子 7 9 によるプレビュー閉じ操作を検知したら、画像処理装置 1 はステップ S 3 0 6 から S 3 1 4 に進み、指定されたプレビュー画像 7 8 の表示を終了させ、その部分は二軸マップ領域 7 1 の表示に戻すようにする。

[0163] 例えば画面切替アイコン 7 3 を用いた正面画表示の操作を検知したら、画像処理装置 1 はステップ S 3 0 7 から S 3 1 5 に進み、指定されたプレビュー画像 7 8 を正面プレビュー画像 7 8 F に変更する。つまりこの場合画像処理装置 1 は、指定されたプレビュー画像 7 8 の元のフレーム画像を表示させる。またこの場合は、二軸マップ領域 7 1 の表示を停止させる。

[0164] 例えば画面切替アイコン 7 3 を用いた立体画表示の操作を検知したら、画像処理装置 1 はステップ S 3 0 8 から S 3 1 6 に進み、指定されたプレビュー画像 7 8 を立体プレビュー画像 7 8 R に変更する。つまりこの場合画像処理装置 1 は、指定されたプレビュー画像 7 8 の元のフレーム画像についての立

体画を生成し、斜視的な二軸マップ領域 7 1 に挿入した状態の表示画像データを生成して、表示部 1 1 で表示させる。

[0165] 画像処理装置 1 は以上の処理により各種の画面状態を遷移させながら、ユーザに二軸マップ領域 7 1 を基本とした編集インタフェースを提供することになる。

[0166] < 9. まとめ及び変形例 >

以上の第 1 ～第 7 の実施の形態によれば次のような効果が得られる。

各実施の形態では、画像処理装置 1 は、動画データに付加する付加画像を生成する付加画像生成部 5 と、付加画像を付加する対象とした動画データにおける画素の深度情報を用いて、動画データに付加画像を付加する画像編集処理を行う画像編集処理部 6 を備える。

これにより被写体の深度に応じた各種の画像編集処理、例えば被写体へのエフェクト画像 E F 等の付加画像の重ね合わせ、付加画像としてのグラフィックスや文字等の合成、光学的効果の付与等の処理を行うことができる。特にこれらの動画編集において、奥行きのある被写体画像の空間に自然になじむような画像表現が可能になる。

特に自動的に深度に応じた付加画像の合成が行われることで、例えば動画放送のための効果的な動画編集が容易に実行でき、例えばリアルタイム中継で視覚効果の高い放送コンテンツを提供できる。もちろん録画放送において容易に画像エフェクトを施した動画を提供できることにもなる。

そして本技術による多彩な画像表現、処理の容易性は、より広いユーザ層に受け入れられ易いと考えられる。映画やドラマは勿論、パーソナルコンテンツやソーシャルメディアのさらなる発展に貢献できる。特にスポーツなどは新たな楽しみ方が増えることでよりファン層の拡大が期待できる。

また自動的に深度に応じた合成が行われることで、一般ユーザにとっても多様かつ高度な画像処理を容易かつ気軽に実行できるものとなる。特にユーザに高度なスキルを要求せずに、より自然な編集動画を生成できる。これによりユーザにとっては画像表現の拡大がもたらされる。

通常、画像表現はユーザのスキルやツールの機能や性能に制約を大きく受けていたが、本実施の形態の技術により表現の幅が大幅に広がることが期待される。ユーザはより容易に自らのイメージに近い表現をすることができるし、一般的なエフェクトと異なり組み合わせで使うことによる効果が期待できるため表現のバリエーションも大きく広がる。

また本実施の形態は、ソーシャルメディアの発展にも寄与できる。

実施の形態の技術は、表現の多彩さ、容易さにより広いユーザ層に受け入れられることが想定される。これによりソーシャルメディアでの表現が多彩となり、おもしろみのある動画によるコミュニケーション環境を提供できる。

なお付加画像としてエフェクト画像EF、情報提示画像IS、テロップ画像TPを例に挙げたが、付加画像はこれらに限定されるものではない。これらに含まれない態様であっても、動画に画像効果や情報提示のために付加される画像が本開示で言う付加画像に相当する。

例えば動画内の被写体をコピーした画像を追加したり、他の動画の被写体を追加するなどの態様の付加画像も想定される。

またエフェクト画像EF、情報提示画像IS、テロップ画像TPなどを複合して動画に合成することも当然考えられる。

[0167] 第1、第2、第3の実施の形態では、動画データの被写体の深度情報と付加画像に設定された深度情報とを用いて被写体と付加画像の前後関係を反映した画像編集を行う例を述べた。

即ち付加画像について設定された深度情報と被写体（動画データの各フレームの画素）の深度情報により、付加画像と被写体の前後関係が規定されるようにする。

この前後関係を反映して、付加画像（エフェクト画像EF、情報提示画像IS、テロップ画像TP等）を動画の元々の被写体に合成することで、立体空間上での正しい位置関係で付加画像を挿入できる。つまり動画編集として付加画像を付加する際に、被写体の空間に対して違和感のない付加画像の合

成が実現できる。

[0168] 第1, 第2, 第3の実施の形態では、付加画像は、動画データにおける当該付加画像の挿入位置における深度情報に応じてサイズ調整されて動画データに合成されることを述べた。

例えば付加画像生成部5もしくは画像編集処理部6は、付加画像を深度情報に基づいてサイズ調整する。これにより付加画像自体が挿入位置の遠近に応じたサイズ（遠近法に沿った状態）で表現されるようにする。

付加画像自体が、その遠近によってサイズ調整されることで、元々の被写体と違和感の無い状態で合成される。これにより高品位な付加画像を合成することができる。例えば図5Aのエフェクト画像EF、図10Aのテロップ画像TPや図13Aの情報提示画像ISは、遠近を反映した画像となっていることで、自然に動画にマッチしている。

[0169] 第1の実施の形態では、付加画像は、動画データにおける動体検出結果を用いて生成されるエフェクト画像EFである例を述べた。

即ち動画内で動きのある被写体の動きの追尾や動きの軌跡などを表現するエフェクト画像EFを付加画像として生成し、動画に合成する（図6、図7、図8、図9参照）。

これにより動画における動体の動きに応じた多様な画像表現を実現できる。例えば図4C、図4D、図4E、図5A、図5Bに例示したように迫力のある動画、動きのわかりやすい動画などを提供できる。

なおエフェクト画像EFとしては動体に応じたものに限られない。例えば図15A、図15Bに例示したように、動体とは特に無関係のエフェクト画像EFも考えられる。動画内の被写体の一部を隠す効果、被写体の一部を強調する効果、特に画像内容とは無関係のグラフィックの挿入、複数の動画の重ね合わせなど、多様な例のエフェクト画像EFが想定される。

[0170] 第1の実施の形態では、エフェクト画像EFには動体に応じた深度情報が設定され、動画データの被写体の深度情報と、エフェクト画像に設定された深度情報とを用いて被写体と付加画像の前後関係を反映した画像編集を行う

例を述べた。

例えばエフェクト画像E Fについて、動体の動きに応じた深度情報を設定し、エフェクト画像が動体の深度に対応した画像となるようにする（図6、図7、図8、図9参照）。

これにより動体に対応するエフェクト画像E Fは、周囲の被写体との前後関係が規定される。この前後関係を反映して、エフェクト画像E Fを動画の元々の被写体に合成することで、立体空間内での正しい位置関係で動体に応じた効果画像を表現するような画像が実現される。従って動画内の被写体の空間に対して違和感のないエフェクト画像E Fとして動体軌跡や動体の動きに応じた画像が表現できる。

[0171] 第1の実施の形態では、エフェクト画像E Fは、動体に応じた深度情報が設定され、深度情報に応じてサイズ調整されて動画データに合成される例を述べた。

つまり付加画像生成部5もしくは画像編集処理部6は、エフェクト画像E Fを深度情報に基づいてサイズ調整する。これによりエフェクト画像E F自体が挿入位置（つまり各フレームにおける動体の位置）の遠近に応じたサイズ（遠近法に沿った状態）で表現されるようにする（図6、図7、図8、図9参照）。

エフェクト画像E Fが、その遠近によってサイズ調整されることで、元々の動画の中で動体が視認される大きさに応じたサイズ感で軌跡等が表現される。例えば図5Aのエフェクト画像E Fとして表現される動体80の軌跡は、各時点（フレーム）の動体位置に応じてサイズ（この場合は軌跡の横幅）が変化する。このように遠近を反映した画像とすることで、自然に動画にマッチしつつ、時間経過に伴う動体80の変位をわかりやすく表現できることになる。

[0172] 第1の実施の形態では、付加画像は、動画データにおける動体の軌跡を表すエフェクト画像E Fである例を述べた。

即ち動画内で動きのある被写体の動き軌跡を表現するエフェクト画像を付

加画像として生成し、動画に合成する（図 9 A、図 9 B 参照）。

これにより動画における動体の動きの軌跡をわかりやすく提示する画像表現を実現できる（図 4 C、図 4 D、図 4 E、図 5 A、図 5 B 参照）。

[0173] 第 1 の実施の形態では、付加画像は、動画データにおける動体の動作態様又は動作種別を表すエフェクト画像 E F である例を述べた。

即ち動画内で動体の動作として各種の態様や種別があるが、それらを表現するエフェクト画像を付加画像として生成し、動画に合成する（図 9 B 参照）。

これにより動画における動体の動きの態様や種別をわかりやすく提示する画像表現を実現できる。例えば図 4 C、図 4 E は略水平方向の移動を表し、図 4 D はスピンを表している。図 5 A の例では、動体 80 としてのボールの回転数によって色や濃度が変化する画像とした。図 5 B の場合、一方の人物の竹刀の軌跡と他方の人物の竹刀の軌跡を異なる色で表現する画像とした。

これらにより、迫力のある画像であるだけでなく、動きの態様や種別がわかりやすい画像を提供できることになる。

[0174] 第 2 の実施の形態では、付加画像は、動画データにおける音声検出結果を用いて生成されるテロップ画像 T P である例を述べた。

つまり動画内で人物の発話などを検出し、そのテロップ画像を付加画像として生成し、動画に合成する（図 11、図 12 A、図 12 B、図 12 C 参照）。

これにより動画において音声テロップを含む画像表現を実現できる。例えば図 10 A、図 10 B、図 10 C、図 10 D に例示したように音声を表現する動画を提供できる。そしてこのテロップ画像 T P が、深度情報を用いて合成されることで、音源（発話者）の深度を反映した画像表現となる。

[0175] 第 2 の実施の形態では、テロップ画像 T P は、動画データからの音声認識に基づいてテキストデータを表示する画像であるとした。

即ち動画内で発話される音声を認識しテキストデータを取得する。そしてそのテキストデータをテロップ表示する（図 11、図 12 A、図 12 B、図

1 2 C 参照)。

これにより動画における被写体人物の発話を明示するとともに、そのテロップは例えば発話した被写体人物の深度に対応した状態で表示されるように合成される。従って発話した被写体人物に応じたテロップとして合成され、発話者がわかりやすい画像表現が実現できる。

[0176] 第2の実施の形態では、テロップ画像TPは、動画データから認識される音声の音量に応じて異なる態様の画像とされる例を述べた。

テロップ画像TPは、例えば認識した音声の音量、声量に応じてサイズ、フォントなどが異なるなど、異なる表示態様とする(図1 2 C参照)。

これにより動画における被写体人物の発話などの音量によって、表示態様が異なるテロップ画像TPを、その発話した被写体人物の深度に対応した状態で表示されるようにすることができ、面白みのある、或いは表現力の高いテロップ表示が実現できる。

[0177] 第2の実施の形態では、テロップ画像TPは、動画データにおける発話者の感情情報に応じて異なる態様の画像とされる例を述べた。

例えば認識した音声や画像から発話者の感情を推定する。そして感情(怒り、喜び、驚き)などにより、サイズ、フォントなどが異なるなど、異なる表示態様とする(図1 2 B参照)。

例えば図1 0 B、図1 0 C、図1 0 Dに例示したように吹き出しの態様や文字サイズ、フォント等を変えることで感情を表現するテロップ画像TPとなる。

これにより動画における被写体人物の感情等を反映し、面白みのある、或いは表現力の高いテロップ表示が実現できる。

[0178] 第3の実施の形態では、付加画像は、取得情報を用いて生成される情報提示画像であるとした。

即ち動画に関して情報ソースから情報を取得し、その情報を提示する画像を付加画像として生成する。そして深度情報を用いて動画に合成する(図1 4 参照)。

これにより動画に関連した情報を、適切な前後関係や遠近状態であって違和感ない状態で画面上に提示することができる。例えば図 13 A のようにタイム表示を行う場合でも、ゴールライン近辺に合成する場合に、その遠近が表現されることで、違和感のない動画を視聴者に提供できることになる。

[0179] 第 5, 第 6, 第 7 の実施の形態では、画像処理装置 1 において、動画データに付加した付加画像を編集するためのユーザインタフェース画像として、動画の時間軸と深度軸が表現される編集操作画像を生成する編集操作画像生成部 8 を備えるものとした。

即ち付加画像の合成状態を編集したい場合に、ユーザが編集に用いる編集画像を生成し、表示されるようにする。この場合に編集画像は、動画の時間軸と深度軸が表現されるものとする。

図 16 の例、図 17 の例、図 18 ~ 図 31 で説明した例では、タイムラインとしての動画の時間軸が存在し、また被写体や付加情報の深度を示す軸が存在する。これによりユーザは、動画の進行とともに深度の状態を把握し易い。これによって付加画像の編集作業が容易化される。

なお、第 4 の実施の形態（図 15）では、編集インタフェース画面上で深度軸を備えるようにして、付加画像の深度を直感的に調整できるようにしている。これによってもユーザの簡易な編集が実現できる。

[0180] 第 6, 第 7 の実施の形態では、編集画面として一方の軸が時間軸で他方の軸が深度軸とされる表示領域を有するものとした。

例えば 1 つの領域として、例えば横方向が時間軸、縦方向が深度軸となるような領域を設けて編集用の画面が形成されるようにする。

図 17 の例や、図 18 ~ 図 31 で説明した例では、二軸マップ領域 68, 71 がこのような領域である。時間軸と深度軸から成る領域を設けることで、動画の進行と被写体や付加画像の深度の関係や変動が極めて認識し易くなり、編集作業の効率化、容易化が実現できる。

[0181] 第 7 の実施の形態の編集操作画像 70 は、時間軸方向の情報と、ある時点の画像（プレビュー画像 78）が同時に表示される画像とした。

例えば時間軸方向の情報がタイムラインとして表示されつつ、その一部の時点のプレビュー画像 7 8 や立体プレビュー画像 7 8 R が同時に表示されるようにする。

例えば図 1 9 や図 2 3 のように表示が行われることで、タイムライン上での或る時点の画像を容易に確認でき、しかも被写体等の前後関係の変化がわかりやすいものとなる。これによって編集作業の効率化、容易化が促進される。

[0182] 第 7 の実施の形態では、編集操作画像 7 0 は、時間軸方向の情報を示す画像上で、ある時点を指定することで、当該時点の画像が時間軸方向の情報と同時に表示される画像であるとした。

例えば時間軸方向の情報がタイムラインとして表示されつつ、或る時点を指定することで、タイムラインの情報が提示されたまま、当該時点のプレビュー画像 7 8 が図 1 9 のように表示されるようにした。

これにより、タイムライン上でユーザが指定した或る時点の画像を容易に確認できる。しかも、その前後での被写体等の前後関係の変化がわかりやすい。これによって編集作業の効率化、容易化が促進される。

[0183] 第 7 の実施の形態では、編集操作画像 7 0 は、時間軸方向のある時点のプレビュー画像を立体画像として表示する画像である例を述べた。

例えば時間軸方向の情報がタイムラインとして表示されつつ、タイムライン上で指定した或る時点の画像が立体プレビュー画像 7 8 R として図 2 2 のように表示される。

これにより被写体や付加画像の深度の関係がわかりやすくなり、これによっても編集作業の効率化、容易化が促進される。

[0184] 実施の形態のプログラムは、情報処理装置に、動画データに付加する付加画像を生成するステップと、付加画像を付加する対象とした動画データにおける画素の深度情報を用いて、動画データに付加画像を付加する画像編集処理を行うステップとを実行させるプログラムである。

即ち図 6、図 7、図 8、図 9、図 1 1、図 1 2、図 1 4 の処理を情報処理

装置に実行させるプログラムである。

[0185] このようなプログラムにより本実施の形態の画像処理装置 1 の実現が容易となる。

そしてこのようなプログラムはコンピュータ装置等の機器に内蔵されている記録媒体や、CPUを有するマイクロコンピュータ内のROM等に予め記憶しておくことができる。あるいはまた、半導体メモリ、メモリカード、光ディスク、光磁気ディスク、磁気ディスクなどのリムーバブル記録媒体に、一時的あるいは永続的に格納（記憶）しておくことができる。またこのようなリムーバブル記録媒体は、いわゆるパッケージソフトウェアとして提供することができる。

また、このようなプログラムは、リムーバブル記録媒体からパーソナルコンピュータ等にインストールする他、ダウンロードサイトから、LAN、インターネットなどのネットワークを介してダウンロードすることもできる。

[0186] なお、本明細書に記載された効果はあくまでも例示であって限定されるものではなく、また他の効果があってもよい。

[0187] なお本技術は以下のような構成も採ることができる。

(1)

動画データに付加する付加画像を生成する付加画像生成部と、
付加画像を付加する対象とした動画データにおける画素の深度情報を用いて、動画データに付加画像を付加する画像編集処理を行う画像編集処理部と、
を備えた
画像処理装置。

(2)

前記画像編集処理部は、動画データの被写体の深度情報と前記付加画像に設定された深度情報とを用いて被写体と付加画像の前後関係を反映した画像編集を行う

上記(1)に記載の画像処理装置。

(3)

前記付加画像は、動画データにおける当該付加画像の挿入位置における深度情報に応じてサイズ調整されて動画データに合成される

上記（１）又は（２）に記載の画像処理装置。

（４）

前記付加画像は、動画データにおける動体検出結果を用いて生成されるエフェクト画像である

上記（１）乃至（３）のいずれかに記載の画像処理装置。

（５）

前記エフェクト画像には動体に応じた深度情報が設定され、

前記画像編集処理部は、動画データの被写体の深度情報と前記エフェクト画像に設定された深度情報とを用いて被写体と付加画像の前後関係を反映した画像編集を行う

上記（４）に記載の画像処理装置。

（６）

前記エフェクト画像は、動体に応じた深度情報が設定され、深度情報に応じてサイズ調整されて動画データに合成される

上記（４）又は（５）に記載の画像処理装置。

（７）

前記付加画像は、動画データにおける動体の軌跡を表すエフェクト画像である

上記（１）乃至（６）のいずれかに記載の画像処理装置。

（８）

前記付加画像は、動画データにおける動体の動作態様又は動作種別を表すエフェクト画像である

上記（１）乃至（７）のいずれかに記載の画像処理装置。

（９）

前記付加画像は、動画データにおける音声検出結果を用いて生成されるテロップ画像である

上記（１）乃至（３）のいずれかに記載の画像処理装置。

（１０）

前記テロップ画像は、動画データからの音声認識に基づいてテキストデータを表示する画像である

上記（９）に記載の画像処理装置。

（１１）

前記テロップ画像は、動画データから認識される音声の音量に応じて異なる態様の画像とされる

上記（９）又は（１０）に記載の画像処理装置。

（１２）

前記テロップ画像は、動画データにおける発話者の感情情報に応じて異なる態様の画像とされる

上記（９）乃至（１１）のいずれかに記載の画像処理装置。

（１３）

前記付加画像は、取得情報を用いて生成される情報提示画像である

上記（１）乃至（３）のいずれかに記載の画像処理装置。

（１４）

動画データに付加した前記付加画像を編集するためのユーザインタフェース画像として、動画の時間軸と深度軸が表現される編集操作画像を生成する編集操作画像生成部を備えた

上記（１）乃至（１３）のいずれかに記載の画像処理装置。

（１５）

前記編集操作画像は、一方の軸が時間軸で他方の軸が深度軸とされる表示領域を有する

上記（１４）に記載の画像処理装置。

（１６）

前記編集操作画像は、時間軸方向の情報と、ある時点の画像が同時に表示される画像である

上記（１４）又は（１５）に記載の画像処理装置。

（１７）

前記編集操作画像は、時間軸方向の情報を示す画像上で、ある時点を指定することで、当該時点の画像が時間軸方向の情報と同時に表示される画像である

上記（１４）乃至（１６）のいずれかに記載の画像処理装置。

（１８）

前記編集操作画像は、時間軸方向のある時点の画像を立体画像として表示する画像である

上記（１４）乃至（１７）のいずれかに記載の画像処理装置。

（１９）

動画データに付加する付加画像を生成する手順と、

付加画像を付加する対象とした動画データにおける画素の深度情報を用いて、動画データに付加画像を付加する画像編集処理を行う手順と、
を情報処理装置が実行する画像処理方法。

（２０）

動画データに付加する付加画像を生成するステップと、

付加画像を付加する対象とした動画データにおける画素の深度情報を用いて、動画データに付加画像を付加する画像編集処理を行うステップと、
を情報処理装置に実行させるプログラム。

符号の説明

[0188] １…画像処理装置、２…画像取得部、３…深度取得部、４Ａ…画像解析部、４Ｂ…音声認識部、４Ｃ…情報取得部、５…付加画像生成部、６…画像編集処理部、７…画像出力部、８…編集操作画像生成部、１０…操作部、１１…表示部、１２…通信部、１３…記憶部、１４…画像ソース、１５…情報ソース、６０…画像モニタ領域、６１…操作領域、６１ａ…操作アイコン画像、６１ｂ…深度軸、６１ｃ…デプスマップ画像、６２…フロントビュー領域、６３…トップビュー領域、６３ａ…深度軸、６４…サイドビュー領域、６

4 a …深度軸、6 5 …設定領域、6 6 …タイムライン領域、6 7 …レイヤー表示領域、6 8 …二軸マップ領域、7 0 …編集操作画像、7 1 …二軸マップ領域、7 2 …ツールアイコン、7 3 …画面切替アイコン、7 4 …現在時間表示、7 5 …動画操作アイコン、7 6 …拡大縮小操作子、7 7 …プレビュー操作子、7 8, 7 8 A …プレビュー画像、7 8 R …立体プレビュー画像、7 8 F …正面プレビュー画像、7 9 …閉じ操作子、8 0 …動体、8 1 …発話者、8 2, 8 3 …被写体、1 0 0 …情報処理装置、E F …エフェクト画像、T P …テロップ画像、I S …情報提示画像、H L 1, H L 2, H L 3 …深度表示ライン

請求の範囲

- [請求項1] 動画データに付加する付加画像を生成する付加画像生成部と、
付加画像を付加する対象とした動画データにおける画素の深度情報を用いて、動画データに付加画像を付加する画像編集処理を行う画像編集処理部と、を備えた
画像処理装置。
- [請求項2] 前記画像編集処理部は、動画データの被写体の深度情報と前記付加画像に設定された深度情報とを用いて被写体と付加画像の前後関係を反映した画像編集を行う
請求項1に記載の画像処理装置。
- [請求項3] 前記付加画像は、動画データにおける当該付加画像の挿入位置における深度情報に応じてサイズ調整されて動画データに合成される
請求項1に記載の画像処理装置。
- [請求項4] 前記付加画像は、動画データにおける動体検出結果を用いて生成されるエフェクト画像である
請求項1に記載の画像処理装置。
- [請求項5] 前記エフェクト画像には動体に応じた深度情報が設定され、
前記画像編集処理部は、動画データの被写体の深度情報と前記エフェクト画像に設定された深度情報とを用いて被写体と付加画像の前後関係を反映した画像編集を行う
請求項4に記載の画像処理装置。
- [請求項6] 前記エフェクト画像は、動体に応じた深度情報が設定され、深度情報に応じてサイズ調整されて動画データに合成される
請求項4に記載の画像処理装置。
- [請求項7] 前記付加画像は、動画データにおける動体の軌跡を表すエフェクト画像である
請求項1に記載の画像処理装置。
- [請求項8] 前記付加画像は、動画データにおける動体の動作態様又は動作種別

を表すエフェクト画像である

請求項 1 に記載の画像処理装置。

[請求項9] 前記付加画像は、動画データにおける音声検出結果を用いて生成されるテロップ画像である

請求項 1 に記載の画像処理装置。

[請求項10] 前記テロップ画像は、動画データからの音声認識に基づいてテキストデータを表示する画像である

請求項 9 に記載の画像処理装置。

[請求項11] 前記テロップ画像は、動画データから認識される音声の音量に応じて異なる態様の画像とされる

請求項 9 に記載の画像処理装置。

[請求項12] 前記テロップ画像は、動画データにおける発話者の感情情報に応じて異なる態様の画像とされる

請求項 9 に記載の画像処理装置。

[請求項13] 前記付加画像は、取得情報を用いて生成される情報提示画像である
請求項 1 に記載の画像処理装置。

[請求項14] 動画データに付加した前記付加画像を編集するためのユーザインタフェース画像として、動画の時間軸と深度軸が表現される編集操作画像を生成する編集操作画像生成部を備えた

請求項 1 に記載の画像処理装置。

[請求項15] 前記編集操作画像は、一方の軸が時間軸で他方の軸が深度軸とされる表示領域を有する

請求項 1 4 に記載の画像処理装置。

[請求項16] 前記編集操作画像は、時間軸方向の情報と、ある時点の画像が同時に表示される画像である

請求項 1 4 に記載の画像処理装置。

[請求項17] 前記編集操作画像は、時間軸方向の情報を示す画像上で、ある時点指定することで、当該時点の画像が時間軸方向の情報と同時に表示

される画像である

請求項 14 に記載の画像処理装置。

[請求項18] 前記編集操作画像は、時間軸方向のある時点の画像を立体画像として表示する画像である

請求項 14 に記載の画像処理装置。

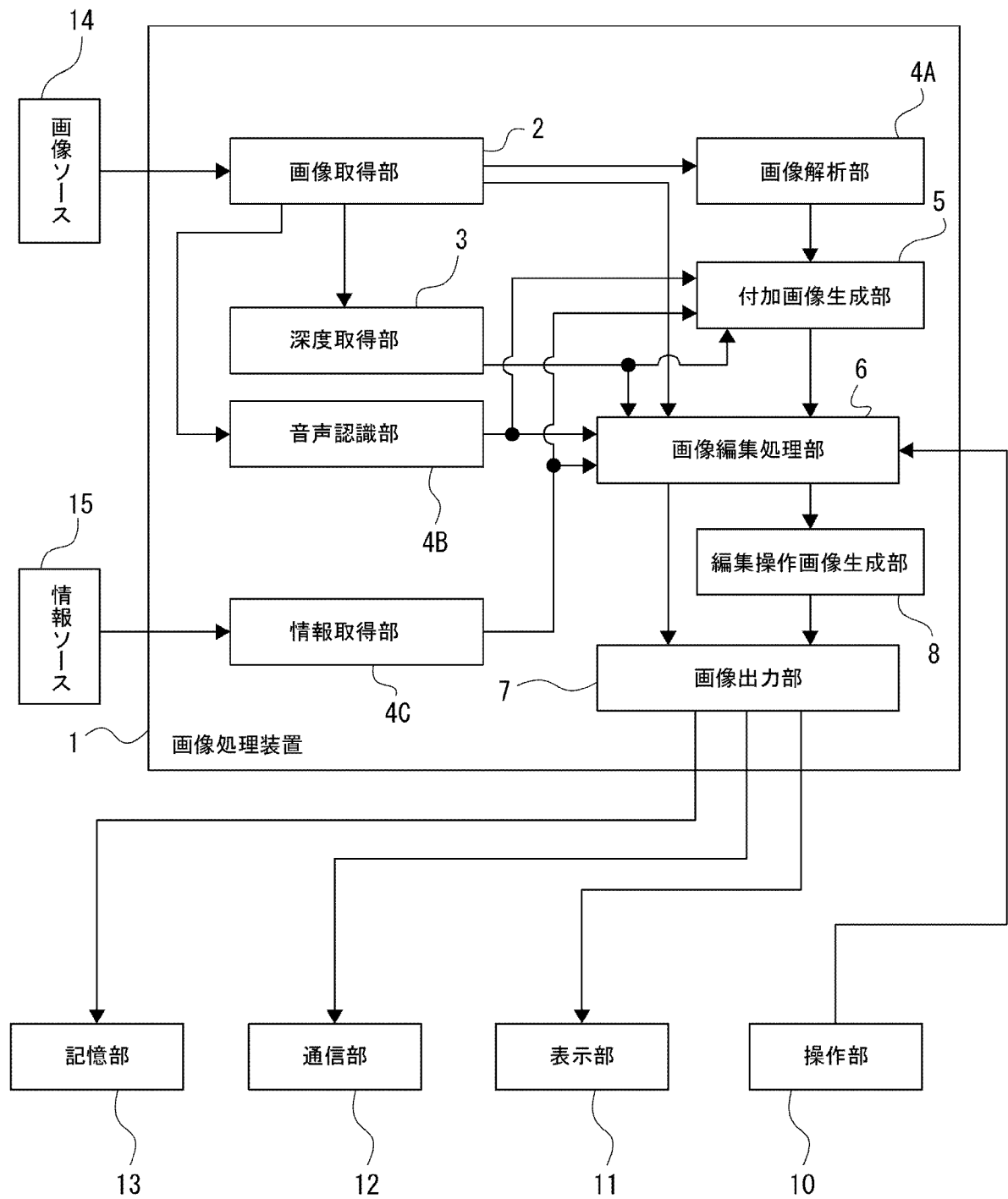
[請求項19] 動画データに付加する付加画像を生成する手順と、
付加画像を付加する対象とした動画データにおける画素の深度情報を用いて、動画データに付加画像を付加する画像編集処理を行う手順と、

を情報処理装置が実行する画像処理方法。

[請求項20] 動画データに付加する付加画像を生成するステップと、
付加画像を付加する対象とした動画データにおける画素の深度情報を用いて、動画データに付加画像を付加する画像編集処理を行うステップと、

を情報処理装置に実行させるプログラム。

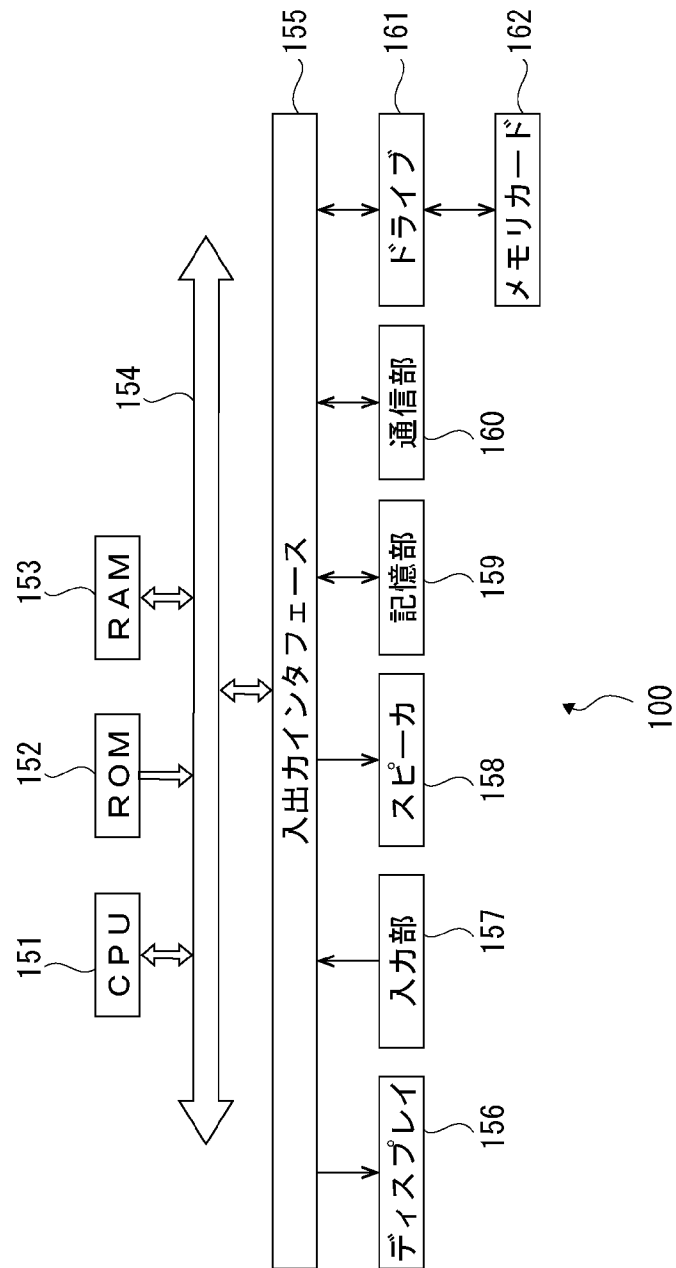
[図1]



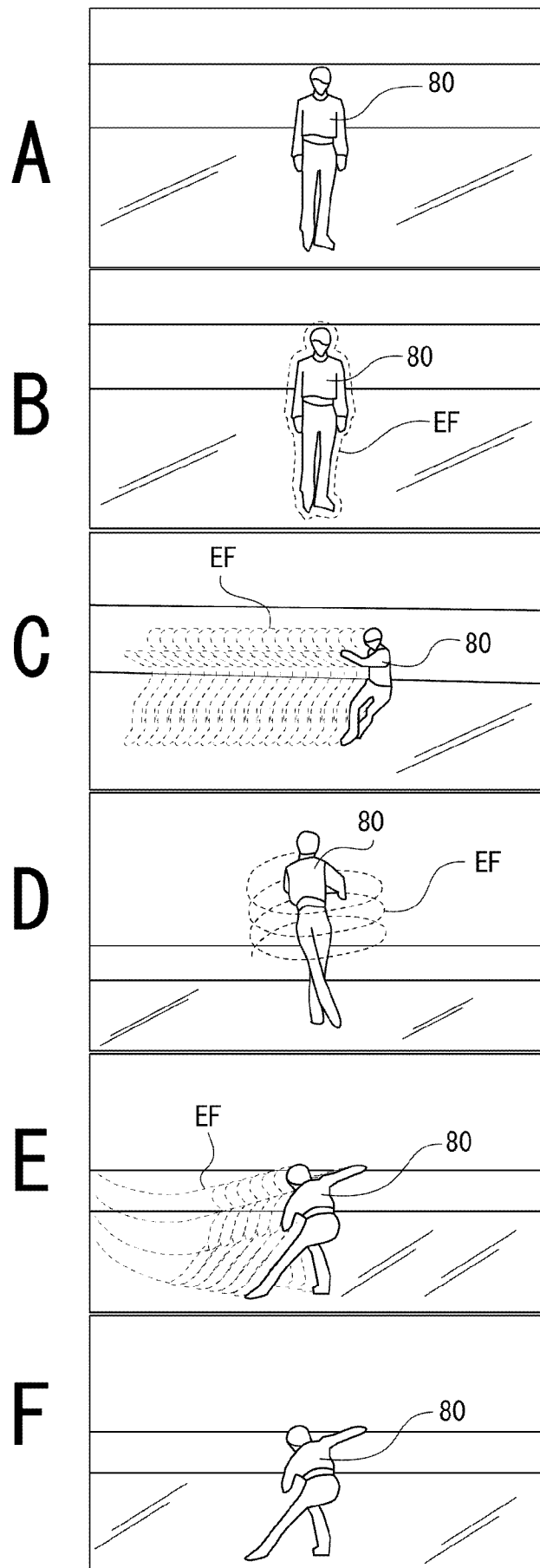
[図2]

	0	1	2	-----	m
0	<PX0-0> DP0-0	<PX0-1> DP0-1	<PX0-2> DP0-2		<PX0-m> DP0-m
1	<PX1-0> DP1-0	<PX1-1> DP1-1	<PX1-2> DP1-2		
⋮					
n	<PXn-0> DPn-0				<PXn-m> DPn-m

[図3]

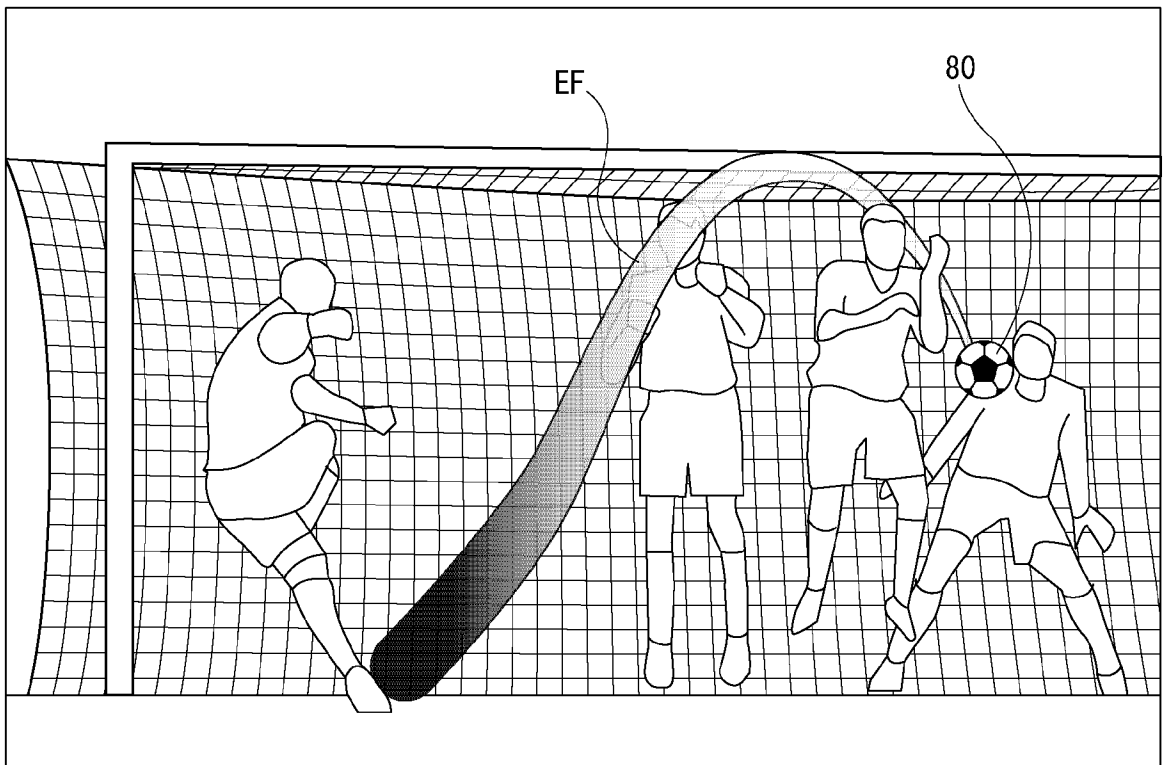


[図4]

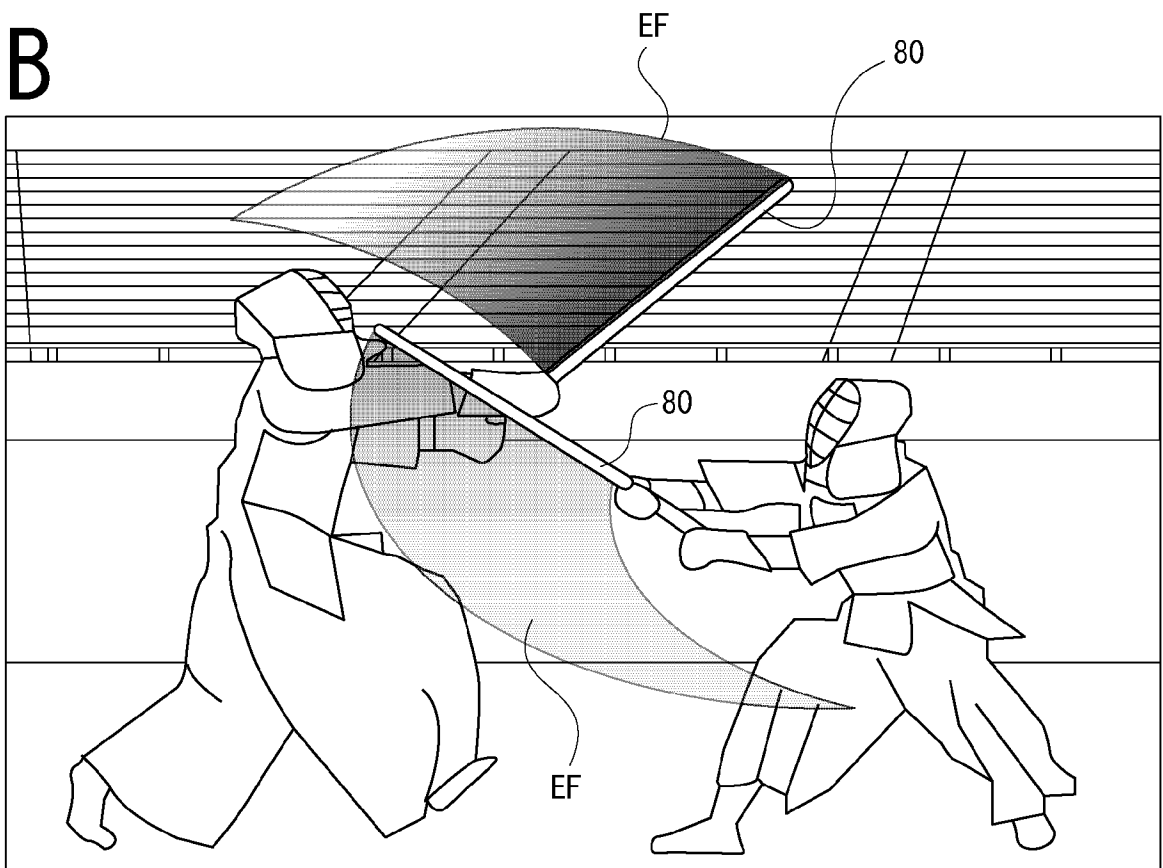


[図5]

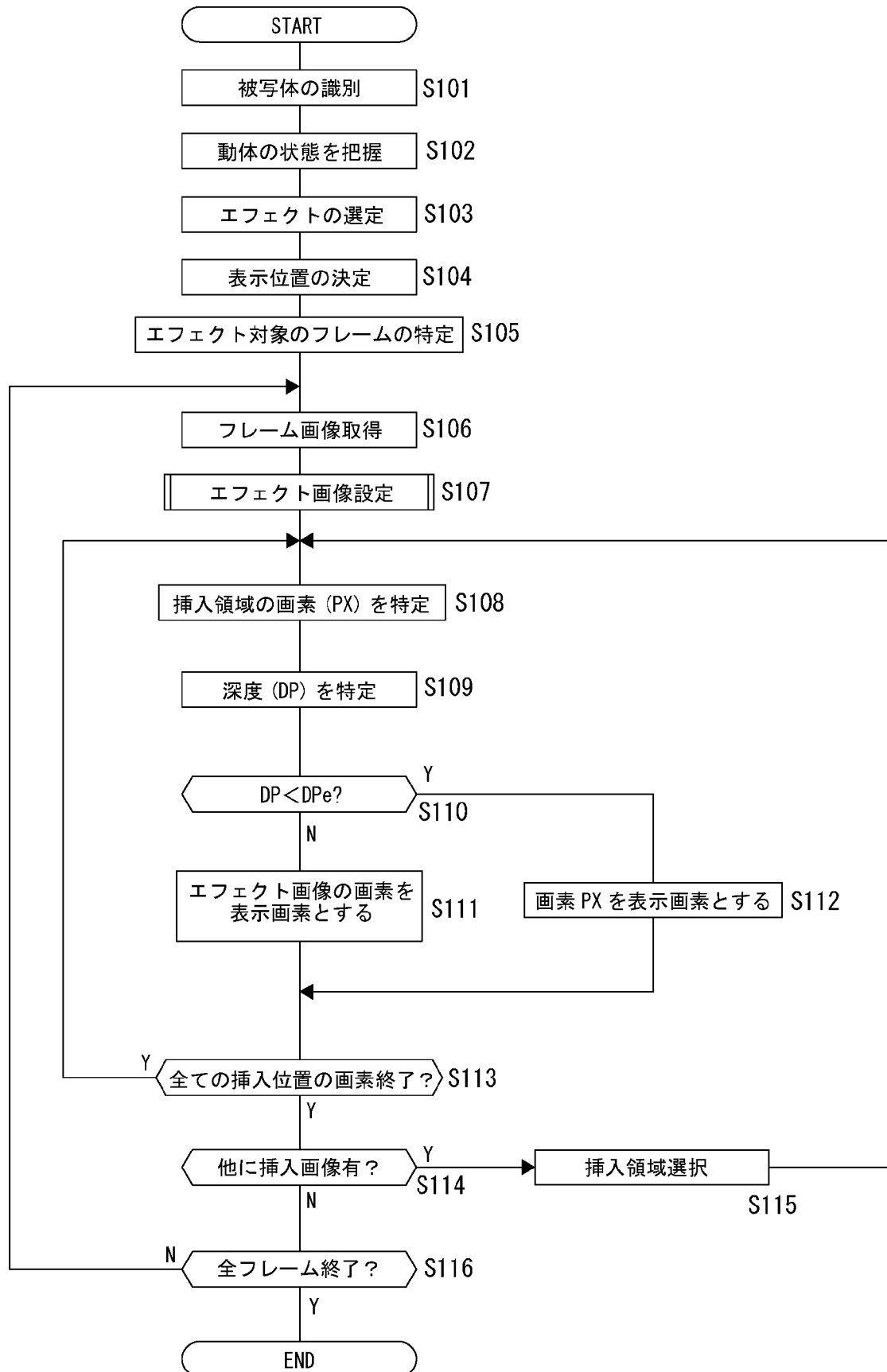
A



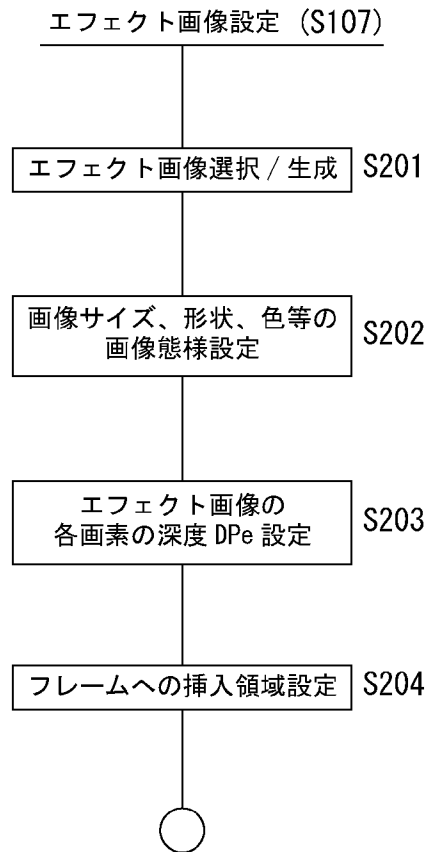
B



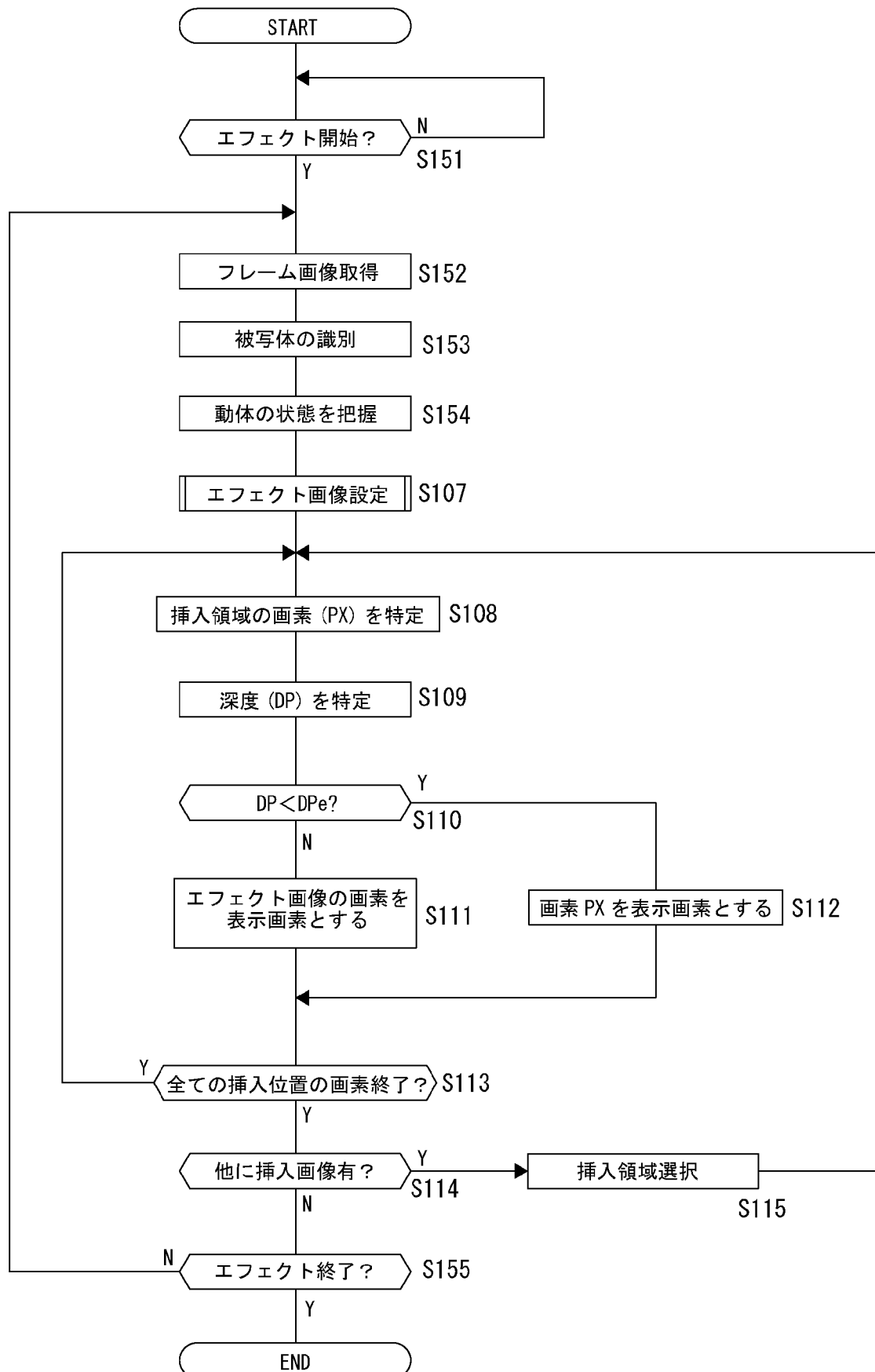
[図6]



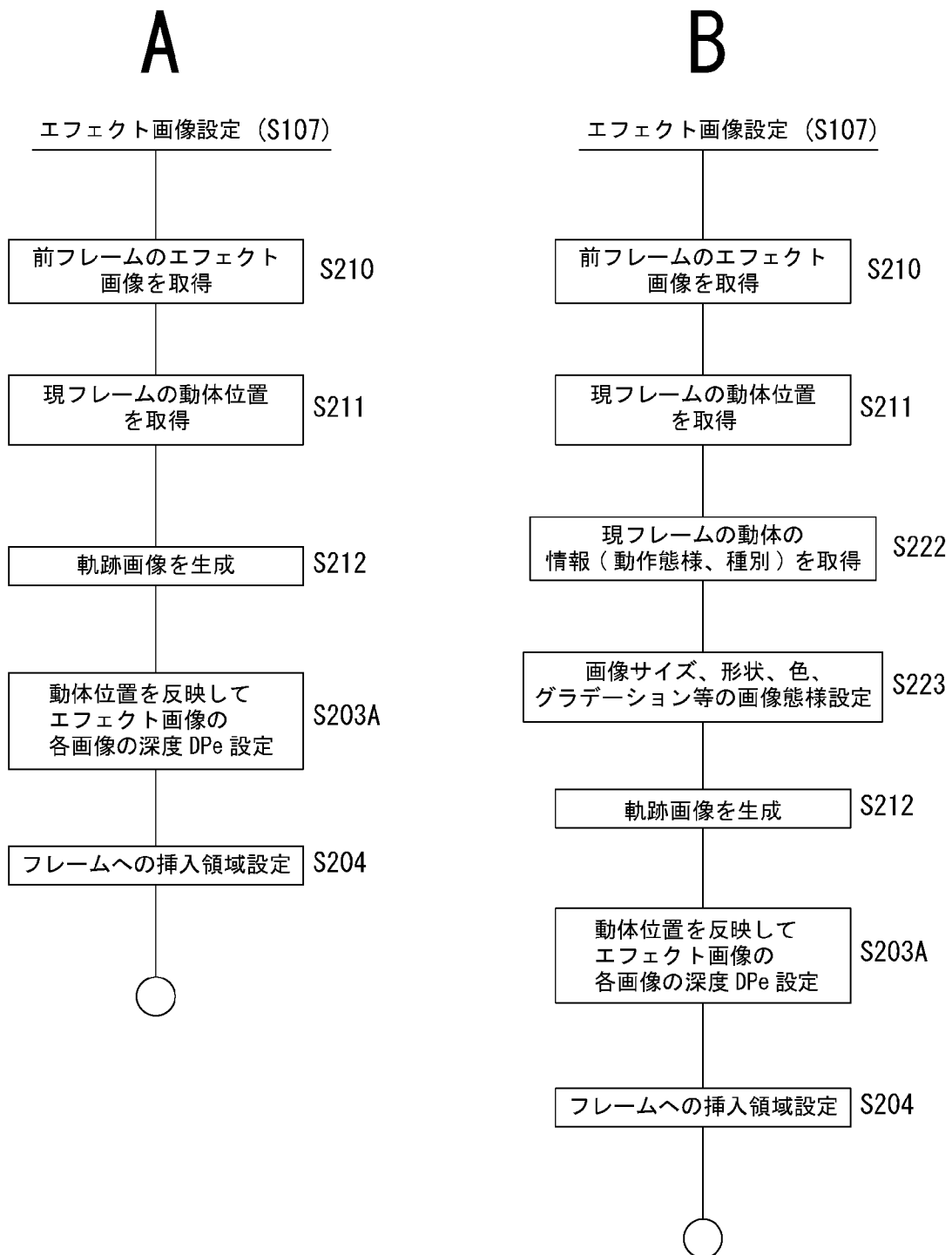
[図7]



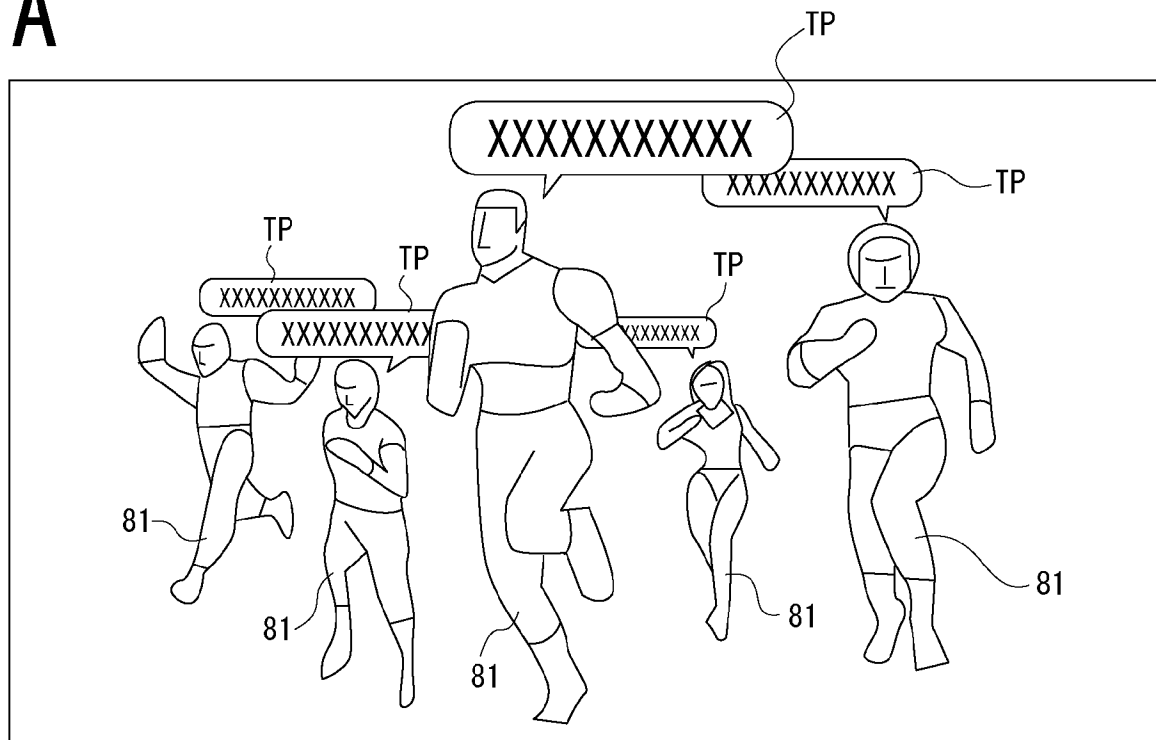
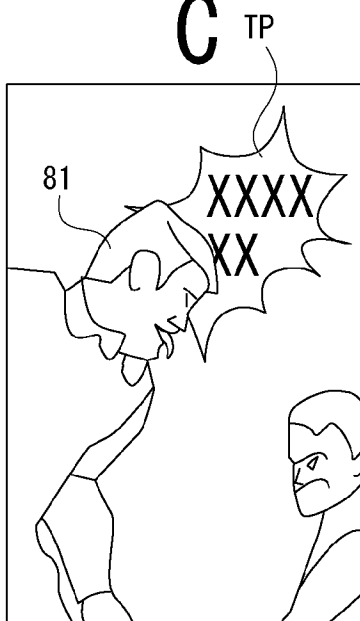
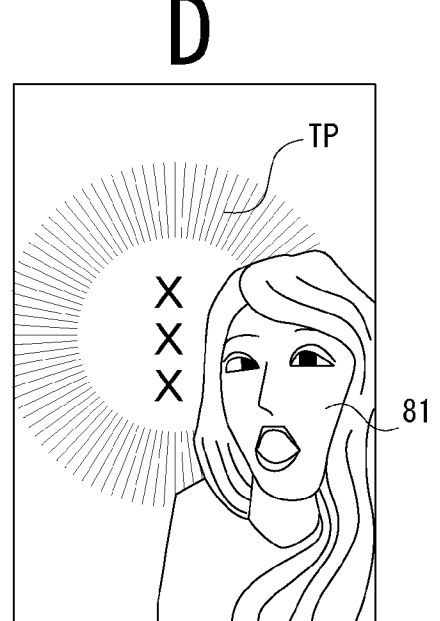
[図8]



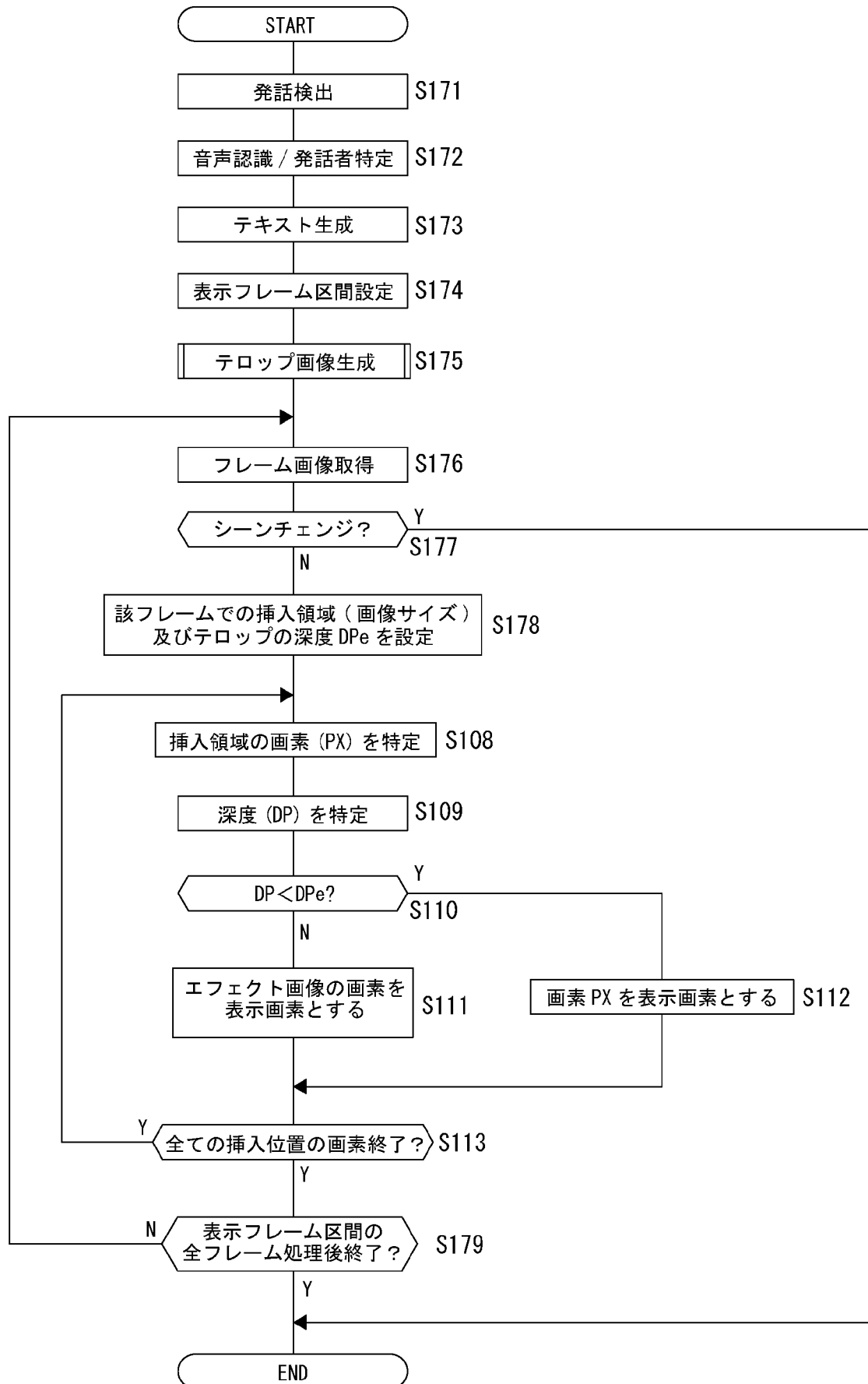
[図9]



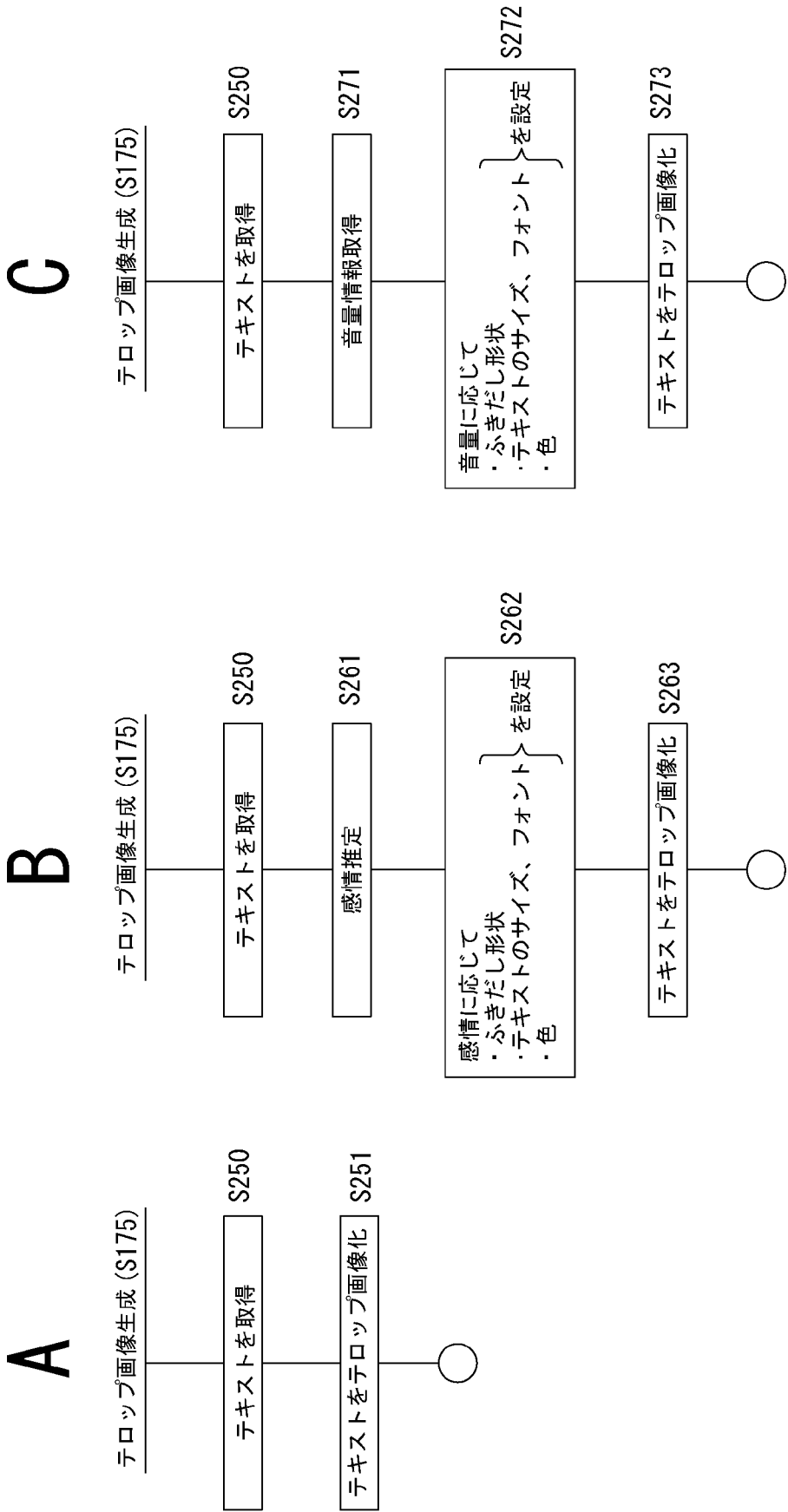
[図10]

A**B****C****D**

[図11]

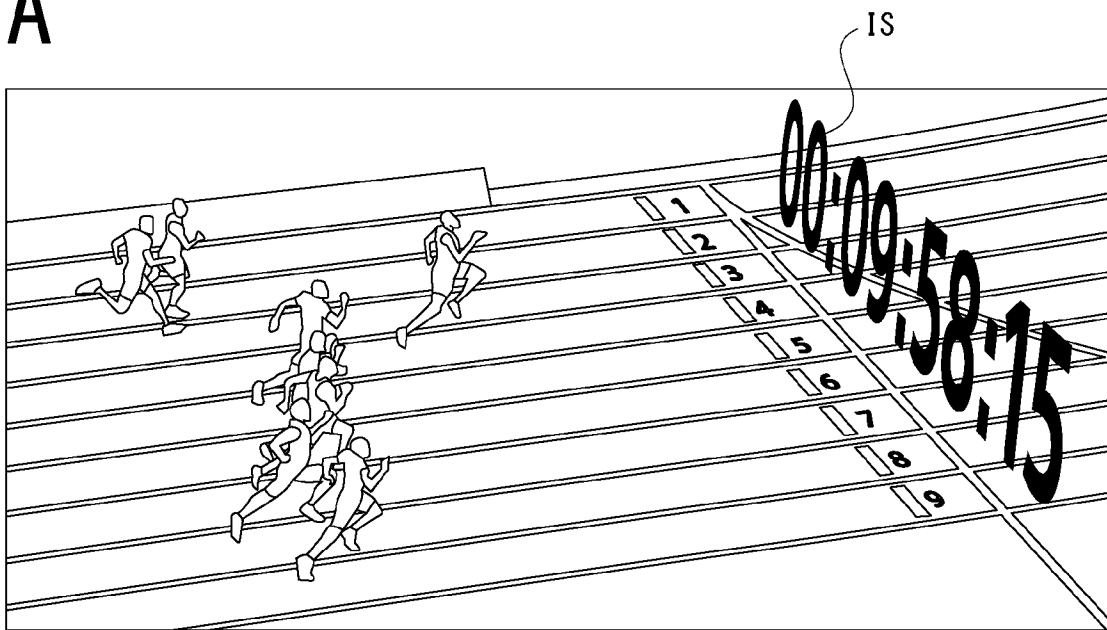


[図12]

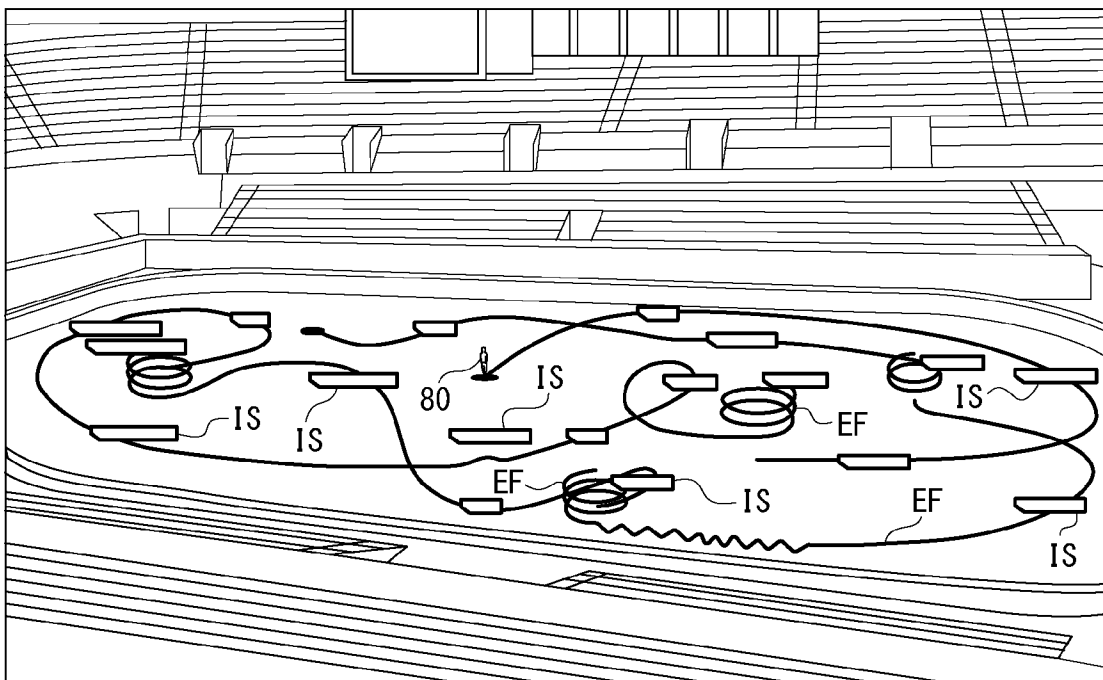


[図13]

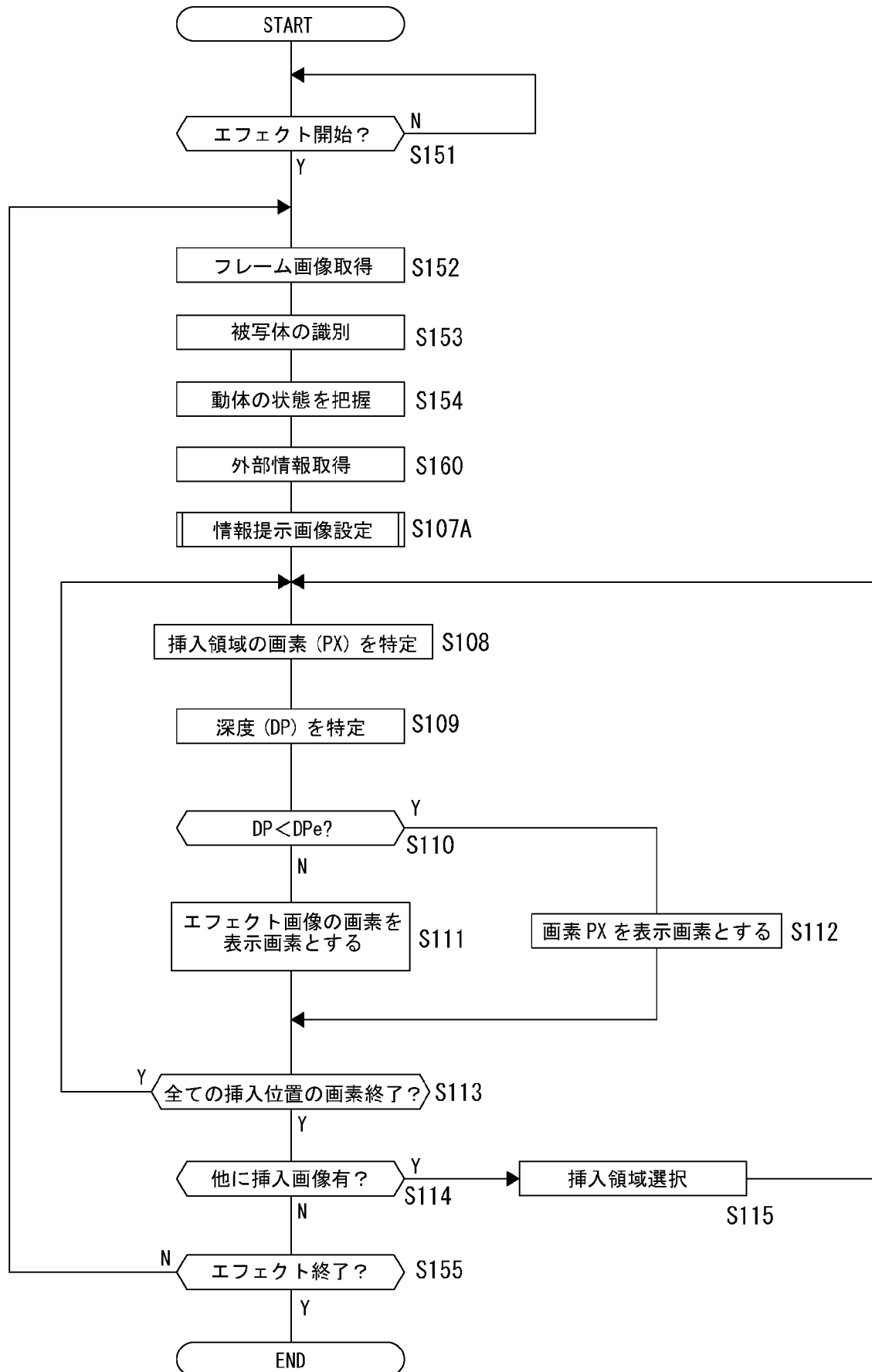
A



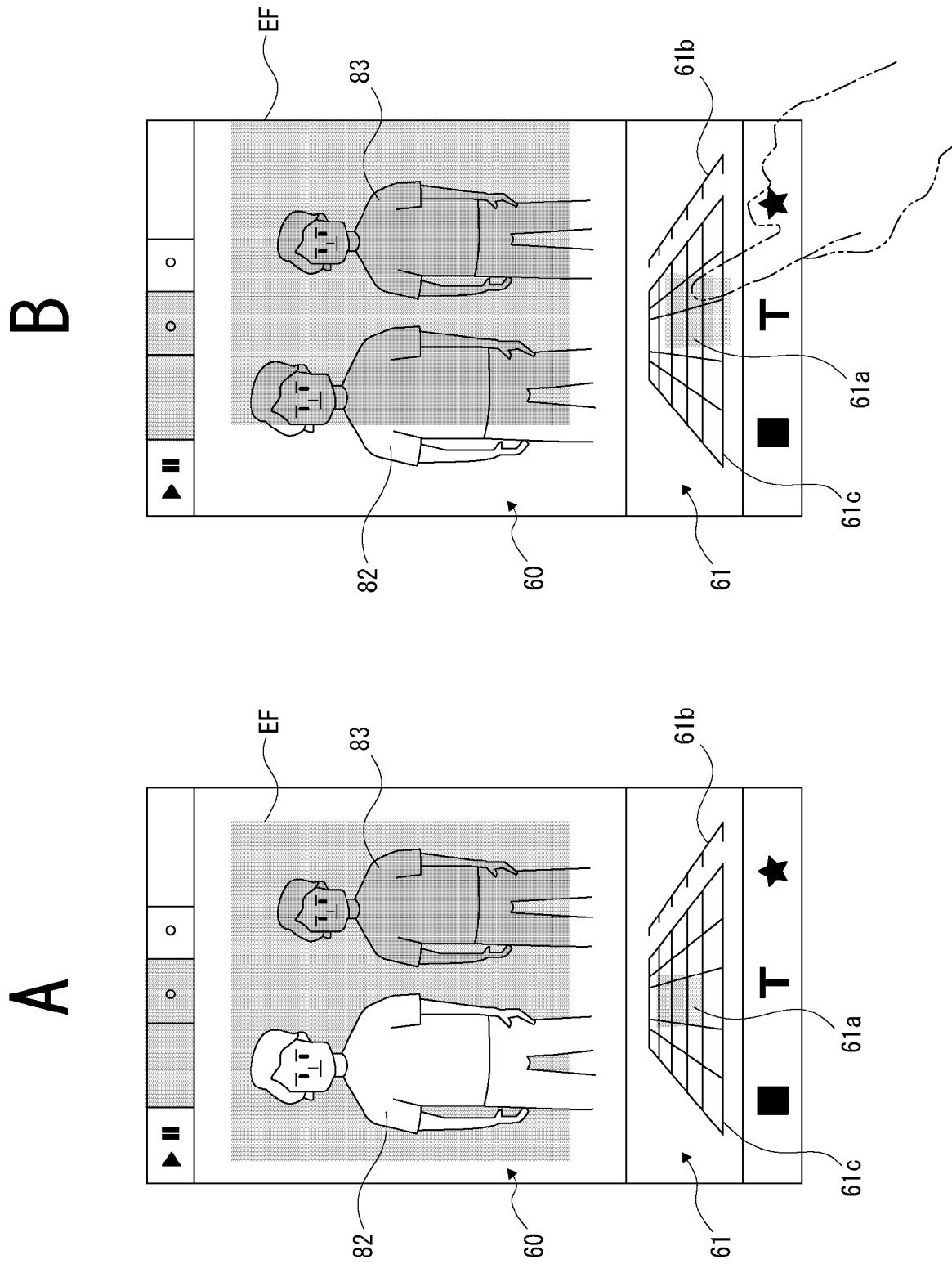
B



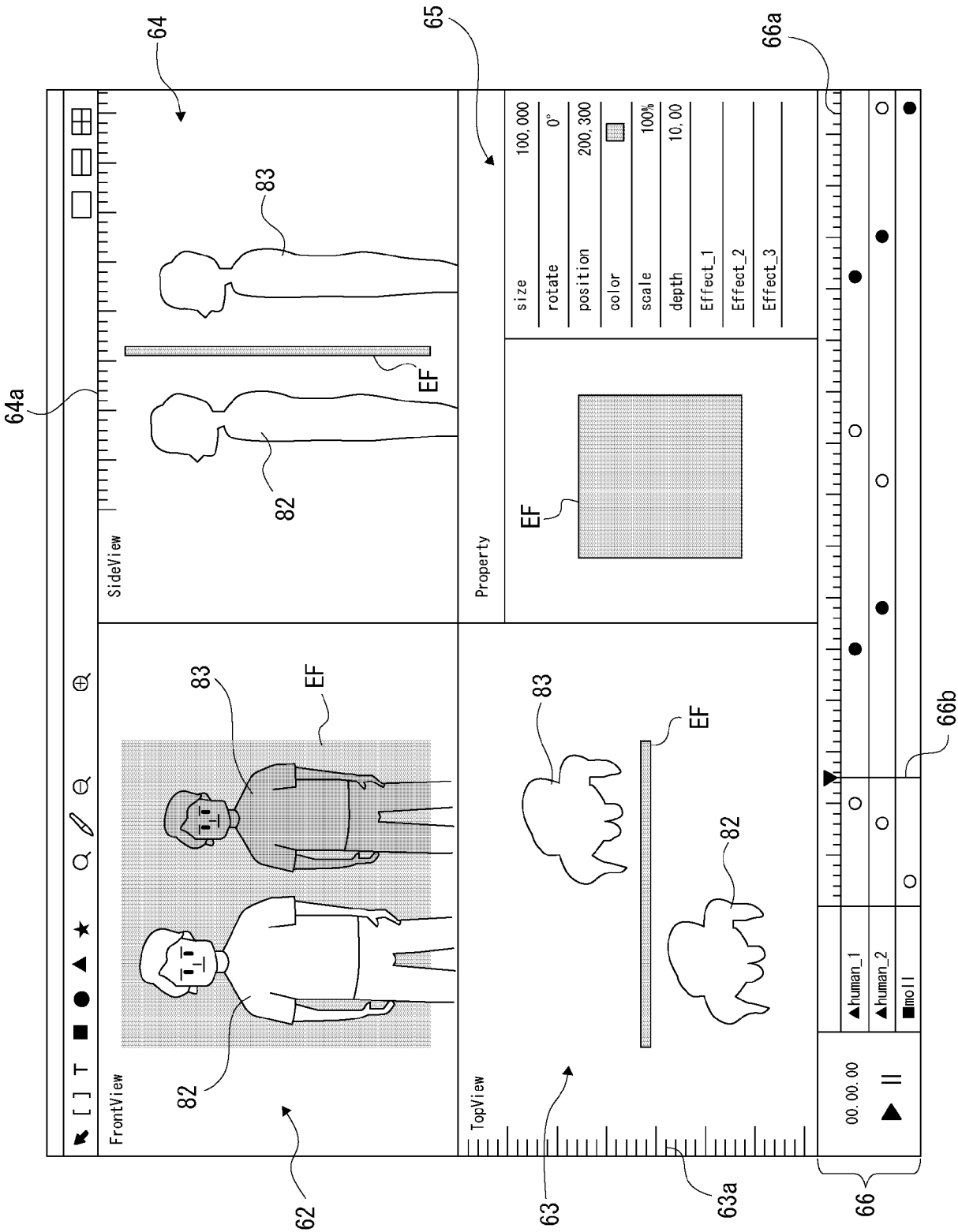
[図14]



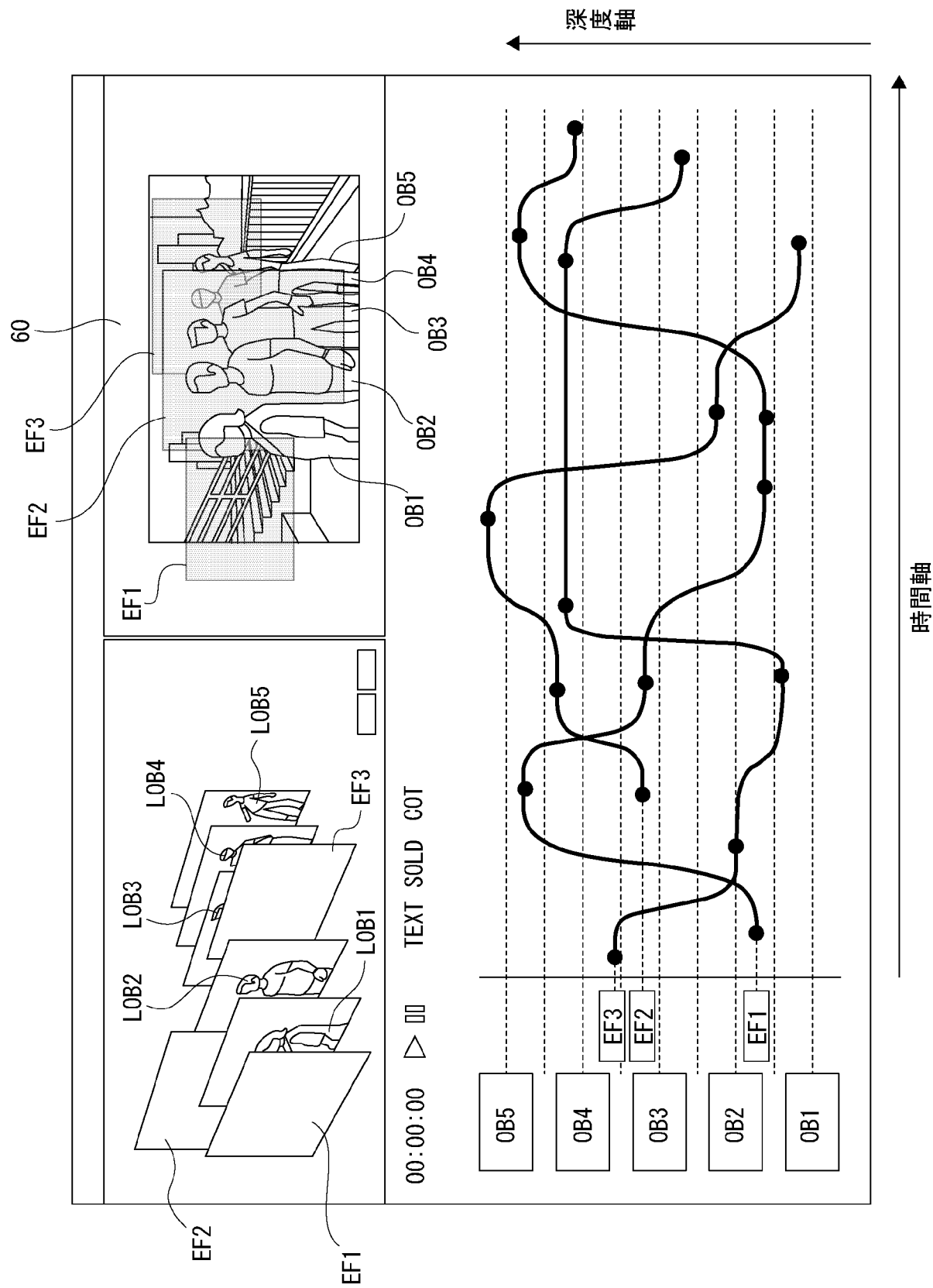
[図15]



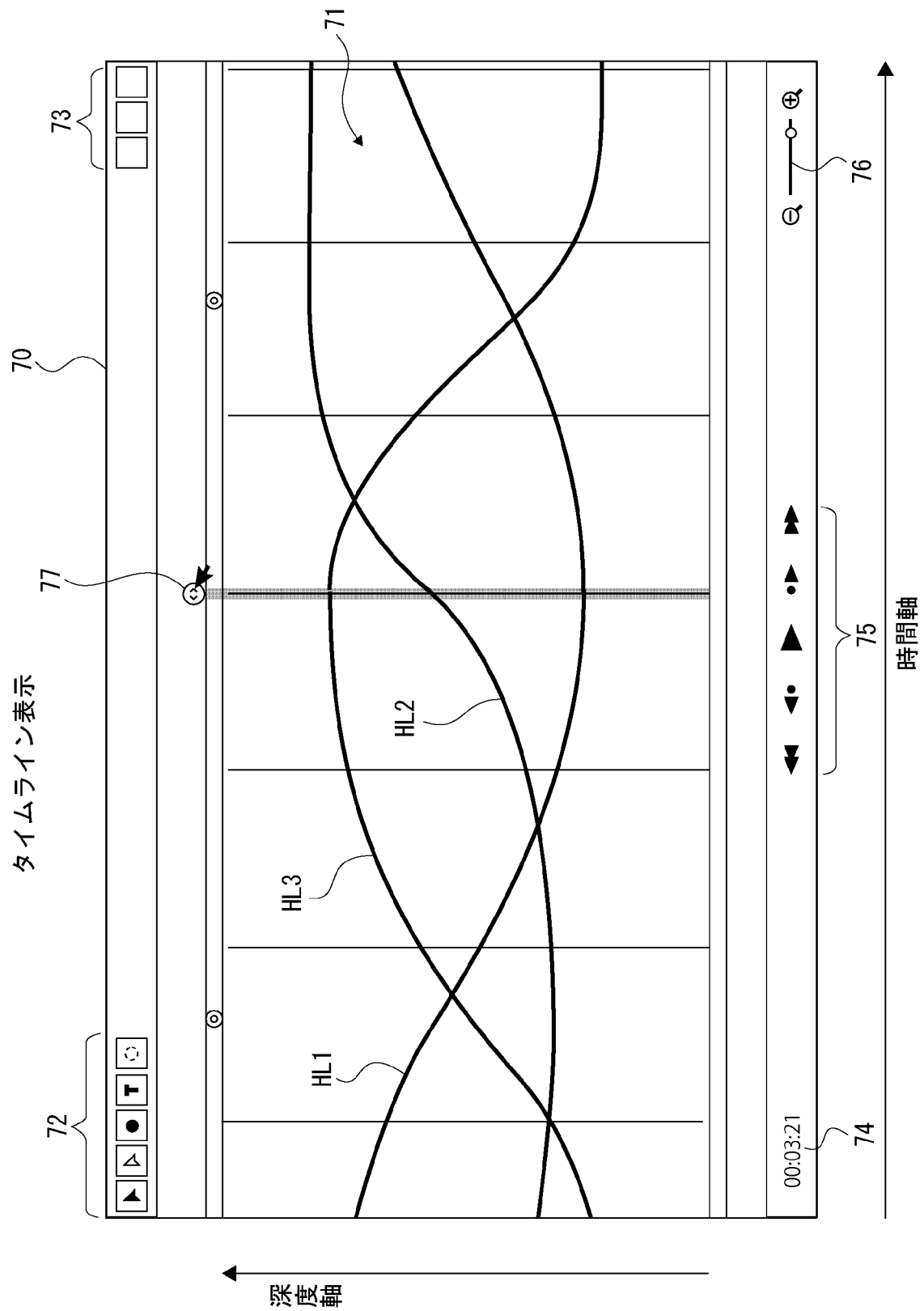
[図16]



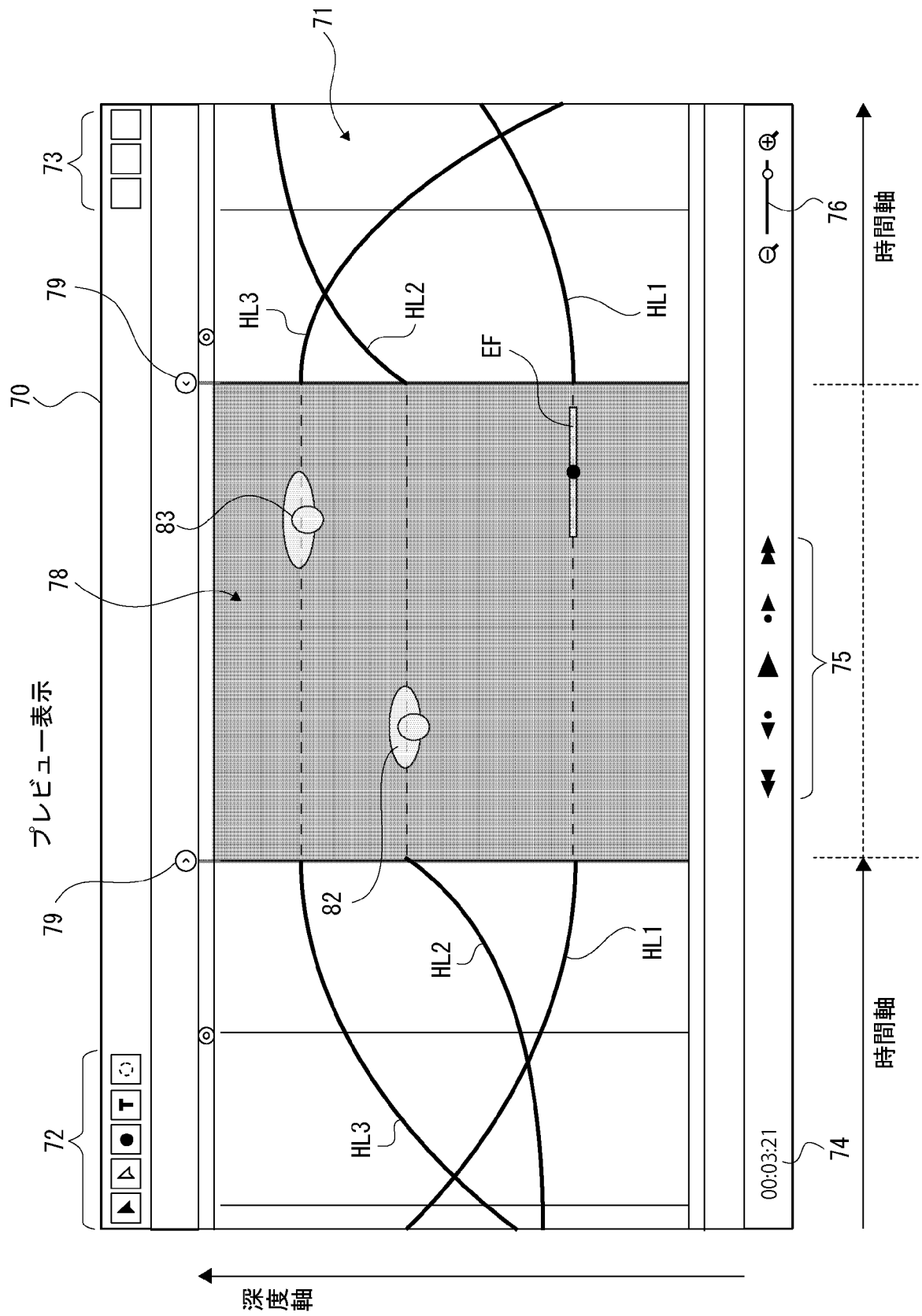
[図17]



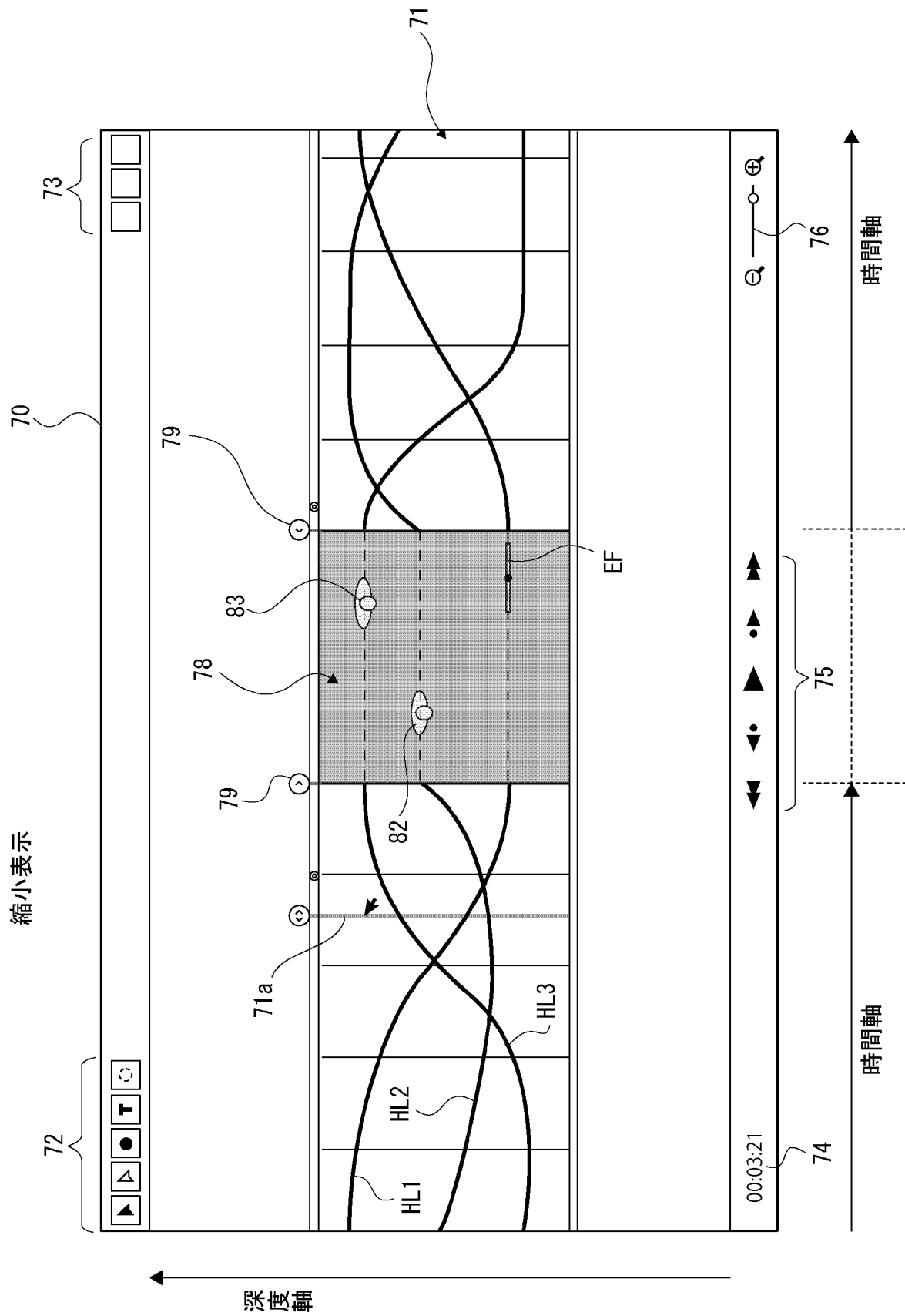
[図18]



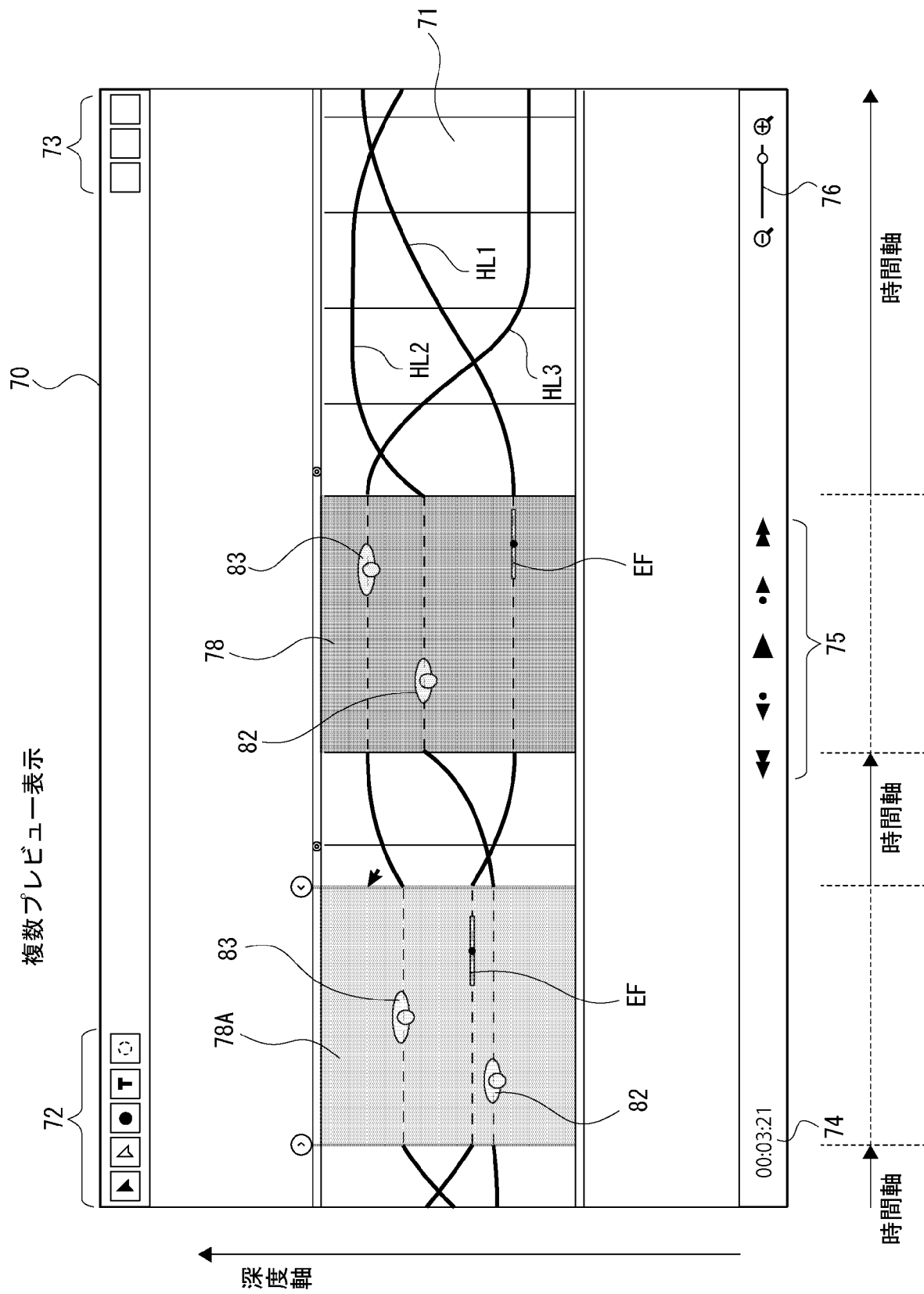
[図19]



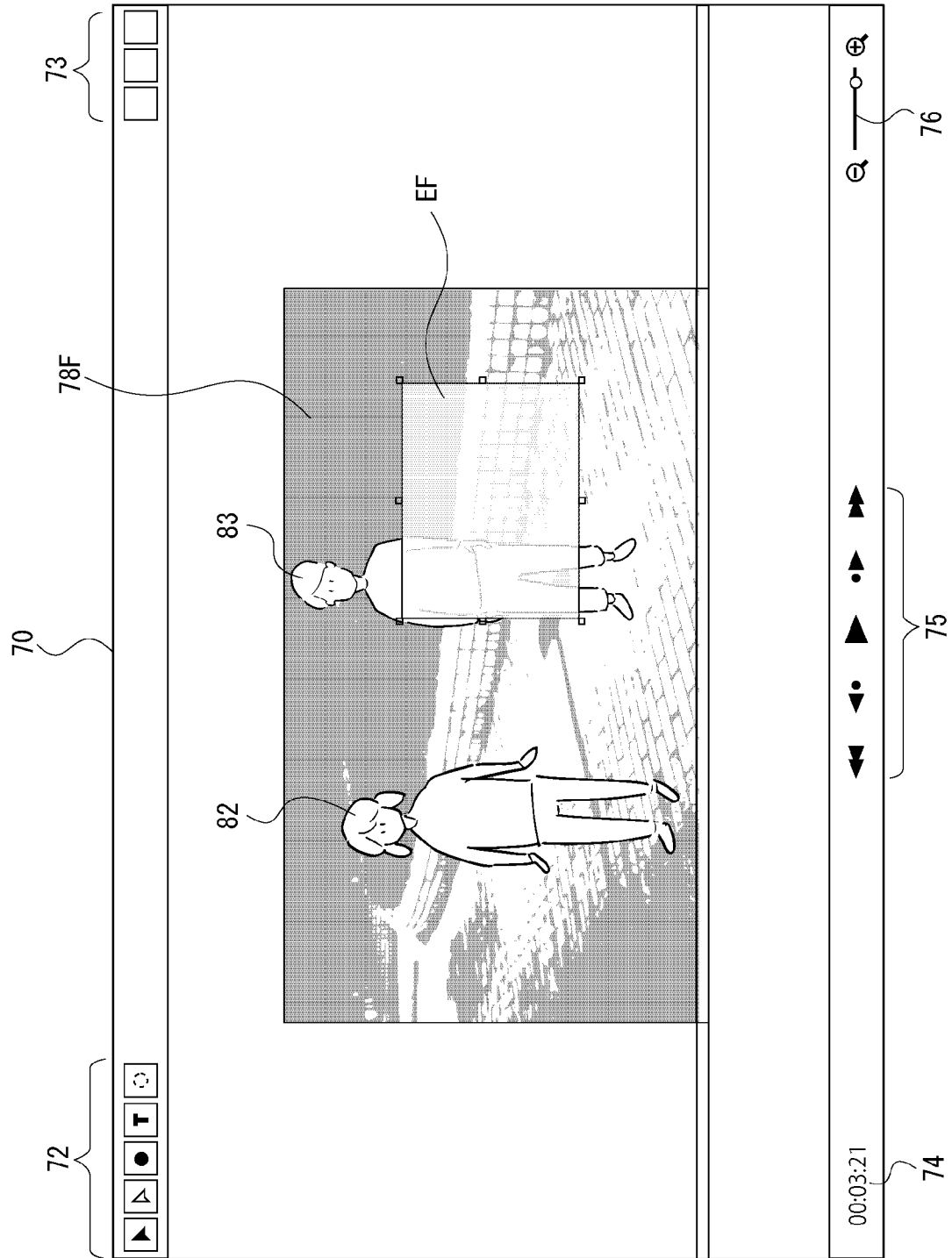
[図20]



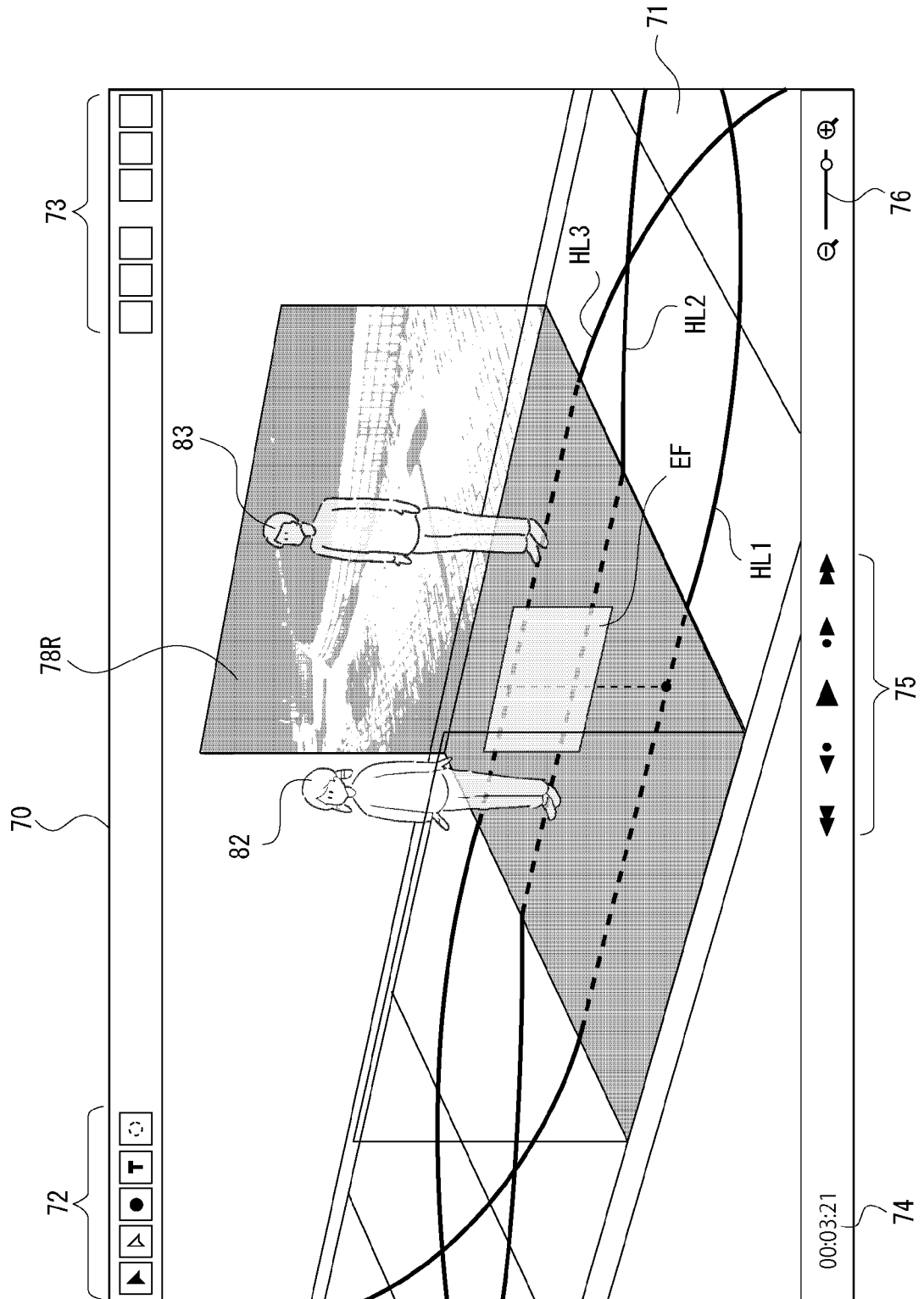
[図21]



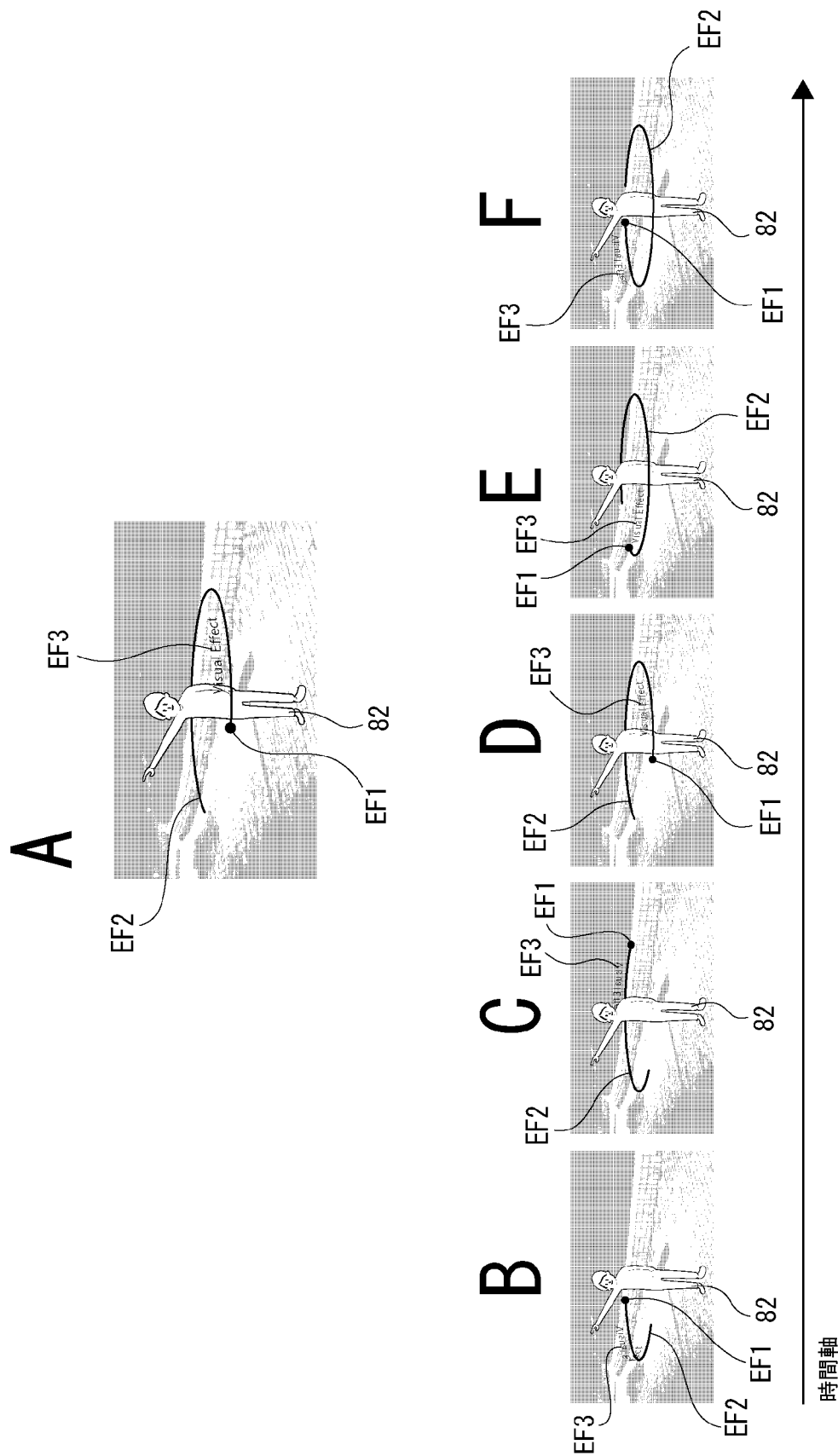
[図22]



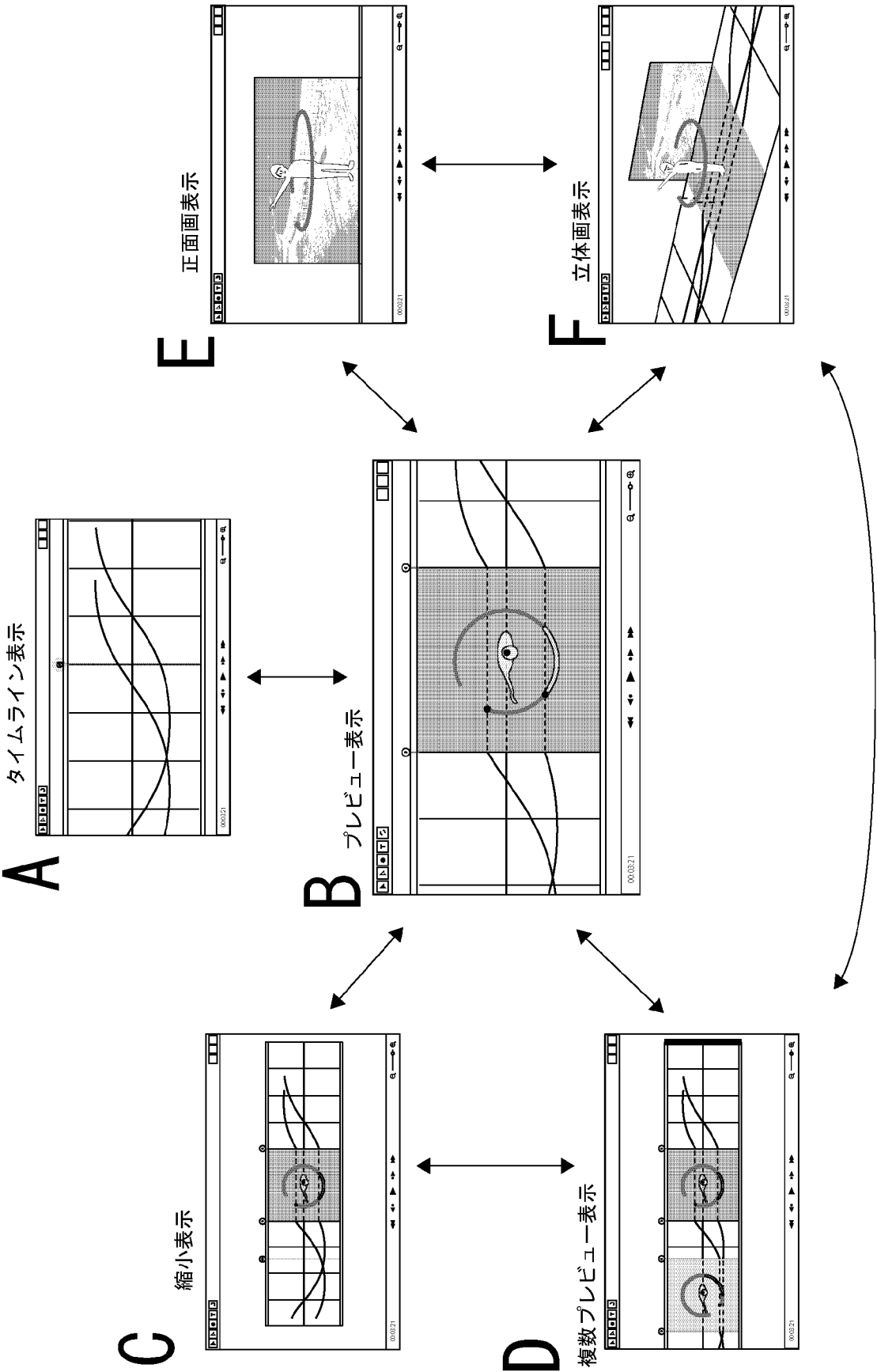
[図23]



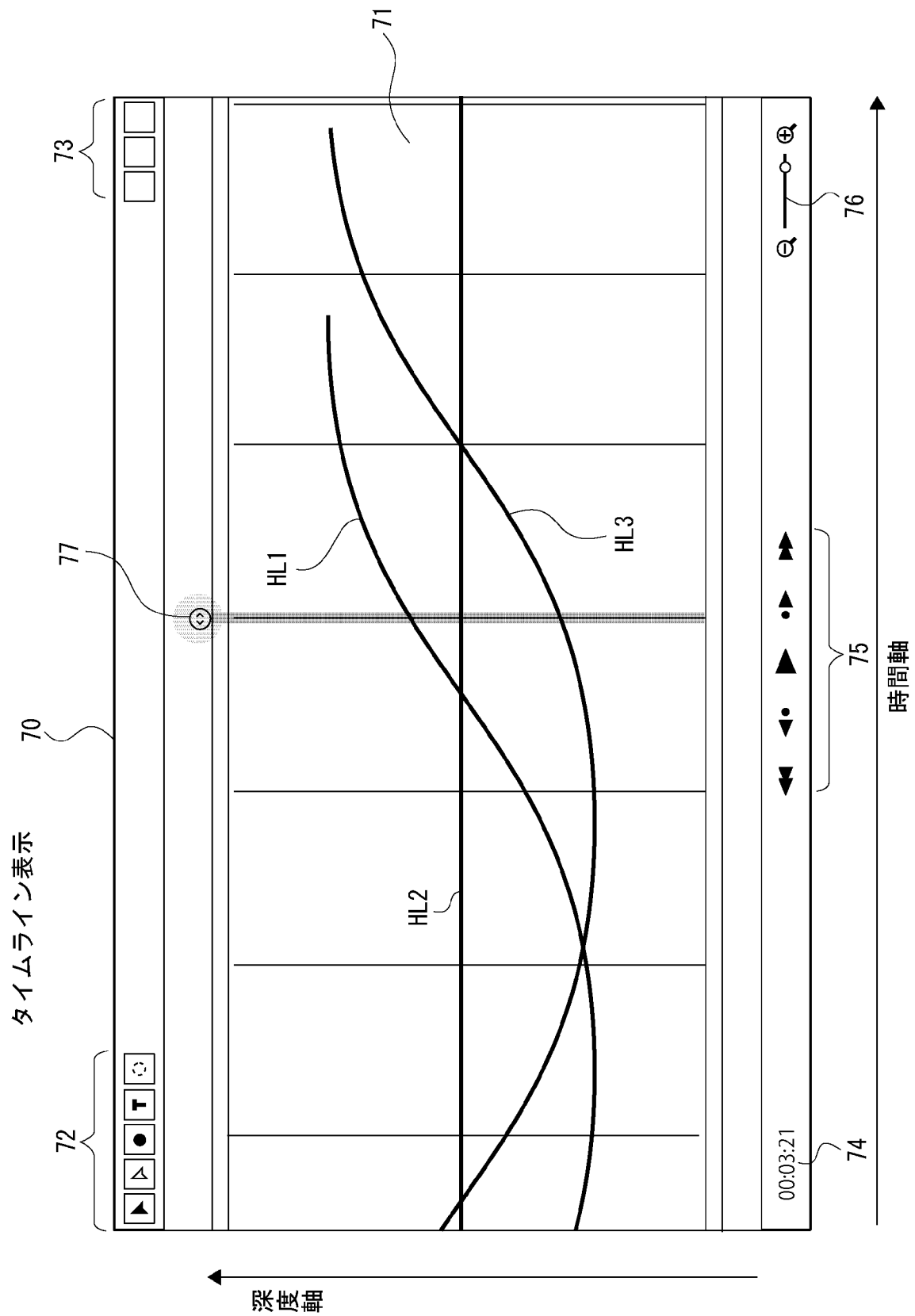
[図24]



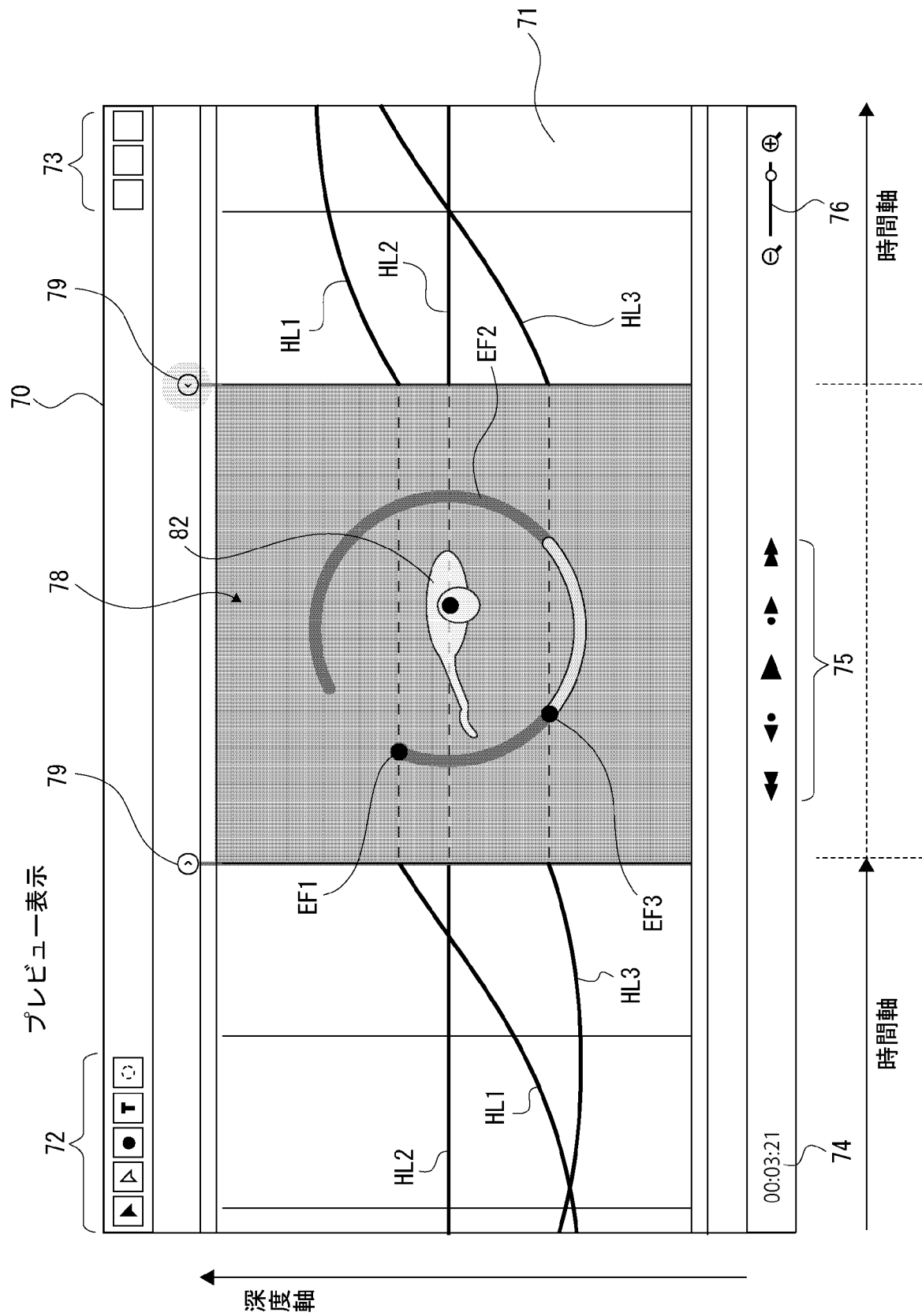
[図25]



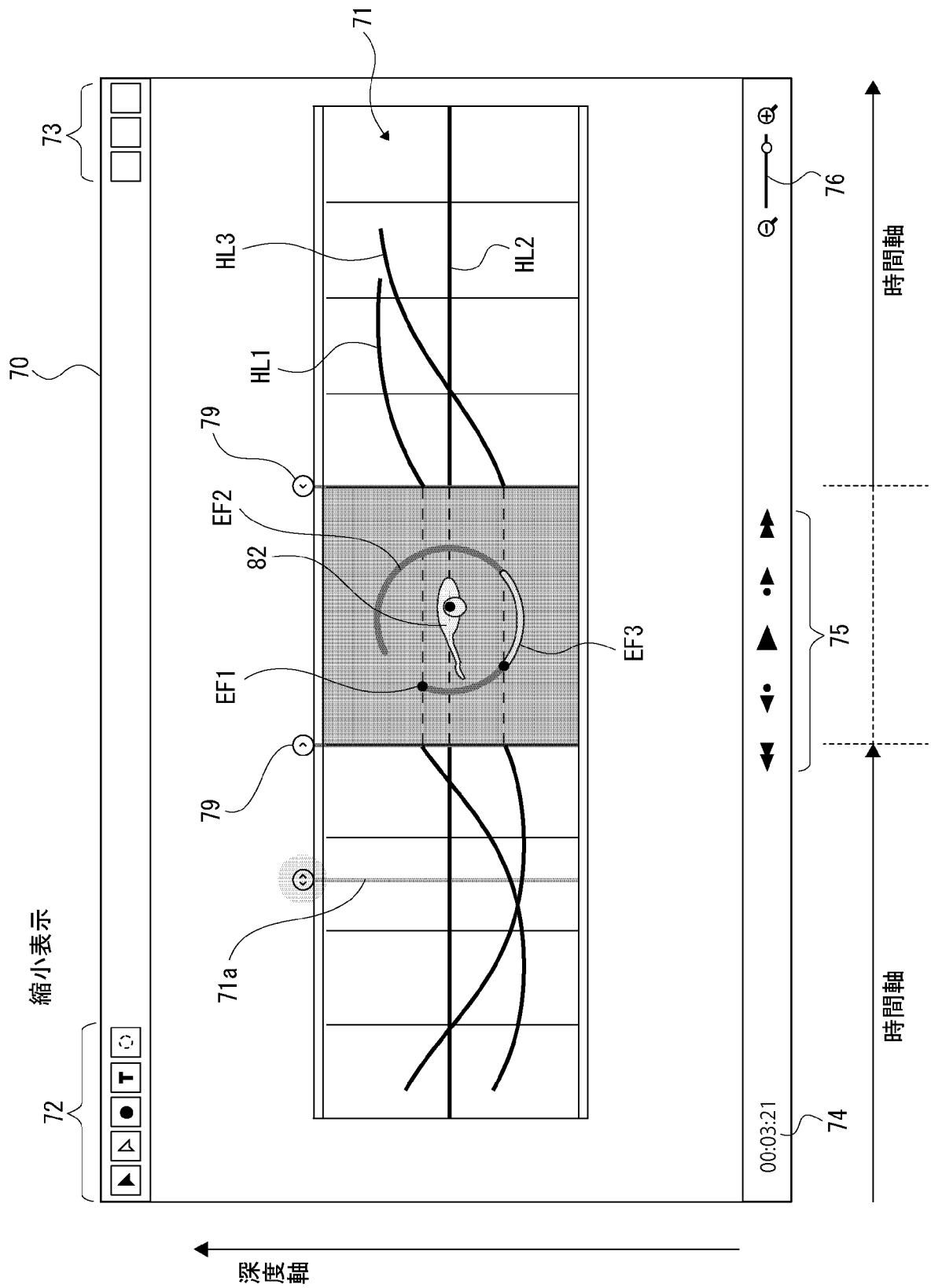
[図26]



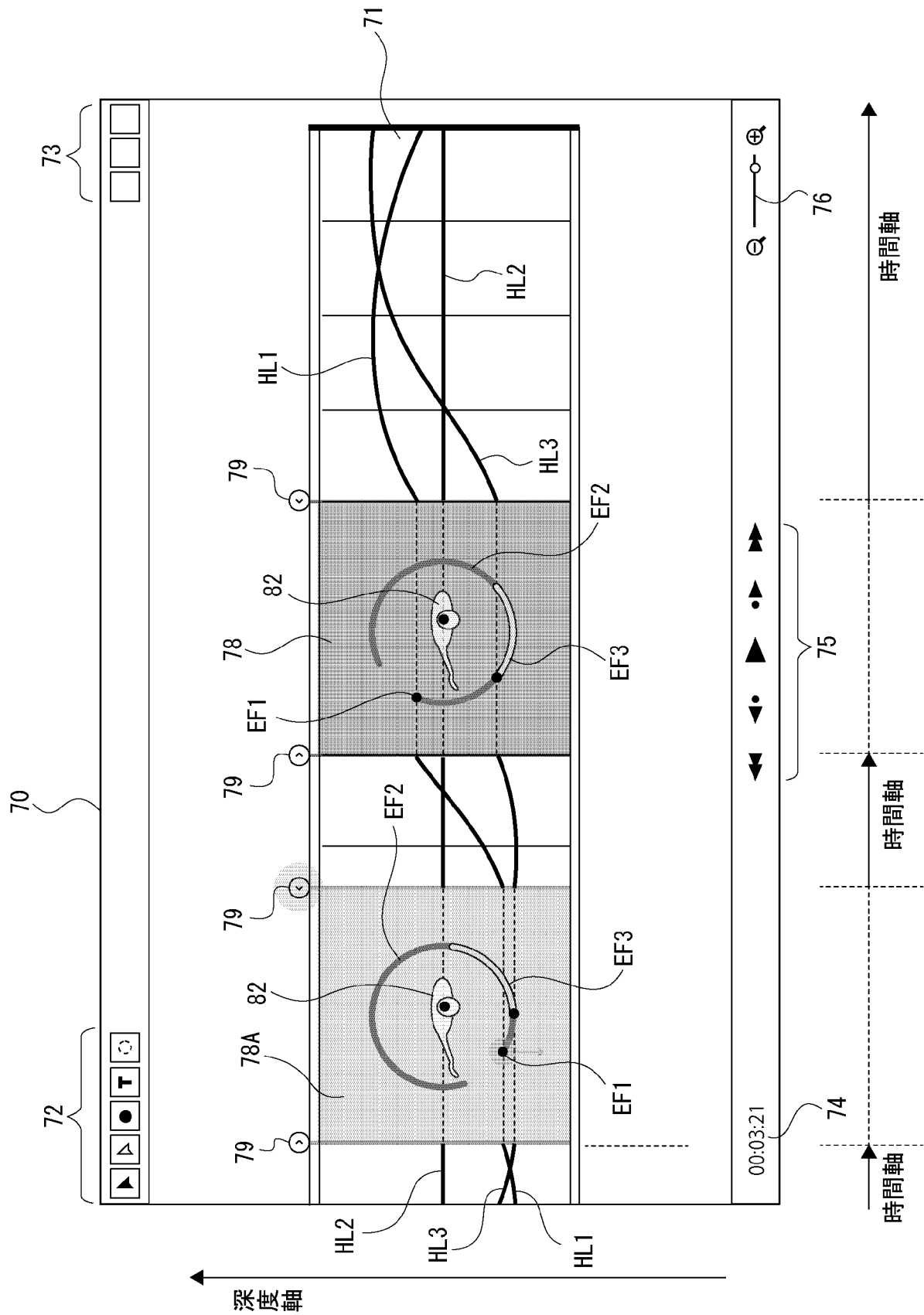
[図27]



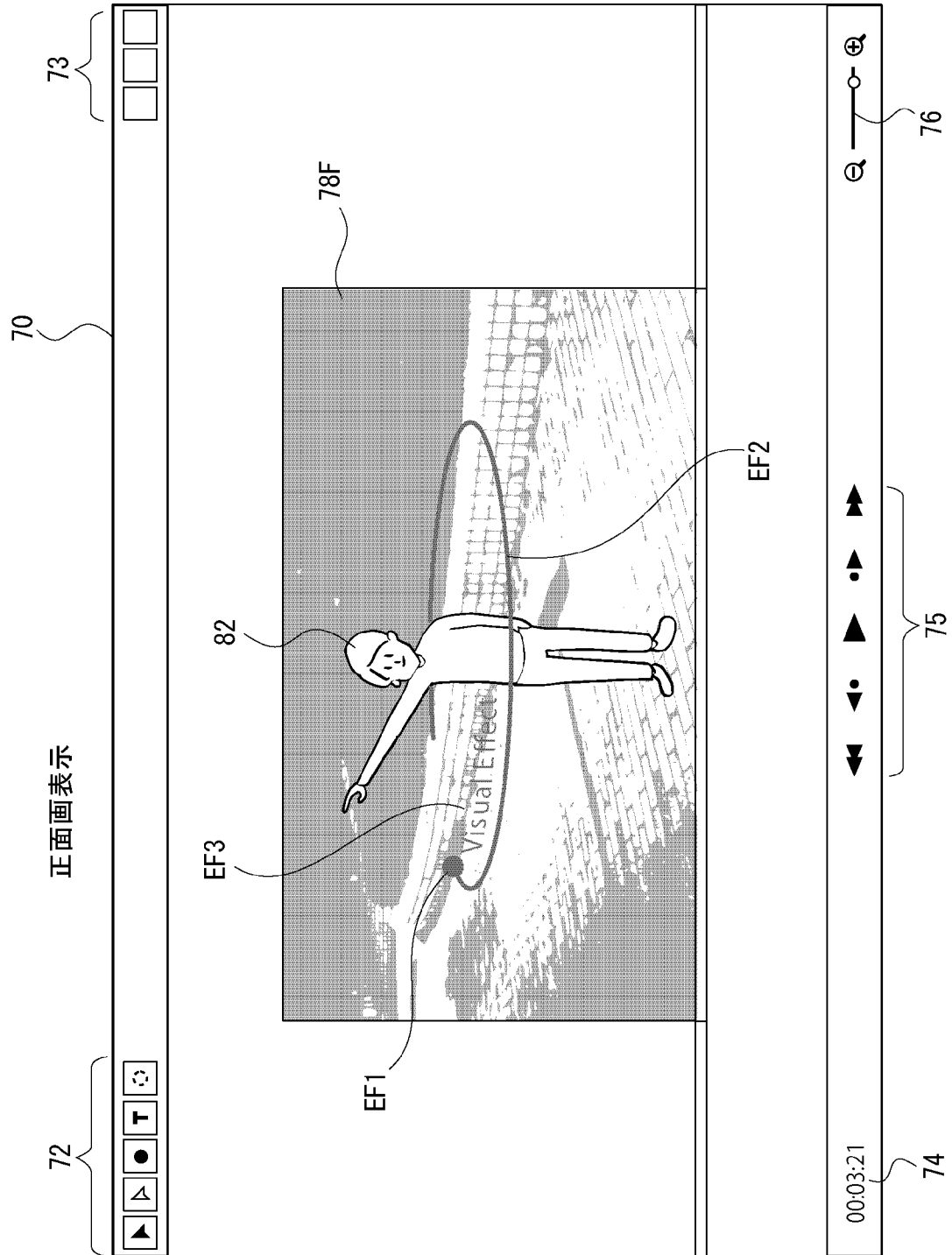
[図28]



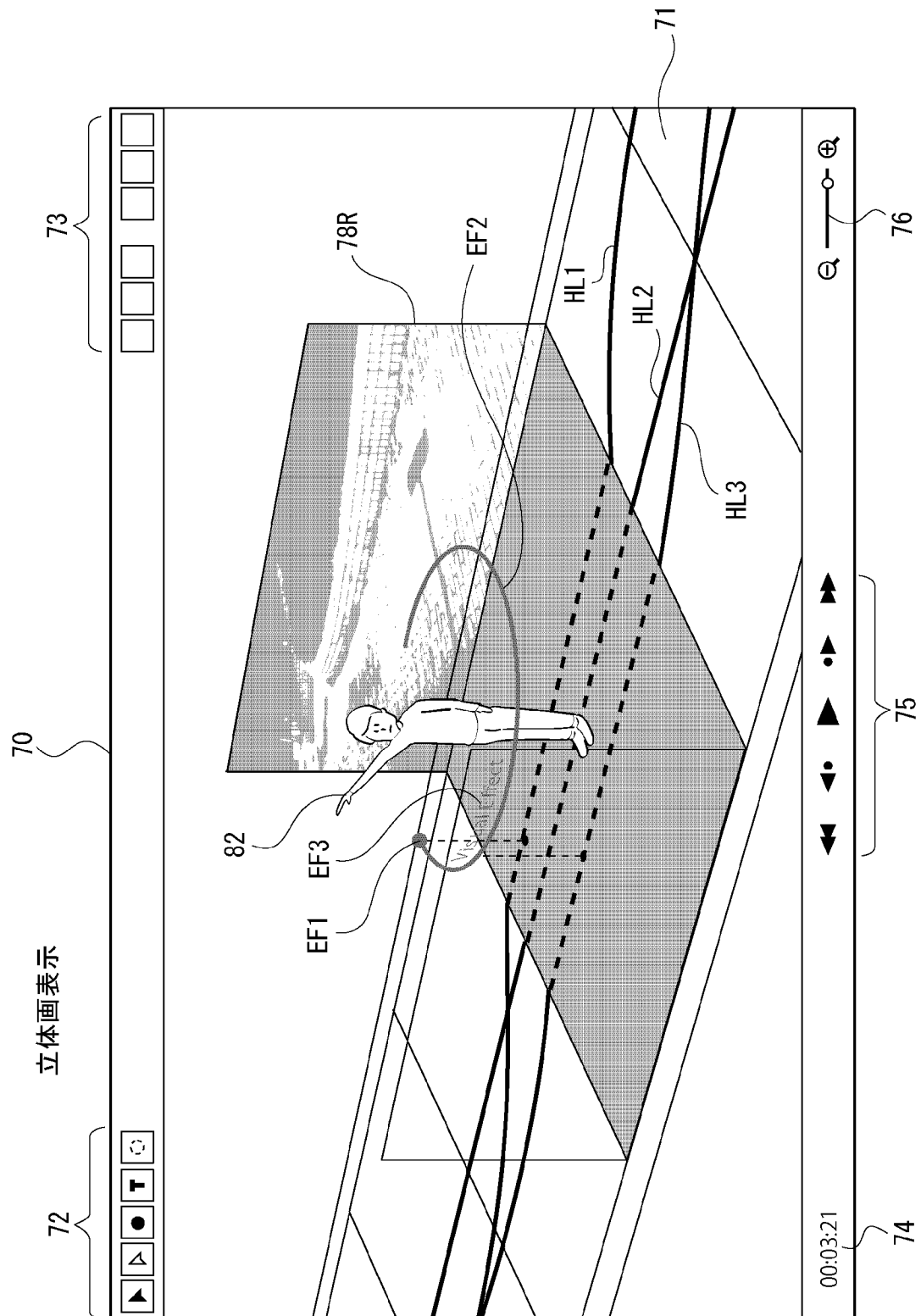
[図29]



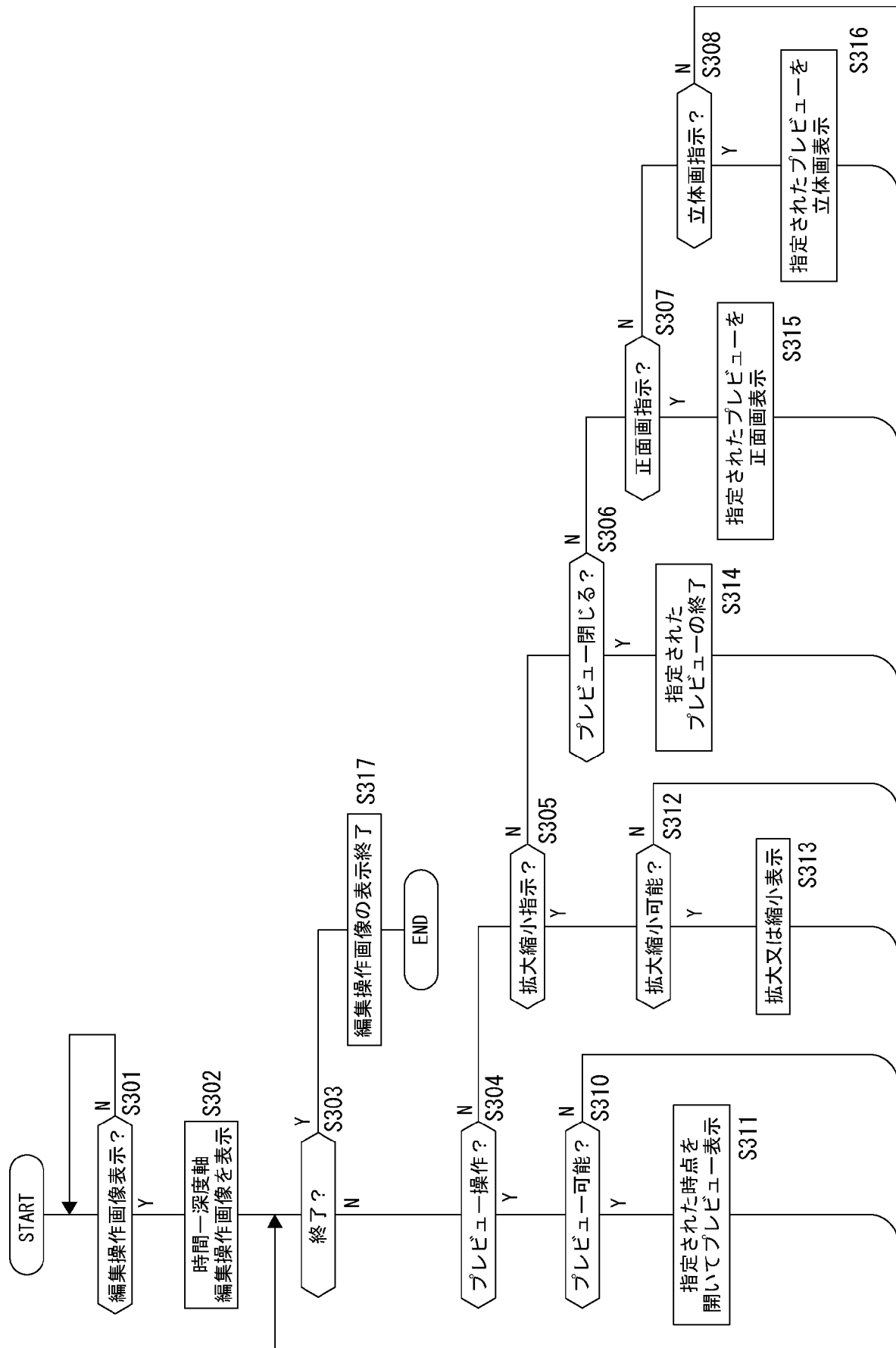
[図30]



[図31]



[図32]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/016198

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. H04N5/91(2006.01)i, G06T13/80(2011.01)i, G09G5/00(2006.01)i, G09G5/36(2006.01)i, G09G5/377(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl. H04N5/91, G06T13/80, G09G5/00, G09G5/36, G09G5/377

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2019
Registered utility model specifications of Japan	1996-2019
Published registered utility model applications of Japan	1994-2019

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	JP 2014-016670 A (NIKON CORPORATION) 30 January 2014, paragraphs [0014], [0022], [0039]-[0072], fig. 2, 4, 5 (Family: none)	1-2, 14-15, 19-20 3-13, 16-18
X Y A	JP 2004-145448 A (TOSHIBA CORPORATION) 20 May 2004, paragraphs [0033]-[0039], [0098]-[0117], fig. 16, 19, 20 (Family: none)	1-2, 4-5, 19- 20 3, 6-8 9-18



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
27.06.2019

Date of mailing of the international search report
09.07.2019

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/016198

C (Continuation). 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 2011-010128 A (CANON INC.) 13 January 2011, paragraphs [0022]-[0025], [0027], [0036], [0045], fig. 7 & US 2010/0328432 A1, paragraphs [0040]-[0043], [0045], [0054], [0063], [0064], fig. 7	1-2, 4-5, 13, 19-20 9-12 3, 6-8, 14-18
X A	JP 2013-118468 A (SONY CORPORATION) 13 June 2013, paragraphs [0014], [0019], [0033], [0036], [0037], fig. 3, 4 & US 2013/0142452 A1, paragraphs [0028], [0031], [0037], [0054], [0057], [0058], fig. 3, 4 & US 2016/0307361 A1 & CN 103136746 A	1-2, 19-20 3-18
Y A	JP 2015-156540 A (FUJITSU LIMITED) 27 August 2015, paragraphs [0058], [0100], [0114] (Family: none)	3, 6 1-2, 4-5, 7-20
Y A	WO 2014/013627 A1 (RAKUTEN, INC.) 23 January 2014, paragraphs [0084]-[0086] & US 2014/0340477 A1, paragraphs [0142]-[0144] & EP 2775452 A1 & JP 5254505 B1	3, 6 1-2, 4-5, 7-20
Y A	JP 2005-123824 A (JAPAN BROADCASTING CORPORATION) 12 May 2005, paragraphs [0080]-[0086], fig. 1, 2 (Family: none)	7-8 1-6, 9-20
Y A	JP 06-105231 A (NAMCO., LTD.) 15 April 1994, paragraphs [0018]-[0068], fig. 4, 11 (Family: none)	7-8 1-6, 9-20
Y A	JP 2002-374494 A (FUJI ELECTRIC CO., LTD.) 26 December 2002, paragraphs [0034], [0036], fig. 1, 4 (Family: none)	9-12 1-8, 13-20
Y A	JP 2007-027990 A (CANON INC.) 01 February 2007, paragraphs [0017]-[0026], [0076], [0077], fig. 6, 8 (Family: none)	9-12 1-8, 13-20
Y A	WO 2006/022071 A1 (MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.) 02 March 2006, paragraphs [0064]-[0067], [0069], fig. 21 (Family: none)	11-12 1-10, 13-20

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/016198

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X	WO 2018/096775 A1 (SONY CORPORATION) 31 May 2018, paragraphs [0082], [0083], [0092], [0093], [0114], fig. 1, 3 (Family: none)	1-2, 13 19-20

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. H04N5/91(2006.01)i, G06T13/80(2011.01)i, G09G5/00(2006.01)i, G09G5/36(2006.01)i, G09G5/377(2006.01)i

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int.Cl. H04N5/91, G06T13/80, G09G5/00, G09G5/36, G09G5/377

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	1971-2019年
日本国実用新案登録公報	1996-2019年
日本国登録実用新案公報	1994-2019年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X	JP 2014-016670 A (株式会社ニコン) 2014.01.30, 段落[0014], [0022], [0039]-[0072], 図 2, 4-5	1-2, 14-15, 19-20
A	(ファミリーなし)	3-13, 16-18
X	JP 2004-145448 A (株式会社東芝) 2004.05.20, 段落[0033]-[0039], [0098]-[0117], 図 16, 19, 20	1-2, 4-5, 19-20
Y	(ファミリーなし)	3, 6-8
A		9-18

☑ 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

27.06.2019

国際調査報告の発送日

09.07.2019

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)

郵便番号 100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

川中 龍太

5 C

5892

電話番号 03-3581-1101 内線 3541

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
X Y A	JP 2011-010128 A (キヤノン株式会社) 2011.01.13, 段落[0022]-[0025], [0027], [0036], [0045], 図 7 & US 2010/0328432 A1, paragraphs [0040]-[0043], [0045], [0054], [0063]-[0064], Fig. 7	1-2, 4-5, 13, 19-20 9-12 3, 6-8, 14-18
X A	JP 2013-118468 A (ソニー株式会社) 2013.06.13, 段落[0014], [0019], [0033], [0036]-[0037], 図 3, 4 & US 2013/0142452 A1, paragraphs [0028], [0031], [0037], [0054], [0057]-[0058], Figs 3, 4 & US 2016/0307361 A1 & CN 103136746 A	1-2, 19-20 3-18
Y A	JP 2015-156540 A (富士通株式会社) 2015.08.27, 段落[0058], [0100], [0114] (ファミリーなし)	3, 6 1-2, 4-5, 7-20
Y A	WO 2014/013627 A1 (楽天株式会社) 2014.01.23, 段落[0084]-[0086] & US 2014/0340477 A1, paragraphs [0142]-[0144] & EP 2775452 A1 & JP 5254505 B1	3, 6 1-2, 4-5, 7-20
Y A	JP 2005-123824 A (日本放送協会) 2005.05.12, 段落[0080]-[0086], 図 1, 2 (ファミリーなし)	7-8 1-6, 9-20
Y A	JP 06-105231 A (株式会社ナムコ) 1994.04.15, 段落[0018]-[0068], 図 4, 11 (ファミリーなし)	7-8 1-6, 9-20
Y A	JP 2002-374494 A (富士電機株式会社) 2002.12.26, 段落[0034], [0036], 図 1, 4 (ファミリーなし)	9-12 1-8, 13-20
Y A	JP 2007-027990 A (キヤノン株式会社) 2007.02.01, 段落[0017]-[0026], [0076]-[0077], 図 6, 8 (ファミリーなし)	9-12 1-8, 13-20
Y A	WO 2006/022071 A1 (松下電器産業株式会社) 2006.03.02, 段落[0064]-[0067], [0069], 図 21 (ファミリーなし)	11-12 1-10, 13-20

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
P, X	WO 2018/096775 A1 (ソニー株式会社) 2018.05.31, 段落[0082]-[0083], [0092]-[0093], [0114], 図 1, 3 (ファミリーなし)	1-2, 13, 19-20