

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11)特許出願公開番号

**特開2013-25748**

(P2013-25748A)

(43) 公開日 平成25年2月4日(2013.2.4)

(51) Int.Cl.

**G O 6 F 17/30 (2006.01)**

F 1

G06 F 17/30 350 C

G06F 17/30 170D

テーマコード (参考)

審査請求 未請求 請求項の数 14 O L (全 40 頁)

(21) 出願番号 特願2011-163044 (P2011-163044)

(22) 出願日 平成23年7月26日 (2011. 7. 26)

(71) 出願人 000002185

ソニー株式会社

東京都港区港南1丁目7番1号

(74) 代理人 100095957

弁理士 亀谷 美明

(74) 代理人 100096389

弁理士 金本 哲男

(74) 代理人 100101557

弁理士 萩原 康司

(74) 代理人 100128587

弁理士 松本 一騎

(72) 発明者 佐藤 浩司

東京都港区港南1丁目7番1号 ソニー株  
式会社内

[最終頁に続く](#)

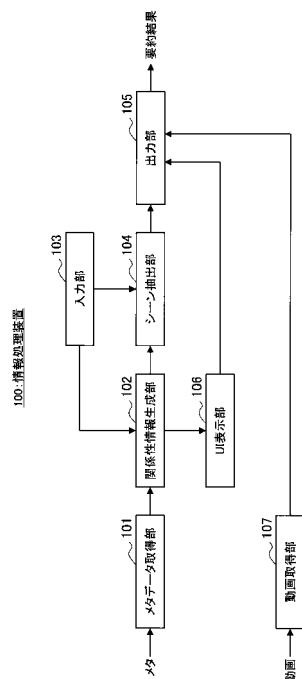
(54) 【発明の名称】 情報処理装置、動画要約方法、及びプログラム

(57) 【要約】

【課題】登場物の選択内容に適合する要約結果が得られる動画要約手法を実現すること。

【解決手段】動画に登場する各登場物間の関係性を示す関係性情報と、前記動画の中で各登場物が登場する区間を示す区間メタデータとを利用し、ユーザにより選択された登場物との間に所定の関係性を有する登場物が登場する区間を検出する区間検出部と、前記区間検出部により検出された区間に対応する画像を出力する画像出力部と、を備える、情報処理装置が提供される。

【選択図】図14



**【特許請求の範囲】****【請求項 1】**

動画に登場する各登場物間の関係性を示す関係性情報と、前記動画の中で各登場物が登場する区間を示す区間メタデータとを利用し、ユーザにより選択された登場物との間に所定の関係性を有する登場物が登場する区間を検出する区間検出部と、

前記区間検出部により検出された区間に対応する画像を出力する画像出力部と、  
を備える、

情報処理装置。

**【請求項 2】**

前記関係性情報は、前記関係性の強さを示す関係性値を含み、

10

前記区間検出部は、前記ユーザにより選択された登場物との間における関係性値の総和が最大となるように所定数の登場物を選択し、選択した登場物が登場する所定数の区間を検出する、

請求項 1 に記載の情報処理装置。

**【請求項 3】**

前記関係性情報は、前記関係性の強さを示す関係性値を含み、

前記区間検出部は、

前記ユーザにより選択された第 1 の登場物が登場する第 1 の区間を検出する処理と、

$n = 2 \sim N$  ( $N \geq 2$ ) に関し、第  $n - 1$  の区間の次に再生される第  $n$  の区間に登場する第  $n$  の登場物を選択する処理と、

20

前記第  $n - 1$  の登場物と前記第  $n$  の登場物との間における関係性値を  $n = 2 \sim N$  に関して積算し、前記関係性値の積算値を算出する処理と、

前記関係性値の積算値が最大となる第 2 ～第  $N$  の登場物を選択する処理と、

選択した前記第  $n$  の登場物 ( $n = 2 \sim N$ ) が登場する第  $n$  の区間を検出する処理と、  
を実行し、前記第 1 ～第  $N$  の区間を検出結果として出力する、

請求項 1 に記載の情報処理装置。

**【請求項 4】**

前記関係性情報は、前記関係性の強さを示す関係性値を含み、

前記区間検出部は、

前記ユーザにより選択された登場物が登場する第 1 の区間を検出する処理と、

30

$n = 2 \sim N$  ( $N \geq 2$ ) に関し、第  $n - 1$  の区間以降の各区間について、各区間内に登場する全ての登場物と、前記第  $n - 1$  の区間に登場する全ての登場物との間における関係性値の総和が最大になる区間を検出して第  $n$  の区間に設定する処理と、

を実行し、前記第 1 ～第  $N$  の区間を検出結果として出力する、

請求項 1 に記載の情報処理装置。

**【請求項 5】**

前記関係性情報は、前記関係性の強さを示す関係性値を含み、

前記区間検出部は、

前記ユーザにより選択された登場物が登場する第 1 の区間を検出する処理と、

$n = 2 \sim N$  ( $N \geq 2$ ) に関し、第  $n - 1$  の区間以降の各区間について、各区間内に登場する全ての登場物と、前記ユーザにより選択された登場物との間における関係性値の総和を算出し、算出した関係性値の総和が前記第  $n - 1$  の区間に関する関係性値の総和に最も近くなる区間を検出して第  $n$  の区間に設定する処理と、

40

を実行し、前記第 1 ～第  $N$  の区間を検出結果として出力する、

請求項 1 に記載の情報処理装置。

**【請求項 6】**

前記関係性情報は、前記関係性の強さを示す関係性値を含み、

前記区間検出部は、前記ユーザにより選択された登場物との間における関係性値の総和が最小となるように所定数の登場物を選択し、選択した登場物が登場する所定数の区間を検出する、

50

請求項 1 に記載の情報処理装置。

【請求項 7】

前記区間検出部は、前記関係性情報に含まれる少なくとも 1 つの関係性値が変更された場合に、変更後の前記関係性情報に基づいて前記区間を再検出し、

前記画像出力部は、前記区間検出部による再検出の結果に応じて出力内容を更新する、  
請求項 2 に記載の情報処理装置。

【請求項 8】

前記ユーザにより選択された登場物に対応する第 1 のオブジェクトを操作領域の中心に表示し、当該登場物との間における関係性値の大きさに応じた距離だけ前記第 1 のオブジェクトから離れた位置に、一又は複数の他の登場物にそれぞれ対応する一又は複数の第 2 のオブジェクトを表示するオブジェクト表示部と、

前記操作領域において前記第 2 のオブジェクトの位置が変更された場合に、変更後の当該第 2 のオブジェクトの位置に応じた関係性値を算出し、算出した関係性値により前記関係性情報を更新する関係性情報更新部と、

をさらに備える、

請求項 7 に記載の情報処理装置。

【請求項 9】

同じ作品に登場する頻度が高いほど、或いは、共に登場する区間の長さが長いほど、前記関係性値が大きくなるように、前記区間メタデータを利用して各登場物の組に関する関係性値を算出して前記関係性情報を生成する関係性情報生成部をさらに備える、

請求項 2 に記載の情報処理装置。

【請求項 10】

前記関係性情報は、前記関係性の種類を示した第 1 ～第 N - 1 の種類情報を含み、

前記区間検出部は、

前記ユーザにより選択された登場物が登場する第 1 の区間を検出する処理と、

$n = 1 \sim N - 1$  に関し、前記ユーザにより選択された登場物との間において第  $n$  の種類情報が示す種類の関係性を有する第  $n$  の登場物を検出し、第  $n$  の区間以降の区間の中から当該第  $n$  の登場物が登場する区間を検出して第  $n + 1$  の区間に設定する処理と、

を実行し、前記第 1 ～第 N の区間を検出結果として出力する、

請求項 1 に記載の情報処理装置。

【請求項 11】

予め用意された複数の動画における特徴的な関係性の時系列変化を検出し、検出した特徴的な関係性の時系列変化に適合するように前記第 1 ～第 N - 1 の種類情報を設定する種類情報設定部をさらに備える、

請求項 10 に記載の情報処理装置。

【請求項 12】

前記第 1 ～第 N - 1 の種類情報には、それぞれカメラワーク又はエフェクトの種類を示す補助情報が対応付けられており、

前記区間検出部は、

前記ユーザにより選択された登場物が登場する第 1 の区間を検出する処理と、

$n = 1 \sim N - 1$  に関し、前記ユーザにより選択された登場物との間において第  $n$  の種類情報が示す種類の関係性を有する第  $n$  の登場物を検出し、第  $n$  の区間以降の区間の中から当該第  $n$  の登場物が登場する区間であり、かつ、前記補助情報に適合する区間を検出して第  $n + 1$  の区間に設定する処理と、

を実行し、前記第 1 ～第 N の区間を検出結果として出力する、

請求項 10 に記載の情報処理装置。

【請求項 13】

動画に登場する各登場物間の関係性を示す関係性情報と、前記動画の中で各登場物が登場する区間を示す区間メタデータとを利用し、ユーザにより選択された登場物との間に所定の関係性を有する登場物が登場する区間を検出するステップと、

検出された区間に対応する画像を出力するステップと、を含む、

動画要約方法。

【請求項 14】

動画に登場する各登場物間の関係性を示す関係性情報と、前記動画の中で各登場物が登場する区間を示す区間メタデータとを利用し、ユーザにより選択された登場物との間に所定の関係性を有する登場物が登場する区間を検出する区間検出機能と、

前記区間検出機能により検出された区間に対応する画像を出力する画像出力機能と、をコンピュータに実現させるためのプログラム。

10

【発明の詳細な説明】

【技術分野】

【0001】

本技術は、情報処理装置、動画要約方法、及びプログラムに関する。

【背景技術】

【0002】

テレビジョン映像やビデオ映像など、人々は、様々な動画を視聴する機会に恵まれている。また、広域通信網の整備が進み、多くの家庭では、インターネットを介してストリーミング配信される動画（以下、ストリーミング映像）を視聴することも可能になっている。こうした動画を視聴する際、ユーザは、リモートコントローラ（以下、リモコン）を操作して放送局を選択したり、ビデオ映像の再生を開始させたりする。視聴する動画が録画映像、DVDビデオ映像、Blu-rayビデオ映像、ストリーミング映像などの場合、ユーザは、再生を開始する位置を指定したり、一部の映像シーンをスキップさせたりすることができる。例えば、下記の特許文献1には、簡単な操作でCMをスキップ再生できるようにする技術が開示されている。

20

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特開2010-277661号公報

【発明の概要】

30

【発明が解決しようとする課題】

【0004】

しかしながら、動画に登場する人物の中から指定した人物に関連する人物が登場する映像シーンを集めたダイジェスト映像を自動再生できるようにしたり、それら映像シーンの概要を俯瞰的に把握できるように表示したりする技術は知られていない。そこで、本技術は、上記のような事情を受けて考案されたものであり、選択した登場人物を基準として登場人物間の関係性を考慮したシーン群の選択方法や、そのシーン群を選択的に再生する仕組みを実現することが可能な、新規かつ改良された情報処理装置、動画要約方法、及びプログラムを提供することを意図するものである。また、本技術は、動画に登場する物品間、或いは、物品と人物との間の関係性をさらに考慮したシーン群の選択方法を実現すること

40

【課題を解決するための手段】

【0005】

本技術のある観点によれば、動画に登場する各登場物間の関係性を示す関係性情報と、前記動画の中で各登場物が登場する区間を示す区間メタデータとを利用し、ユーザにより選択された登場物との間に所定の関係性を有する登場物が登場する区間を検出する区間検出部と、前記区間検出部により検出された区間に対応する画像を出力する画像出力部と、を備える、情報処理装置が提供される。

【0006】

また、本技術の別の観点によれば、動画に登場する各登場物間の関係性を示す関係性情

50

報と、前記動画の中で各登場物が登場する区間を示す区間メタデータとを利用し、ユーザにより選択された登場物との間に所定の関係性を有する登場物が登場する区間を検出するステップと、検出された区間に対応する画像を出力するステップと、を含む、動画要約方法が提供される。

【0007】

また、本技術の別の観点によれば、動画に登場する各登場物間の関係性を示す関係性情報と、前記動画の中で各登場物が登場する区間を示す区間メタデータとを利用し、ユーザにより選択された登場物との間に所定の関係性を有する登場物が登場する区間を検出する区間検出機能と、前記区間検出機能により検出された区間に対応する画像を出力する画像出力機能と、をコンピュータに実現させるためのプログラムが提供される。

10

【0008】

また、本技術の別の観点によれば、上記のプログラムが記録された、コンピュータにより読み取り可能な記録媒体が提供される。

【発明の効果】

【0009】

以上説明したように本技術によれば、選択した登場人物を基準として登場人物間の関係性を考慮したシーン群の選択方法や、そのシーン群を選択的に再生する仕組みを実現することが可能になる。また、本技術によれば、動画に登場する物品間、或いは、物品と人物との間の関係性をさらに考慮したシーン群の選択方法を実現することも可能になる。

【図面の簡単な説明】

20

【0010】

【図1】動画要約手法の概要について説明するための説明図である。

【図2】動画要約手法の概要について説明するための説明図である。

【図3】動画要約手法の概要について説明するための説明図である。

【図4】動画要約手法の概要について説明するための説明図である。

【図5】関係性値の算出方法について説明するための説明図である。

【図6】関係性値の算出方法について説明するための説明図である。

【図7】関係性値の算出方法について説明するための説明図である。

【図8】関係性値の算出方法について説明するための説明図である。

【図9】関係性値の算出方法について説明するための説明図である。

30

【図10】関係性値の算出方法について説明するための説明図である。

【図11】関係性値の算出方法について説明するための説明図である。

【図12】関係性値の算出方法について説明するための説明図である。

【図13】関係性情報について説明するための説明図である。

【図14】本実施形態に係る情報処理装置の構成例について説明するための説明図である。

【図15】本実施形態に係る情報処理装置の構成例（変形例#1）について説明するための説明図である。

【図16】本実施形態に係る情報処理装置の構成例（変形例#2）について説明するための説明図である。

40

【図17】本実施形態に係る情報処理装置の動作例について説明するための説明図である。

【図18】本実施形態に係る情報処理装置の動作例について説明するための説明図である。

【図19】本実施形態に係る情報処理装置の動作例について説明するための説明図である。

【図20】本実施形態に係る情報処理装置の動作例について説明するための説明図である。

【図21】本実施形態に係る情報処理装置の動作例について説明するための説明図である。

50

【図 2 2】本実施形態に係る情報処理装置の動作例について説明するための説明図である。

【図 2 3】本実施形態に係る情報処理装置の動作例について説明するための説明図である。

【図 2 4】本実施形態に係る情報処理装置の動作例について説明するための説明図である。

【図 2 5】本実施形態に係る情報処理装置の動作例について説明するための説明図である。

【図 2 6】本実施形態に係る情報処理装置の動作例について説明するための説明図である。

【図 2 7】本実施形態に係る情報処理装置の動作例について説明するための説明図である。

【図 2 8】本実施形態に係る関係性値の編集方法について説明するための説明図である。

【図 2 9】本実施形態に係る関係性値の編集方法について説明するための説明図である。

【図 3 0】本実施形態に係る関係性値の編集方法について説明するための説明図である。

【図 3 1】本実施形態の一変形例に係る動画要約手法について説明するための説明図である。

【図 3 2】本実施形態の一変形例に係る動画要約手法について説明するための説明図である。

【図 3 3】本実施形態の一変形例に係る動画要約手法について説明するための説明図である。

【図 3 4】本実施形態に係る情報処理装置の機能を実現することが可能なハードウェア構成例について説明するための説明図である。

【図 3 5】顔トラッキング、顔クラスタリング、及び顔識別の内容を示した参考図である。

【発明を実施するための形態】

【0011】

以下に添付図面を参照しながら、本技術に係る好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。

【0012】

[説明の流れについて]

ここで、以下に記載する説明の流れについて簡単に述べる。

【0013】

まず、図 1 ~ 図 4 を参照しながら、本実施形態に係る動画要約手法の概要について説明する。次いで、図 5 ~ 図 1 3 を参照しながら、関係性値の算出方法について説明する。次いで、図 1 4 ~ 図 1 6 を参照しながら、本実施形態に係る情報処理装置 1 0 0 の構成例について説明する。次いで、図 1 7 ~ 図 2 7 を参照しながら、本実施形態に係る情報処理装置 1 0 0 の動作例について説明する。

【0014】

次いで、図 2 8 ~ 図 3 0 を参照しながら、本実施形態に係る関係性値の編集方法について説明する。次いで、図 3 1 ~ 図 3 3 を参照しながら、本実施形態の一変形例に係る動画要約手法について説明する。次いで、図 3 4 を参照しながら、本実施形態に係る情報処理装置 1 0 0 の機能を実現することが可能なハードウェア構成例について説明する。なお、図 3 5 は、説明の中で適宜参照する。

【0015】

最後に、同実施形態の技術的思想について纏め、当該技術的思想から得られる作用効果について簡単に説明する。

【0016】

(説明項目)

10

20

30

40

50

## 1 : はじめに

## 1 - 1 : 登場物間の関係性に基づく動画要約手法の概要

## 1 - 1 - 1 : ダイジェスト再生について

## 1 - 1 - 2 : コミック表示について

## 1 - 2 : 関係性値の算出方法

## 1 - 2 - 1 : 動画タイムラインメタデータについて

## 1 - 2 - 2 : 区間メタデータに基づく共演関係の評価方法

## 1 - 2 - 3 : 共演作品数に基づく関係性値の算出方法

## 1 - 2 - 4 : 共演作品数に基づく関係性値の算出方法 (重み付き)

## 1 - 2 - 5 : 出演時間に基づく関係性値の算出方法

## 1 - 2 - 6 : 出演シーン時間に基づく関係性値の算出方法

## 1 - 3 : 関係性情報の拡張

## 2 : 実施形態

## 2 - 1 : 情報処理装置 100 の構成

## 2 - 1 - 1 : 標準構成

## 2 - 1 - 2 : 変形例 # 1 (動画からメタデータを自動生成する構成)

## 2 - 1 - 3 : 変形例 # 2 (所定の関係性情報を利用する構成)

## 2 - 2 : 情報処理装置 100 の動作

## 2 - 2 - 1 : 基本動作

## 2 - 2 - 2 : 要約処理 # 1 (関係性値の総和を最大化する方法)

## 2 - 2 - 3 : 要約処理 # 2 (シーン切り替え時の関係性値和を最大化 1)

## 2 - 2 - 4 : 要約処理 # 3 (シーン切り替え時の関係性値和を最大化 2)

## 2 - 2 - 5 : 要約処理 # 4 (シーン切り替え時の関係性値差を最小化)

## 2 - 2 - 6 : 要約処理 # 5 (その他の構成)

## 2 - 3 : 関係性値の編集方法

## 2 - 3 - 1 : 操作 UI の構成

## 2 - 3 - 2 : 関係性値の編集処理に伴う情報処理装置 100 の動作

## 2 - 4 : 変形例 (テンプレートを利用した動画要約手法)

## 2 - 4 - 1 : テンプレートの構成

## 2 - 4 - 2 : 要約処理の流れ

## 2 - 4 - 3 : テンプレートの自動生成方法

## 3 : ハードウェア構成例

## 4 : まとめ

## 【0017】

## &lt; 1 : はじめに &gt;

はじめに、本実施形態に係る動画要約手法について詳細に説明するに先立ち、当該動画要約手法の概要及び当該動画要約手法に用いる関係性値の算出方法について説明する。

## 【0018】

## [ 1 - 1 : 登場物間の関係性に基づく動画要約手法の概要 ]

まず、図 1 ~ 図 4 を参照しながら、登場物間の関係性に基づく動画要約手法の概要について説明する。図 1 及び図 2 は、登場物の関係性を利用したダイジェスト再生の仕組みについて説明するための説明図である。また、図 3 及び図 4 は、登場物の関係性を利用した代表シーン画像のコミック表示方法について説明するための説明図である。

## 【0019】

## ( 1 - 1 - 1 : ダイジェスト再生 )

近年、一般家庭においても動画を編集する機会が増えている。例えば、子供の運動会を撮影した映像の中から、自分の子供や自分の子供に関係する人物や物品など (以下、登場物) が写っている映像シーンだけを切り出して繋ぎ合わせる編集作業は、多くの場合、撮影者自身の手で行われる。但し、こうした編集作業は、撮影者にとって負荷の高い作業である。また、子供に関係する登場物を撮影者が十分に把握しているとは限らず、本当に適

10

20

30

40

50

切な編集結果が得られていない可能性もある。そのため、登場物間の関係性を考慮して自動的に適切な映像シーンを抽出する技術が求められている。

#### 【0020】

また、登場物間の関係性を考慮して自動的に適切な映像シーンが抽出されるのであれば、実際には動画を編集せずに、自動抽出した映像シーンだけを時系列に再生することで、視聴者の意図に沿った動画の視聴形態を実現することが可能になる。例えば、自分の子供を指定し、自分の子供と関係する登場物の登場シーンだけを自動抽出して再生すれば、編集した動画を再生するのと同じ結果が得られる。また、ユーザが撮影した動画だけでなく、市場に流通している映画などの動画を対象に、ユーザが指定した登場物に關係する登場物の登場シーンを選択的に再生することも可能になる。

10

#### 【0021】

以下、動画から特定の映像シーンを選択的に再生する再生方法のことをダイジェスト再生と呼ぶことにする。言うまでもないが、動画のダイジェスト再生は、動画の要約に他ならない。さて、本実施形態において、動画のダイジェスト再生は、図1及び図2のような仕組みを用いて実現される。例えば、ある動画に人物A、人物B、人物Cが登場するものとしよう。つまり、人物A、人物B、人物Cは、この動画における登場物である。また、人物Aと人物Bとは良好な関係性を有するものとする。一方、人物Aと人物Cとは陰険な関係性を有するものとする。なお、登場物間の関係性を定量化する方法については後段において詳述する。

#### 【0022】

まず、動画中における人物A、人物B、人物Cの登場シーンが抽出される。また、ユーザが人物Aを選択し、良好な関係性に基づくダイジェスト再生を望んだ場合、人物Aと良好な関係性を有する人物Bが選択される。次いで、人物Aの登場シーンと、人物Bの登場シーンとが時系列に並べられてダイジェスト動画が生成される。但し、ダイジェスト動画は、実際に動画を編集して得られる動画データであってもよいし、再生区間を示した情報であってもよい。図1の例では、人物Aが登場する区間 $S_{11} \sim S_{14}$ と、人物Bが登場する区間 $S_{21}$ 及び $S_{22}$ とが時系列に並べられたダイジェスト動画が得られる。

20

#### 【0023】

また、ユーザが異なる登場物を選択すると、図2に示すように、ダイジェスト動画の内容も変更される。図2の例では、動画中に人物A、人物B、人物C、人物Dが登場するものと仮定されている。また、人物Aと人物Bとは良好な関係性を有し、人物Aと人物Cとは陰険な関係性を有する。そして、人物Aと人物Dとは陰険な関係性を有し、人物Bと人物Dとは良好な関係性を有する。さらに、人物Cと人物Dとは良好な関係性を有する。このような関係性が存在する場合、ユーザが人物Aを選択し、良好な関係性に基づくダイジェスト再生を望むと、ダイジェスト動画は、(1)のような構成になる。一方、ユーザが人物Dを選択すると、ダイジェスト動画は、(2)のような構成になる。

30

#### 【0024】

上記のように、本実施形態に係る動画要約手法を用いると、同じ動画であっても、選択する登場物を変更することにより、ダイジェスト再生される動画の内容も変更される。また、上記の例では良好な関係性に基づくダイジェスト再生について考えたが、関係性の種類が途中で切り替わるように設定してもよい。例えば、「味方の関係 敵対関係 無関係 敵対関係 味方の関係」といった設定を設けておくことにより、特定のストーリー性をダイジェスト動画に与えることが可能になる。但し、このような関係性に関する設定を設けた場合であっても、ユーザが選択する登場物の変更に従ってダイジェスト動画の内容が変更されることに変わりはない。

40

#### 【0025】

(1-1-2: コミック表示)

図1及び図2に示したダイジェスト再生は、要約した動画を時系列で視聴するために利用される方法に関する。一方、ここで説明するコミック表示は、動画の要約結果を俯瞰的に視聴するための利用される方法に関する。但し、映像シーンの抽出方法は、上述したダ

50



イジェスト再生における映像シーンの抽出方法と同じである。例えば、図2の例において人物Aが選択された場合に、人物Aが登場する区間 $S_{11} \sim S_{14}$ と、人物Bが登場する区間 $S_{21}$ 及び $S_{22}$ とが抽出される。しかし、コミック表示の場合、区間 $S_{11} \sim S_{14}$ 、 $S_{21}$ 、 $S_{22}$ が時系列に再生されるのではなく、図3に示すように、各区間を代表する画像が並べて表示される。なお、代表する画像に代えて、各区間の動画が再生されるように構成されていてもよい。

#### 【0026】

また、図2の例において人物Dが選択された場合、人物Dが登場する区間 $S_{41} \sim S_{43}$ と、人物Dの登場シーン以降に人物Bが登場する区間 $S_{22}$ と、人物Dの登場シーン以降に人物Cが登場する区間 $S_{32} \sim S_{34}$ とが抽出される。この場合、図4に示すように、各区間を代表する画像が並べて表示される。なお、1つの枠と1つの区間とが1対1に対応していなくてもよい。例えば、所定時間毎に各枠内に表示された画像が他の区間に対応する画像に切り替わるようにしてもよい。また、ユーザが各枠の画像を選択した場合に、選択された画像に対応する区間の動画が再生されるようにしてもよい。さらに、ユーザにより選択された登場物と関係性の強い登場物の登場シーンに対応する枠が強調表示されるように構成されていてもよい。

#### 【0027】

以上説明したように、本実施形態に係る動画要約手法を用いると、上記のようなダイジェスト再生やコミック表示が可能になる。なお、要約結果の表現方法はこれらの方法に限定されない点に注意されたい。

#### 【0028】

##### [1-2: 関係性値の算出方法]

次に、図5～図12を参照しながら、関係性値の算出方法について説明する。この関係性値は、登場物間の関係性を定量的に表現したものである。この関係性値は、例えば、以下で説明する動画タイムラインメタデータを利用して算出される。

#### 【0029】

##### (1-2-1: 動画タイムラインメタデータについて)

まず、動画タイムラインメタデータについて説明する。動画タイムラインメタデータは、図5に示すように、領域メタデータ、区間メタデータ、オブジェクトメタデータにより構成される。領域メタデータは、動画フレーム内に登場する人物や物品など（登場物）の位置及び範囲を示すメタデータである。図5の例では、動画フレーム内に登場する人物の顔領域を示す領域メタデータが例示されている。また、区間メタデータは、動画の中で、登場物が登場する区間を示すメタデータである。そして、オブジェクトメタデータは、動画に登場する登場物に関する関連情報を示すメタデータである。

#### 【0030】

なお、動画タイムラインメタデータは、動画に登場する登場物毎に設定される。また、動画タイムラインメタデータは、動画1つ1つについて設定される。そのため、動画タイムラインメタデータのうち、区間メタデータを利用することにより、ある登場物がどの動画のどの区間に登場するかを把握することが可能になる。また、領域メタデータを利用することにより、ユーザが指定した画面上の領域と、画面上に表示されている登場物との対応関係を把握することが可能になる。例えば、領域メタデータ及びオブジェクトメタデータを利用すれば、ユーザが登場人物の顔領域を指定した場合に、その登場人物に関する関連情報が表示されるといったユーザインターフェースを実現することが可能になる。

#### 【0031】

なお、動画タイムラインメタデータは、手入力により生成してもよいが、図35に示すような顔検出、顔トラッキング、顔クラスタリング、顔識別などの方法を用いて自動生成することもできる。また、手入力による生成方法と自動生成方法とを補完的に組み合わせることにより、高精度の動画タイムラインメタデータを生成することができる。例えば、特開2005-44330号公報（物体検出・トラッキング）、特開2010-3021号公報（クラスタリング）、特開2007-65766号公報（認識）などに記載の技術

10

20

30

40

50

を利用することが可能である。

#### 【0032】

(1-2-2: 区間メタデータに基づく共演関係の評価方法)

次に、図6～図8を参照しながら、区間メタデータに基づく共演関係の評価方法について説明する。なお、ここで言う共演とは、同じ動画又は同じ動画フレームに複数の登場物が登場する状態のことを意味する。また、以下では、説明の都合上、登場物が人物である場合を例に挙げて説明を進める。まず、区間メタデータから得られる主な情報の種類と、人物間の関係性を評価する際の評価基準とについて説明する。図6～図8は、区間メタデータに基づく共演関係の評価方法について説明するための説明図である。

#### 【0033】

まず、図6を参照する。先に述べたように、区間メタデータは、動画毎及び人物毎に用意されている。また、各区間メタデータは、人物が登場する動画中の区間を表す。図6には、動画 $M_1$ について、人物Aの登場区間を表した区間メタデータと、人物Bの登場区間を表した区間メタデータと、人物Cの登場区間を表した区間メタデータとが例示されている。図6の例では、 $t_2 \sim t_4$ の区間、 $t_9 \sim t_{11}$ の区間、 $t_{13} \sim t_{16}$ の区間が人物Aの登場区間である。また、 $t_1 \sim t_3$ の区間、 $t_5 \sim t_7$ の区間、 $t_{10} \sim t_{12}$ の区間が人物Bの登場区間である。さらに、 $t_6 \sim t_8$ の区間、 $t_{14} \sim t_{15}$ の区間が人物Cの登場区間である。

#### 【0034】

図6に例示した3つの区間メタデータからは、動画 $M_1$ における各登場人物の出演時間(動画別出演時間)、動画 $M_1$ における共演者、動画 $M_1$ における共演シーンの長さ(以下、共演シーン時間)が分かる。例えば、動画 $M_1$ における人物Aの動画別出演時間 $t_A$ は、 $t_A = |t_4 - t_2| + |t_{11} - t_9| + |t_{16} - t_{13}|$ と算出される。人物Bの動画別出演時間 $t_B$ 、人物Cの動画別出演時間 $t_C$ についても同様である。また、動画 $M_1$ において人物A、B、Cの登場区間が存在することから、人物Aと人物Bとは共演者であり、人物Bと人物Cとは共演者であり、人物Cと人物Aとは共演者であることが分かる。

#### 【0035】

さらに、 $t_2 \sim t_3$ の区間及び $t_{10} \sim t_{11}$ の区間で人物Aと人物Bとが共に登場しているから、動画 $M_1$ における人物Aと人物Bとの共演シーン時間 $t_{AB}$ は、 $t_{AB} = |t_3 - t_2| + |t_{11} - t_{10}|$ と算出される。同様に、 $t_6 \sim t_7$ の区間で人物Bと人物Cとが共に登場しているから、動画 $M_1$ における人物Bと人物Cとの共演シーン時間 $t_{BC}$ は、 $t_{BC} = |t_7 - t_6|$ と算出される。また、 $t_{15} \sim t_{14}$ の区間で人物Aと人物Cとが共に登場しているから、動画 $M_1$ における人物Aと人物Cとの共演シーン時間 $t_{CA}$ は、 $t_{CA} = |t_{15} - t_{14}|$ と算出される。

#### 【0036】

このように、各動画の区間メタデータを分析することにより、各動画における各登場人物の動画別出演時間、各動画における共演者、各動画における共演シーン時間が得られる。また、複数の動画について得られた動画別出演時間、共演者、共演シーン時間などの情報を組み合わせると、図7に示すように、対象とする動画群(図7の例では動画 $M_1 \sim M_n$ )について各人物の総出演時間、各人物の組に関する共演作品数、各人物の組に関する総共演時間、各人物の組に関する総共演シーン時間といった情報が得られる。

#### 【0037】

例えば、対象とする動画群について同じ人物の動画別出演時間を合計すると、その人物に関する総出演時間が得られる。また、各動画における共演者が分かっているため、ある人物の組が共演者となっている動画の数を計算すれば、その人物の組に関する共演作品数が得られる。さらに、ある人物の組に注目し、その人物の組が共演する動画群について各人物の動画別出演時間を合計することで、その人物の組に関する共演時間が得られる。また、ある人物の組に注目し、その人物の組が共演する動画群について共演シーン時間を合計することで、その人物の組に関する共演シーン時間が得られる。

## 【0038】

例えば、人物Dと人物Eとが動画M<sub>2</sub>、M<sub>5</sub>で共演しているとしよう。この場合、人物Dと人物Eとの組に関する共演時間は、動画M<sub>2</sub>及びM<sub>5</sub>における人物Dの動画別出演時間( $t_{2D}$ 及び $t_{5D}$ )の合計値と、動画M<sub>2</sub>及びM<sub>5</sub>における人物Eの動画別出演時間( $t_{2E}$ 及び $t_{5E}$ )の合計値との和( $t_{2D} + t_{5D} + t_{2E} + t_{5E}$ )となる。また、人物Dと人物Eとの組に関する共演シーン時間は、動画M<sub>2</sub>における共演シーン時間 $t_{2DE}$ と、動画M<sub>5</sub>における共演シーン時間 $t_{5DE}$ との和( $t_{2DE} + t_{5DE}$ )となる。

## 【0039】

共演作品数が多い人物の組、共演時間の長い人物の組、共演シーン時間の長い人物の組は、人物間の関係性が強い組であると考えられる。また、お互いに主役として共演している作品が多いほど、人物間の関係性が強いと考えることもできる。さらに、表示される面積なども考慮して人物間の関係性を評価する方が好ましいと考えられる。表示面積を考慮しないと、例えば、エキストラとして出演している人物と主演俳優との間の関係性が高く、主演俳優と主演女優との間の関係性が低くなるといったことも起こりうる。主役や脇役などの情報は、オブジェクトメタデータから得られる。また、子供、大人、男性、女性、年齢などの情報がオブジェクトデータに含まれている場合には、これらの情報を利用してよい。また、表示面積などは、領域メタデータから得られる。

## 【0040】

上記のように、区間メタデータを利用することで、人物間の関係性を評価するための様々な情報が得られる。また、その情報を利用することにより、人物間の関係性を評価することができる。さらに、オブジェクトメタデータや領域メタデータを併せて利用することにより、人物間の関係性をより適切に評価することが可能になる。例えば、人物間の関係性は、図8に示すように、マトリックスの形式で表現できる。なお、マトリックスの各枠には、その要素に対応する人物の組について関係性の強さを表す関係性値が入る。なお、このマトリックスのことを関係性マトリックスと呼ぶことにする。また、人物以外の登場物についても、同様に関係性マトリックスを求めることができる。以下、関係性値の具体的な算出方法について説明する。

## 【0041】

(1 - 2 - 3 : 共演作品数に基づく関係性値の算出方法)

まず、図9を参照しながら、共演作品数に基づく関係性値の算出方法について説明する。図9は、共演作品数に基づく関係性値の算出方法について説明するための説明図である。なお、対象とする動画群は、動画M<sub>1</sub> ~ M<sub>6</sub>であるとする。また、動画M<sub>1</sub> ~ M<sub>6</sub>に登場する人物は、人物A、人物B、人物Cであるとする。

## 【0042】

既に述べたように、区間メタデータを利用すると、図6に示すように、人物A、人物B、人物Cの共演関係が得られる。図9の例において、動画M<sub>1</sub>の出演者は、人物A、人物B、人物Cである。また、動画M<sub>2</sub>の出演者は、人物A、人物Cである。さらに、動画M<sub>3</sub>の出演者は、人物Bだけである。そして、動画M<sub>4</sub>の出演者は、人物Aだけである。また、動画M<sub>5</sub>の出演者は、人物A、人物Bである。さらに、動画M<sub>6</sub>の出演者は、人物A、人物Bである。つまり、人物Aと人物Bとが共演する作品は、動画M<sub>1</sub>、動画M<sub>5</sub>、動画M<sub>6</sub>である。また、人物Aと人物Cとが共演する作品は、動画M<sub>1</sub>、動画M<sub>2</sub>である。さらに、人物Bと人物Cとが共演する作品は、動画M<sub>1</sub>だけである。

## 【0043】

人物Aと人物Bとが共演する作品数は3、人物Aと人物Cとが共演する作品数は2、人物Bと人物Cとが共演する作品数は1であるため、人物Aと人物Bとの関係性を表す関係性値は3、人物Aと人物Cとの関係性を表す関係性値は2、人物Bと人物Cとの関係性を表す関係性値は1となる。なお、同じ人物間の関係性値も形式的に算出することができる。例えば、人物Aと人物Aとが共演する作品は、人物Aが出演する作品と同義であり、動画M<sub>1</sub>、M<sub>2</sub>、M<sub>4</sub> ~ M<sub>6</sub>が該当する。つまり、人物Aと人物Aとの共演作品数は5とな

10

20

30

40

50

り、人物 A と人物 A との関係性を表す関係性値は 5 となる。人物 B、人物 C についても同様である。

#### 【 0 0 4 4 】

同じ人物に関する関係性値も含め、上記の方法で算出された関係性値をまとめると、図 9 に示すような関係性マトリックスが得られる。但し、この関係性マトリックスの対角成分は、各枠に対応する人物の出演作品数を表す。なお、ここでは出演作品数をそのまま関係性値として記載しているが、所定の係数をかけたり、正規化したりするなど、加工した数値を関係性値として利用することが望ましい。例えば、対角成分が全て 1 になるように正規化することが望ましい。

#### 【 0 0 4 5 】

以上、共演作品数に基づく関係性の算出方法について説明した。

#### 【 0 0 4 6 】

( 1 - 2 - 4 : 共演作品数に基づく関係性値の算出方法 ( 重み付き ) )

次に、図 10 を参照しながら、共演作品数に基づく関係性値の算出方法 ( 重み付き ) について説明する。図 10 は、共演作品数に基づく関係性値の算出方法 ( 重み付き ) について説明するための説明図である。ここでは、主役 / 脇役の区別を重みで表現し、その重み値と共演作品数とに基づいて関係性値を算出する方法について述べる。なお、子供、大人、男性、女性、年齢などの区別を重みで表現した場合も、同様の方法にて関係性マトリックスを求めることができる。ここでは、対象とする動画群は、動画  $M_1 \sim M_6$  であるとする。また、動画  $M_1 \sim M_6$  に登場する人物は、人物 A、人物 B、人物 C であるとする。

#### 【 0 0 4 7 】

区間メタデータを利用すると、図 10 に示すように、人物 A、人物 B、人物 C の共演関係が得られる。図 10 の例において、動画  $M_1$  の出演者は、人物 A、人物 B、人物 C である。また、動画  $M_2$  の出演者は、人物 A、人物 C である。さらに、動画  $M_3$  の出演者は、人物 B だけである。そして、動画  $M_4$  の出演者は、人物 A だけである。また、動画  $M_5$  の出演者は、人物 A、人物 B である。さらに、動画  $M_6$  の出演者は、人物 A、人物 B である。つまり、人物 A と人物 B とが共演する作品は、動画  $M_1$ 、動画  $M_5$ 、動画  $M_6$  である。また、人物 A と人物 C とが共演する作品は、動画  $M_1$ 、動画  $M_2$  である。さらに、人物 B と人物 C とが共演する作品は、動画  $M_1$  だけである。

#### 【 0 0 4 8 】

また、オブジェクトメタデータを利用すると、動画  $M_1$  において人物 A 及び人物 B が主役、人物 C が脇役であることが分かる。同様に、動画  $M_2$  において人物 A が主役、人物 C が脇役であることが分かる。また、動画  $M_3$  において人物 B が主役であることが分かる。さらに、動画  $M_4$  において人物 A が脇役であることが分かる。そして、動画  $M_5$  において人物 A、人物 B が脇役であることが分かる。また、動画  $M_6$  において人物 A が主役、人物 B が脇役であることが分かる。

#### 【 0 0 4 9 】

ここで、重みについて考える。共演者が共に主役の場合、共演作品 1 つにつきスコア = 4 を与えるものとする。また、共演者の一方が主役、かつ、他方が脇役の場合、共演作品 1 つにつきスコア = 2 を与えるものとする。さらに、共演者が共に脇役の場合、共演作品 1 つにつきスコア = 1 を与えるものとする。

#### 【 0 0 5 0 】

図 10 の例において、人物 A と人物 B とが共に主役を務める作品は、動画  $M_1$  である。また、人物 A が主役を務め、かつ、人物 B が脇役を務める作品、或いは、人物 A が脇役を務め、かつ、人物 B が主役を務める作品は、動画  $M_6$  である。そして、人物 A と人物 B とが共に脇役を務める作品は、動画  $M_5$  である。これらの結果から、人物 A と人物 B との組について、人物 A と人物 B とが共に主役を務める作品の数は 1 である。また、人物 A が主役を務め、かつ、人物 B が脇役を務める作品、或いは、人物 A が脇役を務め、かつ、人物 B が主役を務める作品の数は 1 である。そして、人物 A と人物 B とが共に脇役を務める作品の数は 1 である。従って、スコアの合計は、 $4 \times 1 + 2 \times 1 + 1 \times 1 = 7$  となる。つま

10

20

30

40

50

り、人物 A と人物 B との関係性を表す関係性値は 7 となる。

【 0 0 5 1 】

同様に、人物 A と人物 C とが共に主役を務める作品は、動画  $M_2$  である。また、人物 A が主役を務め、かつ、人物 C が脇役を務める作品、或いは、人物 A が脇役を務め、かつ、人物 C が主役を務める作品は、動画  $M_1$  である。そして、人物 A と人物 C とが共に脇役を務める作品は存在しない。これらの結果から、人物 A と人物 C との組について、人物 A と人物 C とが共に主役を務める作品の数は 1 である。また、人物 A が主役を務め、かつ、人物 C が脇役を務める作品、或いは、人物 A が脇役を務め、かつ、人物 C が主役を務める作品の数は 1 である。そして、人物 A と人物 C とが共に脇役を務める作品の数は 0 である。従って、スコアの合計は、 $4 \times 1 + 2 \times 1 + 1 \times 0 = 6$  となる。つまり、人物 A と人物 C との関係性を表す関係性値は 6 となる。

10

【 0 0 5 2 】

同様に、人物 B と人物 C とが共に主役を務める作品は存在しない。また、人物 B が主役を務め、かつ、人物 C が脇役を務める作品、或いは、人物 B が脇役を務め、かつ、人物 C が主役を務める作品は、動画  $M_1$  である。そして、人物 B と人物 C とが共に脇役を務める作品は存在しない。これらの結果から、人物 B と人物 C との組について、人物 B と人物 C とが共に主役を務める作品の数は 0 である。また、人物 B が主役を務め、かつ、人物 C が脇役を務める作品、或いは、人物 B が脇役を務め、かつ、人物 C が主役を務める作品の数は 1 である。そして、人物 B と人物 C とが共に脇役を務める作品の数は 0 である。従って、スコアの合計は、 $4 \times 0 + 2 \times 1 + 1 \times 0 = 2$  となる。つまり、人物 B と人物 C との関係性を表す関係性値は 2 となる。

20

【 0 0 5 3 】

同じ人物に関する関係性値も同様にして求め、これらの関係性値をまとめると、図 10 に示すような関係性マトリックスが得られる。なお、ここではスコアの合計値をそのまま関係性値としたが、合計値の平方根を関係性値としてもよい。また、所定の係数をかけたり、正規化したりするなど、加工した数値を関係性値として利用することが望ましい。例えば、対角成分が全て 1 になるように正規化することが望ましい。また、上記の説明では、役柄の組み合わせに応じてスコア付けする方法を例示したが、例えば、次のように役柄重み  $RW$  を定義し、下記の式 (1) に基づいて関係性値  $Rel$  を算出してもよい。役柄重み  $RW(M_k, A)$  は、動画  $M_k$  における人物 A の役柄が主役の場合に 2、脇役の場合に 1、出演していない場合に 0 をとる。また、 $Rel(A, B)$  は、人物 A と人物 B との関係性を示す関係性値を表す。

30

【 0 0 5 4 】

【 数 1 】

$$Rel(A, B) = \sqrt{\sum_k \{RW(M_k, A) \cdot RW(M_k, B)\}}$$

… (1)

40

【 0 0 5 5 】

以上、共演作品数に基づく関係性の算出方法 (重み付き) について説明した。

【 0 0 5 6 】

( 1 - 2 - 5 : 出演時間に基づく関係性値の算出方法 )

次に、図 11 を参照しながら、出演時間に基づく関係性値の算出方法について説明する。図 11 は、出演時間に基づく関係性値の算出方法について説明するための説明図である。ここでは、各動画における各人物の出演時間を利用して関係性値を算出する方法について述べる。なお、対象とする動画群は、動画  $M_1 \sim M_6$  であるとする。また、動画  $M_1 \sim M_6$  に登場する人物は、人物 A、人物 B、人物 C であるとする。

50

## 【 0 0 5 7 】

区間メタデータを利用すると、図 1 1 に示すように、人物 A、人物 B、人物 C の共演関係が得られる。図 1 1 の例において、動画 M<sub>1</sub> の出演者は、人物 A、人物 B、人物 C である。また、動画 M<sub>2</sub> の出演者は、人物 A、人物 C である。さらに、動画 M<sub>3</sub> の出演者は、人物 B だけである。そして、動画 M<sub>4</sub> の出演者は、人物 A だけである。また、動画 M<sub>5</sub> の出演者は、人物 A、人物 B である。さらに、動画 M<sub>6</sub> の出演者は、人物 A、人物 B である。つまり、人物 A と人物 B とが共演する作品は、動画 M<sub>1</sub>、動画 M<sub>5</sub>、動画 M<sub>6</sub> である。また、人物 A と人物 C とが共演する作品は、動画 M<sub>1</sub>、動画 M<sub>2</sub> である。さらに、人物 B と人物 C とが共演する作品は、動画 M<sub>1</sub> だけである。

## 【 0 0 5 8 】

また、区間メタデータを利用すると、図 1 1 に示すように、動画 M<sub>1</sub> における人物 A の出演時間は 4 0、人物 B の出演時間は 3 0、人物 C の出演時間は 1 0 であることが分かる。同様に、動画 M<sub>2</sub> における人物 A の出演時間は 3 0、人物 C の出演時間は 4 0 であることが分かる。また、動画 M<sub>3</sub> における人物 B の出演時間は 2 0 であることが分かる。さらに、動画 M<sub>4</sub> における人物 A の出演時間は 1 0 であることが分かる。そして、動画 M<sub>5</sub> における人物 A の出演時間は 5、人物 B の出演時間は 1 0 であることが分かる。また、動画 M<sub>6</sub> における人物 A の出演時間は 4 0、人物 B の出演時間は 5 であることが分かる。

## 【 0 0 5 9 】

出演時間が長い人物は、その動画の中で重要な役を担当していると言える。また、互いに重要な役で出演している人物間の関係性は強いと考えられる。逆に、互いに出演時間が短い人物間の関係性は低いと考えられる。例えば、ちょい役で登場する人物間の関係性は低く評価されるべきであろう。こうした考えから、出演時間 P S L を定義し、下記の式 ( 2 ) に基づいて関係性値 R e l を算出する方法を提案する。但し、出演時間 P S L ( M<sub>k</sub> , A ) は、動画 M<sub>k</sub> における人物 A の出演時間を表し、人物 A が出演していない場合には 0 をとる。また、R e l ( A , B ) は、人物 A と人物 B との関係性を示す関係性値を表す。なお、右辺の平方根を関係性値としてもよい。

## 【 0 0 6 0 】

## 【 数 2 】

$$Rel(A, B) = \sum_k \{PSL(M_k, A) \cdot PSL(M_k, B)\}$$

… ( 2 )

## 【 0 0 6 1 】

図 1 1 の例において、人物 A と人物 B との関係性を示す関係性値は、4 0 × 3 0 ( 動画 M<sub>1</sub> ) + 3 0 × 0 ( 動画 M<sub>2</sub> ) + 0 × 2 0 ( 動画 M<sub>3</sub> ) + 1 0 × 0 ( 動画 M<sub>4</sub> ) + 5 × 1 0 ( 動画 M<sub>5</sub> ) + 4 0 × 5 ( 動画 M<sub>6</sub> ) = 1 4 5 0 となる。同様に計算すると、人物 A と人物 C との関係性を示す関係性値は 1 6 0 0、人物 B と人物 C との関係性を示す関係性値は 3 0 0 となる。これらの数値をまとめると、図 1 1 に示すような関係性マトリックスが得られる。なお、ここではスコアをそのまま関係性値として記載しているが、所定の係数をかけたり、正規化したりするなど、加工した数値を関係性値として利用することが望ましい。例えば、対角成分が全て 1 になるように正規化することが望ましい。

## 【 0 0 6 2 】

以上、出演時間に基づく関係性マトリックスの計算方法について説明した。

## 【 0 0 6 3 】

( 1 - 2 - 6 : 出演シーン時間に基づく関係性値の算出方法 )

次に、図 1 2 を参照しながら、出演シーン時間に基づく関係性値の算出方法について説明する。図 1 2 は、出演シーン時間に基づく関係性値の算出方法について説明するための説明図である。ここでは、各動画における共演シーン時間を利用して関係性値を算出する

10

20

30

40

50

方法について述べる。なお、対象とする動画群は、動画  $M_1 \sim M_6$  であるとする。また、動画  $M_1 \sim M_6$  に登場する人物は、人物 A、人物 B、人物 C であるとする。

【0064】

区間メタデータを利用すると、図 12 に示すように、各動画について各人物の組に関する共演シーン時間が得られる。なお、図 12 の中では、人物 X と人物 Y との組に関する共演シーン時間を  $[X, Y]$  と表現している。例えば、動画  $M_1$  における人物 A と人物 B との組に関する共演シーン時間は  $[A, B] = 20$  である。同様に、動画  $M_1$  に関しては、 $[A, A] = 40$ 、 $[B, B] = 30$ 、 $[C, C] = 10$ 、 $[A, C] = 5$ 、 $[B, C] = 5$  という結果が得られている。なお、 $[A, A]$  とは、人物 A と人物 A とが共演する区間の長さを示すものであるが、同じ人物が対象であるから、人物 A の出演時間に一致する。 $[B, B]$ 、 $[C, C]$  についても同様である。

10

【0065】

同じ映像シーンに登場する人物間の関係性は強いと考えられる。例えば、会話を交わす関係にある人物は、当然のことながら同じ映像シーンに登場する。また、敵対していたり、或いは、味方の関係にある人物は、動画の中で同じ映像シーンに登場する頻度が高い。ヒーロー役の人物とヒロイン役の人物も同じ映像シーンに登場する頻度が高い。こうした考えから、共演シーン時間  $CSL$  を定義し、下記の式 (3) に基づいて関係性値  $Rel$  を算出する方法を提案する。但し、共演シーン時間  $CSL(M_k, A, B)$  は、動画  $M_k$  において人物 A と人物 B とが共演している区間の長さを表し、いずれかの人物が出演していない場合には 0 をとる。また、 $Rel(A, B)$  は、人物 A と人物 B との関係性を示す関係性値を表す。なお、右辺の平方根を関係性値としてもよい。

20

【0066】

【数 3】

$$Rel(A, B) = \sum_k CSL(M_k, A, B)$$

… (3)

【0067】

30

図 12 の例において、人物 A と人物 B との関係性を示す関係性値は、 $20$  (動画  $M_1$ ) +  $0$  (動画  $M_2$ ) +  $0$  (動画  $M_3$ ) +  $0$  (動画  $M_4$ ) +  $5$  (動画  $M_5$ ) +  $5$  (動画  $M_6$ ) =  $30$  となる。同様に計算すると、人物 A と人物 C との関係性を示す関係性値は  $25$ 、人物 B と人物 C との関係性を示す関係性値は  $5$  となる。これらの数値をまとめると、図 12 に示すような関係性マトリックスが得られる。なお、ここでは共演シーン時間の合計値をそのまま関係性値として記載しているが、所定の係数をかけたり、正規化したりするなど、加工した数値を関係性値として利用することが望ましい。例えば、対角成分が全て 1 になるように正規化することが望ましい。

【0068】

以上、共演シーン時間に基づく関係性マトリックスの計算方法について説明した。

40

【0069】

以上説明したように、区間メタデータを利用して得られる情報から、様々な観点で人物間の関係性を評価することが可能である。なお、ここで説明した関係性値の算出方法は一例であり、例えば、顔領域の面積により重み付けした出演時間や共演シーン時間などを利用して関係性値を算出することも可能である。また、主役 / 脇役の区別、或いは、子供、大人、男性、女性、年齢などの区別を考慮して重み付けした出演時間や共演シーン時間などを利用して関係性値を算出することも可能である。

【0070】

(その他の方法 1 : 共演シーン時間 + 役柄重み)

一例として、共演シーン時間  $CSL$  及び役柄重み  $RW$  を組み合わせて関係性値  $Rel$  を

50

算出する方法を紹介する。なお、CSL及びRWの定義は既に説明したものと同一である。この例において、関係性値Rel(A, B)は、下記の式(4)により算出される。なお、右辺の平方根は省略してもよい。この方法によると、各動画における各人物の重要度を示す役柄重みと、各動画における人物間の関係性の強さを示す共演シーン時間とを共に考慮した関係性値が算出される。

【0071】

【数4】

$$Rel(A, B) = \sqrt{\sum_k \{RW(M_k, A) \cdot RW(M_k, B) \cdot CSL(M_k, A, B)\}}$$

10

… (4)

【0072】

(その他の方法2：共演シーン時間＋出演時間)

別の例として、共演シーン時間CSL及び出演時間PSLを組み合わせて関係性値Relを算出する方法を紹介する。なお、CSL及びPSLの定義は既に説明したものと同一である。この例において、関係性値Rel(A, B)は、下記の式(5)により算出される。なお、右辺の平方根は省略してもよい。この方法によると、各動画における各人物の重要度を示す出演時間と、各動画における人物間の関係性の強さを示す共演シーン時間とを共に考慮した関係性値が算出される。

20

【0073】

【数5】

$$Rel(A, B) = \sqrt{\sum_k \{PSL(M_k, A) \cdot PSL(M_k, B) \cdot CSL(M_k, A, B)\}}$$

… (5)

【0074】

30

[1-3：関係性情報の拡張]

さて、これまで説明してきた関係性値は、関係性の強弱を所定の条件下で定量化した数値であった。確かに、この数値を利用することにより、登場物間の関係性を客観的に判断することが可能になる。但し、動画を要約する際に関係性の種類を考慮したいケースが生じることも想定されるため、ここで関係性値の定義を拡張しておきたい。

【0075】

例えば、親友と家族とが同等の関係性値で表現される可能性が考えられるであろう。これらを区別する方法としては、例えば、図13に示すように、関係性値に属性を設定しておき、その属性を利用して関係性の種類を判別する方法が考えられる。また、人物Aが人物Bに感じる親密度と、人物Bが人物Aに感じる親密度とは異なる可能性が考えられる。そこで、図13に示すように、関係性マトリックスの非対称性を許容することとする。図13の例では、人物Aにとっての人物Cは敵であり、人物Cにとっての人物Aは友人である。このように関係性マトリックスの構成を拡張することで、より柔軟に要約の条件を設定することが可能になる。

40

【0076】

以上、本実施形態に係る動画要約手法の概要及び関連技術について説明した。

【0077】

<2：実施形態>

本技術の一実施形態について説明する。本実施形態は、動画要約手法に関する。

【0078】

50



## [ 2 - 1 : 情報処理装置 1 0 0 の構成 ]

まず、図 1 4 ~ 図 1 6 を参照しながら、本実施形態に係る情報処理装置 1 0 0 の構成例について説明する。図 1 4 は、情報処理装置 1 0 0 の標準的な構成について説明するための説明図である。図 1 5 は、変形例 ( 変形例 # 1 ) に係る情報処理装置 1 0 0 の構成について説明するための説明図である。図 1 6 は、他の変形例 ( 変形例 # 2 ) に係る情報処理装置 1 0 0 の構成について説明するための説明図である。

### 【 0 0 7 9 】

#### ( 2 - 1 - 1 : 標準構成 )

図 1 4 に示すように、情報処理装置 1 0 0 は、主に、メタデータ取得部 1 0 1 と、関係性情報生成部 1 0 2 と、入力部 1 0 3 と、シーン抽出部 1 0 4 と、出力部 1 0 5 と、UI 表示部 1 0 6 と、動画取得部 1 0 7 とにより構成される。

10

### 【 0 0 8 0 】

動画の要約処理を開始すると、メタデータ取得部 1 0 1 は、動画タイムラインメタデータを取得する。例えば、メタデータ取得部 1 0 1 は、情報処理装置 1 0 0 の筐体内に設けられた記憶装置 ( 非図示 )、ネットワークに接続された記憶装置 ( 非図示 )、或いは、動画タイムラインメタデータを提供するサービスなどから動画タイムラインメタデータを取得する。そして、メタデータ取得部 1 0 1 により取得された動画タイムラインメタデータは、関係性情報生成部 1 0 2 に入力される。

### 【 0 0 8 1 】

動画タイムラインメタデータが入力されると、関係性情報生成部 1 0 2 は、入力された動画タイムラインメタデータを利用して関係性情報 ( 関係性値や属性などを含む情報 ) を生成する。そして、関係性情報生成部 1 0 2 により生成された関係性情報は、シーン抽出部 1 0 4 及び UI 表示部 1 0 6 に入力される。なお、入力部 1 0 3 を介してユーザにより関係性情報の変更操作が行われた場合、関係性情報生成部 1 0 2 は、関係性情報に変更操作の内容を反映する。また、変更後の関係性情報は、シーン抽出部 1 0 4 及び UI 表示部 1 0 6 に入力される。

20

### 【 0 0 8 2 】

関係性情報が入力されると、シーン抽出部 1 0 4 は、入力された関係性情報、及び入力部 1 0 3 を介してユーザにより選択された登場物の情報に基づいてシーンを抽出する。なお、シーン抽出部 1 0 4 による抽出処理の内容については後段において詳述する。

30

### 【 0 0 8 3 】

シーン抽出部 1 0 4 により抽出されたシーンの情報は、出力部 1 0 5 に入力される。抽出されたシーンの情報が入力されると、出力部 1 0 5 は、入力されたシーンの情報に基づいて要約結果を出力する。例えば、出力部 1 0 5 は、動画取得部 1 0 7 により取得された動画のうち、シーン抽出部 1 0 4 により抽出されたシーンを再生 ( ダイジェスト再生 ) したり、そのシーンを代表する画像を並べて表示 ( コミック表示 ) したりする。また、出力部 1 0 5 は、シーン抽出部 1 0 4 により抽出されたシーンに対応する区間の情報を外部の機器に向けて出力するように構成されていてもよい。なお、動画取得部 1 0 7 は、情報処理装置 1 0 0 の筐体内に設けられた記憶装置 ( 非図示 )、ネットワークに接続された記憶装置 ( 非図示 )、或いは、動画を提供するサービスなどから動画を取得する。

40

### 【 0 0 8 4 】

また、出力部 1 0 5 は、UI 表示部 1 0 6 により生成されるユーザインターフェースの画像を表示する。このユーザインターフェースは、ユーザに対して関係性情報を提示するために利用されるものである。また、このユーザインターフェースは、ユーザが関係性情報を変更する際に利用されるものである。このユーザインターフェースを構成する部品の生成及び表示制御は、UI 表示部 1 0 6 により実行される。なお、このユーザインターフェースの構成については後段において詳述する。

### 【 0 0 8 5 】

以上、情報処理装置 1 0 0 の標準的な構成について説明した。

### 【 0 0 8 6 】

50

( 2 - 1 - 2 : 変形例 # 1 ( 動画からメタデータを自動生成する構成 ) )

次に、図 1 5 を参照しながら、変形例 # 1 に係る情報処理装置 1 0 0 の構成について説明する。変形例 # 1 の構成は、動画タイムラインメタデータを動画から自動抽出する点で図 1 4 に示した標準的な構成と相違する。

【 0 0 8 7 】

図 1 5 に示すように、情報処理装置 1 0 0 は、主に、動画解析部 1 1 1 と、関係性情報生成部 1 0 2 と、入力部 1 0 3 と、シーン抽出部 1 0 4 と、出力部 1 0 5 と、UI 表示部 1 0 6 と、動画取得部 1 0 7 とにより構成される。

【 0 0 8 8 】

動画の要約処理を開始すると、動画解析部 1 1 1 は、動画取得部 1 0 7 により取得された動画を解析し、動画タイムラインメタデータを生成する。なお、動画取得部 1 0 7 は、情報処理装置 1 0 0 の筐体内に設けられた記憶装置 ( 非図示 ) 、ネットワークに接続された記憶装置 ( 非図示 ) 、或いは、動画を提供するサービスなどから動画を取得する。動画解析部 1 1 1 により生成された動画タイムラインメタデータは、関係性情報生成部 1 0 2 に入力される。

【 0 0 8 9 】

動画タイムラインメタデータが入力されると、関係性情報生成部 1 0 2 は、入力された動画タイムラインメタデータを利用して関係性情報を生成する。そして、関係性情報生成部 1 0 2 により生成された関係性情報は、シーン抽出部 1 0 4 及び UI 表示部 1 0 6 に入力される。なお、入力部 1 0 3 を介してユーザにより関係性情報の変更操作が行われた場合、関係性情報生成部 1 0 2 は、関係性情報に変更操作の内容を反映する。また、変更後の関係性情報は、シーン抽出部 1 0 4 及び UI 表示部 1 0 6 に入力される。

【 0 0 9 0 】

関係性情報が入力されると、シーン抽出部 1 0 4 は、入力された関係性情報、及び入力部 1 0 3 を介してユーザにより選択された登場物の情報に基づいてシーンを抽出する。なお、シーン抽出部 1 0 4 による抽出処理の内容については後段において詳述する。

【 0 0 9 1 】

シーン抽出部 1 0 4 により抽出されたシーンの情報は、出力部 1 0 5 に入力される。抽出されたシーンの情報が入力されると、出力部 1 0 5 は、入力されたシーンの情報に基づいて要約結果を出力する。例えば、出力部 1 0 5 は、動画取得部 1 0 7 により取得された動画のうち、シーン抽出部 1 0 4 により抽出されたシーンを再生 ( ダイジェスト再生 ) したり、そのシーンを代表する画像を並べて表示 ( コミック表示 ) したりする。また、出力部 1 0 5 は、シーン抽出部 1 0 4 により抽出されたシーンに対応する区間の情報を外部の機器に向けて出力するように構成されていてもよい。

【 0 0 9 2 】

また、出力部 1 0 5 は、UI 表示部 1 0 6 により生成されるユーザインターフェースの画像を表示する。このユーザインターフェースは、ユーザに対して関係性情報を提示するために利用されるものである。また、このユーザインターフェースは、ユーザが関係性情報を変更する際に利用されるものである。このユーザインターフェースを構成する部品の生成及び表示制御は、UI 表示部 1 0 6 により実行される。なお、このユーザインターフェースの構成については後段において詳述する。

【 0 0 9 3 】

以上、変形例 # 1 に係る情報処理装置 1 0 0 の構成について説明した。

【 0 0 9 4 】

( 2 - 1 - 3 : 変形例 # 2 ( 所定の関係性情報を利用する構成 ) )

次に、図 1 6 を参照しながら、変形例 # 2 に係る情報処理装置 1 0 0 の構成について説明する。変形例 # 2 の構成は、関係性情報を外部から取得する点で図 1 4 に示した標準的な構成と相違する。

【 0 0 9 5 】

図 1 6 に示すように、情報処理装置 1 0 0 は、主に、関係性情報取得部 1 2 1 と、入力

10

20

30

40

50

部 1 0 3 と、シーン抽出部 1 0 4 と、出力部 1 0 5 と、UI 表示部 1 0 6 と、動画取得部 1 0 7 とにより構成される。

#### 【 0 0 9 6 】

動画の要約処理を開始すると、関係性情報取得部 1 2 1 は、関係性情報を取得する。なお、関係性情報取得部 1 2 1 は、情報処理装置 1 0 0 の筐体内に設けられた記憶装置（非図示）、ネットワークに接続された記憶装置（非図示）、或いは、関係性情報を提供するサービスなどから関係性情報を取得する。そして、関係性情報取得部 1 2 1 により取得された関係性情報は、シーン抽出部 1 0 4 及び UI 表示部 1 0 6 に入力される。なお、入力部 1 0 3 を介してユーザにより関係性情報の変更操作が行われた場合、関係性情報取得部 1 2 1 は、関係性情報に変更操作の内容を反映する。また、変更後の関係性情報は、シーン抽出部 1 0 4 及び UI 表示部 1 0 6 に入力される。

10

#### 【 0 0 9 7 】

関係性情報が入力されると、シーン抽出部 1 0 4 は、入力された関係性情報、及び入力部 1 0 3 を介してユーザにより選択された登場物の情報に基づいてシーンを抽出する。なお、シーン抽出部 1 0 4 による抽出処理の内容については後段において詳述する。

#### 【 0 0 9 8 】

シーン抽出部 1 0 4 により抽出されたシーンの情報は、出力部 1 0 5 に入力される。抽出されたシーンの情報が入力されると、出力部 1 0 5 は、入力されたシーンの情報に基づいて要約結果を出力する。例えば、出力部 1 0 5 は、動画取得部 1 0 7 により取得された動画のうち、シーン抽出部 1 0 4 により抽出されたシーンを再生（ダイジェスト再生）したり、そのシーンを代表する画像を並べて表示（コミック表示）したりする。また、出力部 1 0 5 は、シーン抽出部 1 0 4 により抽出されたシーンに対応する区間の情報を外部の機器に向けて出力するように構成されていてもよい。

20

#### 【 0 0 9 9 】

また、出力部 1 0 5 は、UI 表示部 1 0 6 により生成されるユーザインターフェースの画像を表示する。このユーザインターフェースは、ユーザに対して関係性情報を提示するために利用されるものである。また、このユーザインターフェースは、ユーザが関係性情報を変更する際に利用されるものである。このユーザインターフェースを構成する部品の生成及び表示制御は、UI 表示部 1 0 6 により実行される。なお、このユーザインターフェースの構成については後段において詳述する。

30

#### 【 0 1 0 0 】

以上、変形例 # 2 に係る情報処理装置 1 0 0 の構成について説明した。

#### 【 0 1 0 1 】

##### [ 2 - 2 : 情報処理装置 1 0 0 の動作 ]

次に、図 1 7 ~ 図 2 7 を参照しながら、本実施形態に係る情報処理装置 1 0 0 の動作について説明する。なお、以下では、説明の都合上、情報処理装置 1 0 0 の標準的な構成を想定して説明を進める。

#### 【 0 1 0 2 】

##### ( 2 - 2 - 1 : 基本動作 )

まず、図 1 7 を参照しながら、情報処理装置 1 0 0 の基本的な動作について説明する。

40

#### 【 0 1 0 3 】

図 1 7 に示すように、動画の要約処理を開始すると、情報処理装置 1 0 0 は、メタデータ取得部 1 0 1 の機能により、動画タイムラインメタデータを取得する（S 1 0 1）。次いで、情報処理装置 1 0 0 は、関係性情報生成部 1 0 2 の機能により、動画タイムラインメタデータを用いて関係性情報を生成する（S 1 0 2）。次いで、情報処理装置 1 0 0 は、ユーザにより登場物が選択されたか否かを判定する（S 1 0 3）。登場物が選択された場合、情報処理装置 1 0 0 は、処理をステップ S 1 0 4 に進める。一方、登場物が選択されていない場合、情報処理装置 1 0 0 は、処理をステップ S 1 0 3 に戻す。

#### 【 0 1 0 4 】

処理をステップ S 1 0 4 に進めた場合、情報処理装置 1 0 0 は、シーン抽出部 1 0 4 の

50

機能により、関係性情報に基づく動画の要約処理を実行する（S 1 0 4）。なお、ステップ S 1 0 4 にて実行される要約処理の詳細については後述する。次いで、情報処理装置 1 0 0 は、出力部 1 0 5 の機能により、要約処理の結果を出力する（S 1 0 5）。要約処理の結果を出力すると、情報処理装置 1 0 0 は、動画の要約処理に関する一連の処理を終了する。

#### 【0 1 0 5】

以上、情報処理装置 1 0 0 の基本的な動作について説明した。

#### 【0 1 0 6】

（2 - 2 - 2：要約処理 # 1（関係性値の総和を最大化する方法））

次に、図 1 8 及び図 1 9 を参照しながら、ステップ S 1 0 4 にて実行される要約処理の一例について、より詳細に説明する。ここでは、関係性値の総和を最大化するようにシーンを選択する方法（以下、要約処理 # 1）について説明する。図 1 8 は、要約処理 # 1 の概要について説明するための説明図である。また、図 1 9 は、要約処理 # 1 に係る情報処理装置 1 0 0 の動作について説明するための説明図である。

10

#### 【0 1 0 7】

まず、図 1 8 を参照する。ユーザにより登場物 A が選択された場合、図 1 8 に示すように、情報処理装置 1 0 0 は、動画に登場する各登場物と登場物 A との間関係性値 R を算出する。そして、情報処理装置 1 0 0 は、関係性値 R が大きい順に所定数の登場物を選択する。図 1 8 の例では、関係性値 R が大きい順に、登場物 C、登場物 D、登場物 B、登場物 G が選択されている。このようにして登場物を選択することにより、選択した登場物と登場物 A との間における関係性値 R の総和が最大化される。登場物を選択すると、情報処理装置 1 0 0 は、選択した各登場物が登場する動画中の区間を抽出する。そして、情報処理装置 1 0 0 は、抽出した区間の動画を時系列に沿って再生したり、各区間を代表する画像を並べて表示したりする。

20

#### 【0 1 0 8】

次に、図 1 9 を参照する。図 1 8 の例では単純に関係性値 R が大きい順に所定数の登場物を選択したが、関係性値 R の総和が最大化されるように所定数の区間を抽出したいケースも考えられる。また、各区間に登場する全ての登場物に関する関係性値 R を考慮したいケースも考えられる。このようなケースに適応するには、例えば、図 1 9 に示すような処理の流れに沿って動画の要約処理を実行すればよい。

30

#### 【0 1 0 9】

図 1 9 に示すように、まず、情報処理装置 1 0 0 は、ユーザにより選択された登場物が登場する区間（第 1 の区間）を検出する（S 1 1 1）。次いで、情報処理装置 1 0 0 は、 $n = 2 \sim N$  に関する処理ループを開始する。まず、情報処理装置 1 0 0 は、第  $n - 1$  の区間以降の区間に登場する登場物を選択する（S 1 1 2）。例えば、情報処理装置 1 0 0 は、ユーザにより選択された登場物との間における関係性値 R が所定の閾値を越える登場物を選択する。但し、ステップ S 1 1 2 で登場物が選択できない場合、情報処理装置 1 0 0 は、処理をステップ S 1 1 4 に進める。登場物を選択した情報処理装置 1 0 0 は、選択した登場物が登場する区間を検出し、検出した区間を第  $n$  の区間に設定する（S 1 1 3）。 $n = 2 \sim N$  についてステップ S 1 1 2 及び S 1 1 3 の処理を実行した後、情報処理装置 1 0 0 は、処理をステップ S 1 1 4 に進める。

40

#### 【0 1 1 0】

処理をステップ S 1 1 4 に進めた情報処理装置 1 0 0 は、ユーザにより選択された登場物と第 2 ～第  $n$  の区間に登場する各登場物との間における関係性値 R を算出し、算出した関係性値 R の総和を計算する（S 1 1 4）。次いで、情報処理装置 1 0 0 は、ステップ S 1 1 4 にて算出した関係性値 R の総和が最大か否かを判定する（S 1 1 5）。関係性値 R の総和が最大になった場合、情報処理装置 1 0 0 は、処理をステップ S 1 1 6 に進める。一方、関係性値 R の総和が最大になっていない場合、情報処理装置 1 0 0 は、処理をステップ S 1 1 1 の後段に戻す。処理をステップ S 1 1 6 に進めた場合、情報処理装置 1 0 0 は、抽出したシーンの情報として第 1 ～第  $N$  の区間を出力し（S 1 1 6）、要約処理 # 1

50

に係る一連の処理を終了する。

【0111】

以上、ステップS104にて実行される要約処理の一例について説明した。

【0112】

(2-2-3: 要約処理#2 (シーン切り替え時の関係性値和を最大化1))

次に、図20及び図21を参照しながら、ステップS104にて実行される要約処理の他の一例について、より詳細に説明する。ここでは、シーン切り替え時の関係性値和が最大化されるように登場物を選択する方法(以下、要約処理#2)について説明する。図20は、要約処理#2の概要について説明するための説明図である。また、図21は、要約処理#2に係る情報処理装置100の動作について説明するための説明図である。

10

【0113】

まず、図20を参照する。図20に示すように、例えば、登場物Aが登場する区間 $S_A$ 、登場物Fが登場する区間 $S_F$ 、登場物Dが登場する区間 $S_D$ 、登場物Bが登場する区間 $S_B$ 、登場物Gが登場する区間 $S_G$ を順次選択した場合について考えてみよう。この場合、シーンの切り替え前後に登場する登場物の組は、(A, F)、(F, D)、(D, B)、(B, G)となる。そこで、情報処理装置100は、登場物Aと登場物Fとの間における関係性値R、登場物Fと登場物Fとの間における関係性値R、登場物Dと登場物Bとの間における関係性値R、登場物Bと登場物Gとの間における関係性値Rを算出し、それらの総和(図20の例では3.15)を計算する。

【0114】

20

図20には区間 $S_A$ 、区間 $S_F$ 、区間 $S_D$ 、区間 $S_B$ 、区間 $S_G$ を順次選択するケースを示したが、同様にして、情報処理装置100は、選択する区間の組み合わせを変更しながら関係性値Rの総和を算出する。そして、情報処理装置100は、関係性値Rの総和が最大となる区間の組み合わせを検出する。区間の組み合わせを検出した情報処理装置100は、検出した区間の動画を時系列に沿って再生したり、各区間を代表する画像を並べて表示したりする。要約処理#2の手法を適用すると、ユーザにより選択された登場物と間接的に強い関係性を有する登場物の登場シーンをダイジェスト再生やコミック表示に含めることが可能になる。例えば、仲が良い友人の繋がりを考慮したダイジェスト再生などを実現することが可能になる。

【0115】

30

ここで、図21を参照しながら、要約処理#2に係る情報処理装置100の動作について、より詳細に説明する。図21に示すように、まず、情報処理装置100は、ユーザにより選択された登場物が登場する区間(第1の区間)を検出する(S121)。次いで、情報処理装置100は、 $n = 2 \sim N$ に関する処理ループを開始する。まず、情報処理装置100は、第 $n - 1$ の区間以降の区間に登場する登場物を選択する(S122)。例えば、情報処理装置100は、ユーザにより選択された登場物との間における関係性値Rが所定の閾値を越える登場物を選択する。但し、ステップS122で登場物が選択できない場合、情報処理装置100は、処理をステップS124に進める。

【0116】

40

登場物を選択した情報処理装置100は、選択した登場物が登場する区間を検出し、検出した区間を第 $n$ の区間に設定する(S123)。 $n = 2 \sim N$ についてステップS122及びS123の処理を実行した後、情報処理装置100は、処理をステップS124に進める。処理をステップS124に進めた情報処理装置100は、 $n = 2 \sim N$ について第 $n - 1$ の区間に登場する登場物と第 $n$ の区間に登場する登場物との間における関係性値Rを算出し、算出した関係性値Rの総和を計算する(S124)。

【0117】

次いで、情報処理装置100は、ステップS124にて算出した関係性値Rの総和が最大か否かを判定する(S125)。関係性値Rの総和が最大になった場合、情報処理装置100は、処理をステップS126に進める。一方、関係性値Rの総和が最大になっていない場合、情報処理装置100は、処理をステップS121の後段に戻す。処理をステッ

50

ブ S 1 2 6 に進めた場合、情報処理装置 1 0 0 は、抽出したシーンの情報として第 1 ~ 第 N の区間を出力し ( S 1 2 6 )、要約処理 # 2 に係る一連の処理を終了する。

【 0 1 1 8 】

以上、ステップ S 1 0 4 にて実行される要約処理の他の一例について説明した。

【 0 1 1 9 】

( 2 - 2 - 4 : 要約処理 # 3 ( シーン切り替え時の関係性値和を最大化 2 ) )

次に、図 2 2 及び図 2 3 を参照しながら、ステップ S 1 0 4 にて実行される要約処理の他の一例について、より詳細に説明する。ここでは、シーン切り替え時の関係性値和が最大化されるように登場物を選択する方法 ( 以下、要約処理 # 3 ) について説明する。図 2 2 は、要約処理 # 3 の概要について説明するための説明図である。また、図 2 3 は、要約処理 # 3 に係る情報処理装置 1 0 0 の動作について説明するための説明図である。

10

【 0 1 2 0 】

上記の要約処理 # 2 の場合、情報処理装置 1 0 0 は、登場物及びその登場物の登場シーンを逐次選択していき、シーンの切り替わりタイミングにおける登場物間の関係性値を求める。このとき、情報処理装置 1 0 0 は、選択された登場物間の関係性値を求めている。そのため、上記の要約処理 # 2 においては、各シーンに登場する他の登場物に関する関係性値が考慮されていない。一方、ここで説明する要約処理 # 3 は、各シーンに登場する全ての登場物について関係性を考慮する。

【 0 1 2 1 】

例えば、図 2 2 に示すように、ユーザにより登場物 A が選択された場合、情報処理装置 1 0 0 は、登場物 A が登場する区間  $S_1$  を抽出する。次いで、情報処理装置 1 0 0 は、区間  $S_1$  に登場する全ての登場物を検出する。図 2 2 の例では、区間  $S_1$  に登場物 A と登場物 G とが登場している。次いで、情報処理装置 1 0 0 は、抽出した区間  $S_1$  以降の区間  $S_2$  を選択する。そして、情報処理装置 1 0 0 は、区間  $S_2$  に登場する全ての登場物を検出する。図 2 2 の例では、区間  $S_2$  に登場物 F が登場している。そこで、情報処理装置 1 0 0 は、登場物 A と登場物 F との間における関係性値 R、及び、登場物 G と登場物 F との間における関係性値 R を算出し、それら関係性値 R の総和を計算する。

20

【 0 1 2 2 】

情報処理装置 1 0 0 は、区間  $S_2$  の位置を移動させながら同様の方法で関係性値 R の総和を計算し、関係性値 R の総和が最大となる区間  $S_2$  を探索する。ここでは、登場物 F が登場する区間  $S_2$  が、関係性値 R の総和が最大となる区間  $S_2$  として確定されたものとする。区間  $S_2$  の位置を確定させると、情報処理装置 1 0 0 は、区間  $S_2$  以降の区間  $S_3$  を選択する。そして、情報処理装置 1 0 0 は、区間  $S_3$  に登場する全ての登場物を検出する。図 2 2 の例では、区間  $S_3$  に登場物 D と登場物 C とが登場している。そこで、情報処理装置 1 0 0 は、登場物 F と登場物 D との間における関係性値 R、及び、登場物 F と登場物 C との間における関係性値 R を算出し、それら関係性値 R の総和を計算する。

30

【 0 1 2 3 】

区間  $S_2$  を確定された場合と同様に、情報処理装置 1 0 0 は、関係性値 R の総和が最大となる区間  $S_3$  を探索する。区間  $S_3$  を確定させると、情報処理装置 1 0 0 は、順次、区間  $S_4$  及び区間  $S_5$  についても同様の方法で位置を確定させる。区間の組み合わせを確定させた情報処理装置 1 0 0 は、確定した区間の動画を時系列に沿って再生したり、各区間を代表する画像を並べて表示したりする。要約処理 # 3 の手法を適用すると、共演者との関係性も考慮に入れて、ユーザにより選択された登場物と間接的に強い関係性を有する登場物の登場シーンをダイジェスト再生やコミック表示に含めることが可能になる。

40

【 0 1 2 4 】

ここで、図 2 3 を参照しながら、要約処理 # 3 に係る情報処理装置 1 0 0 の動作について、より詳細に説明する。図 2 3 に示すように、まず、情報処理装置 1 0 0 は、ユーザにより選択された登場物が登場する区間 ( 第 1 の区間 ) を検出する ( S 1 3 1 )。次いで、情報処理装置 1 0 0 は、 $n = 2 \sim N$  に関する処理ループを開始する。まず、情報処理装置 1 0 0 は、第  $n - 1$  の区間以降の区間を選択し、選択した区間に登場する全ての登場物と

50

、第  $n - 1$  の区間に登場する全ての登場物との間における関係性値の総和を算出する ( S 1 3 2 )。次いで、情報処理装置 1 0 0 は、関係性値の総和が最大になる区間を検出し、検出した区間を第  $n$  の区間に設定する ( S 1 3 3 )。

#### 【 0 1 2 5 】

$n = 2 \sim N$  についてステップ S 1 3 2 及び S 1 3 3 の処理を実行した後、情報処理装置 1 0 0 は、処理をステップ S 1 3 4 に進める。処理をステップ S 1 3 4 に進めた情報処理装置 1 0 0 は、抽出したシーンの情報として第 1 ~ 第  $N$  の区間を出力し ( S 1 3 4 )、要約処理 # 3 に係る一連の処理を終了する。

#### 【 0 1 2 6 】

以上、ステップ S 1 0 4 にて実行される要約処理の他の一例について説明した。

10

#### 【 0 1 2 7 】

( 2 - 2 - 5 : 要約処理 # 4 ( シーン切り替え時の関係性値差を最小化 ) )

次に、図 2 4 及び図 2 5 を参照しながら、ステップ S 1 0 4 にて実行される要約処理の他の一例について、より詳細に説明する。ここでは、シーン切り替え時における関係性値の差が最小化されるように登場物を選択する方法 ( 以下、要約処理 # 4 ) について説明する。図 2 4 は、要約処理 # 4 の概要について説明するための説明図である。また、図 2 5 は、要約処理 # 4 に係る情報処理装置 1 0 0 の動作について説明するための説明図である。

#### 【 0 1 2 8 】

まず、図 2 4 を参照する。ユーザにより登場物 A が選択された場合、情報処理装置 1 0 0 は、登場物 A が登場する区間  $S_1$  を選択する。そして、情報処理装置 1 0 0 は、選択した区間  $S_1$  に登場する全ての登場物を検出し、検出した各登場物と登場物 A との間における関係性値を算出する。図 2 4 の例では、登場物 A と登場物 A との間における関係性値、及び登場物 A と登場物 G との間における関係性値が算出される。関係性値を算出した情報処理装置 1 0 0 は、算出した関係性値の総和を計算する。次いで、情報処理装置 1 0 0 は、区間  $S_1$  以降の区間  $S_2$  を選択し、区間  $S_2$  に登場する全ての登場物に関して、各登場物と登場物 A との間における関係性値を算出する。そして、情報処理装置 1 0 0 は、区間  $S_2$  に登場する登場物に関する関係性値の総和を計算する。

20

#### 【 0 1 2 9 】

さらに、情報処理装置 1 0 0 は、区間  $S_1$  に関して算出された関係性値の総和と区間  $S_2$  に関して算出された関係性値の総和との差を算出する。この差が最小でない場合、情報処理装置 1 0 0 は、区間  $S_2$  の位置を変更して区間  $S_2$  に関する関係性値の総和を算出し、区間  $S_1$  に関して算出された関係性値の総和と区間  $S_2$  に関して算出された関係性値の総和との差を算出する。差が最小になった場合、情報処理装置 1 0 0 は、区間  $S_2$  の位置を確定させ、同様の方法で区間  $S_2$  に次ぐ区間  $S_3$  の位置を探索する。その後も、情報処理装置 1 0 0 は、同様の処理を逐次実行して区間  $S_4$  及び区間  $S_5$  の位置を確定させる。このようにして区間の組み合わせを確定させた情報処理装置 1 0 0 は、確定した区間の動画を時系列に沿って再生したり、各区間を代表する画像を並べて表示したりする。

30

#### 【 0 1 3 0 】

要約処理 # 4 の手法を適用すると、ユーザにより選択された登場物と強い関係性を有する登場物の登場シーンを用いてダイジェスト再生やコミック表示を実現することができる。また、シーンの切り替え前後で関係性値の差が小さくなることから、自然なシーンの切り替えが実現される。

40

#### 【 0 1 3 1 】

ここで、図 2 5 を参照しながら、要約処理 # 4 に係る情報処理装置 1 0 0 の動作について、より詳細に説明する。図 2 5 に示すように、まず、情報処理装置 1 0 0 は、ユーザにより選択された登場物が登場する区間 ( 第 1 の区間 ) を検出する ( S 1 4 1 )。次いで、情報処理装置 1 0 0 は、 $n = 2 \sim N$  に関する処理ループを開始する。まず、情報処理装置 1 0 0 は、第  $n - 1$  の区間以降の区間を選択し、選択した区間に登場する全ての登場物と、ユーザにより選択された登場物との間における関係性値の総和を算出する ( S 1 4 2 )

50

。次いで、情報処理装置 100 は、選択した区間のうち、関係性値の総和が第  $n - 1$  の区間に関する関係性値の総和に最も近い区間を検出し、検出した区間を第  $n$  の区間に設定する (S143)。

#### 【0132】

$n = 2 \sim N$  についてステップ S142 及び S143 の処理を実行した後、情報処理装置 100 は、処理をステップ S144 に進める。処理をステップ S144 に進めた情報処理装置 100 は、抽出したシーンの情報として第 1 ~ 第  $N$  の区間を出力し (S144)、要約処理 # 4 に係る一連の処理を終了する。

#### 【0133】

以上、ステップ S104 にて実行される要約処理の他の一例について説明した。

10

#### 【0134】

(2-2-6: 要約処理 # 5 (その他の構成))

これまで説明してきた要約処理は、ユーザにより選択された登場物との間で直接的又は間接的に強い関係性を有する登場物の登場シーンを抽出するものであった。しかし、関係性値の総和が最小になるように区間を選択するなどの条件変更を行うことも可能である。また、要約処理 # 2 及び # 3 のようにシーン切り替え時の関係性に基づいて区間を逐次選択する手法の場合、ユーザにより選択された登場物との関係性が次第に薄れていってしまう可能性がある。そこで、ランダムなタイミングでユーザにより選択された登場物との直接的な関係性が強い登場物が登場する区間を挿入するなどの工夫を追加してもよい。また、上記の説明では比較的単純なアルゴリズムを例示したが、最適な区間の組み合わせを検出する方法として、例えば、局所探索、焼きなまし法、遺伝的アルゴリズムなどを利用する方法も考えられる。このように、上記の要約処理は必要に応じて変形可能である。

20

#### 【0135】

以上、本実施形態に係る情報処理装置 100 の動作について説明した。なお、図 26 及び図 27 に示すように、いずれの方法を用いても、ダイジェスト動画における関係性値の時系列変化は、ユーザにより選択される登場物に応じて変わる。また、ダイジェスト動画の構成も、ユーザにより選択される登場物に応じて変化する。もちろん、要約結果をコミック表示にした場合も、ユーザにより選択される登場物に応じて表示内容が変化する。

#### 【0136】

[2-3: 関係性値の編集方法]

30

次に、図 28 ~ 図 30 を参照しながら、本実施形態に係る関係性値の編集方法について説明する。図 28 は、関係性値の編集操作に用いるユーザインターフェースの構成例を示した説明図である。図 29 は、関係性値の編集方法について説明するための説明図である。図 30 は、関係性値の編集に伴う情報処理装置 100 の動作について説明するための説明図である。

#### 【0137】

(2-3-1: 操作 UI の構成)

まず、図 28 及び図 29 を参照しながら、関係性値の編集操作に用いるユーザインターフェースの構成例について説明する。

#### 【0138】

40

図 28 に示すように、情報処理装置 100 は、UI 表示部 106 の機能により、編集領域、及び各登場物を表すオブジェクトを表示画面に表示する。この編集領域は、コミック表示やダイジェスト再生用のウィンドウと同じ表示画面上に表示されてもよいし、或いは、単独で表示画面上に表示されてもよい。編集領域の内部には、動画に登場する登場物のオブジェクトが表示される。また、編集領域の内部には、図 28 に示すような同心円状のガイドが表示されていてもよい。図 28 の例では、動画に登場する登場人物 A ~ D のオブジェクト及び同心円状のガイドが表示されている。

#### 【0139】

図 28 の例において、ユーザにより選択された登場物は、登場人物 A である。そのため、登場人物 A のオブジェクトは、編集領域の中央に表示される。また、登場人物 B ~ D の

50



オブジェクトは、各登場物間の関係性値に基づいて配置される。例えば、登場人物 A のオブジェクトと登場人物 B のオブジェクトとの間の距離は、登場人物 A のオブジェクトと登場人物 C のオブジェクトとの間の距離よりも短い。この配置は、登場人物 A と登場人物 B との間における関係性値の方が、登場人物 A と登場人物 C との間における関係性値よりも大きいことを示している。また、登場人物 B ~ D のオブジェクトが表示される位置は、登場人物 B と登場人物 C との間における関係性値、登場人物 C と登場人物 D との間における関係性値、登場人物 B と登場人物 D との間における関係性値も考慮して決定される。

#### 【 0 1 4 0 】

上記のように、動画に登場する各登場物のオブジェクトを関係性値に基づいて配置することにより、一見して登場物間の関係性を把握することが可能になる。また、同心円状のガイドを併せて表示することにより、ユーザにより選択された登場物と他の登場物との関係性を正確に把握することが可能になる。なお、ユーザにより選択された登場物との間における関係性値が所定の閾値を下回る登場物のオブジェクトを編集領域に表示しないようにしてもよい。このような表示方法を適用すると、見やすい表示となり、編集対象として注目すべき登場物を素早く見つけることが可能になる。

#### 【 0 1 4 1 】

ここで関係性値の編集操作について説明する。関係性値を編集する場合、ユーザは、図 2 9 に示すように、編集対象の登場物を表すオブジェクトの位置を変えることで、その登場物と他の登場物との間における関係性値を変更することができる。例えば、図 2 9 に示すように、登場人物 B のオブジェクトを左下方向へ、登場人物 C のオブジェクトを中心方向へと移動させると、登場人物 A と登場人物 B との間における関係性値が減少し、登場人物 A と登場人物 C との間における関係性値が増加する。また、図 2 9 の例では、登場人物 B と登場人物 D との間における関係性値が減少し、登場人物 C と登場人物 D との間における関係性値が増大する。

#### 【 0 1 4 2 】

( 2 - 3 - 2 : 関係性値の編集処理に伴う情報処理装置 1 0 0 の動作 )

図 2 9 に示すような編集操作を行った場合、情報処理装置 1 0 0 は、図 3 0 に示すように、オブジェクトの位置関係に基づいて関係性マトリックスを更新する。例えば、図 2 9 に例示した編集操作を行った場合、情報処理装置 1 0 0 は、登場人物 A と登場人物 B との間における関係性値を距離に応じた値に減少させ ( 0 . 8    0 . 1 )、登場人物 A と登場人物 C との間における関係性値を距離に応じた値に増加させる ( 0 . 2    0 . 7 )。さらに、情報処理装置 1 0 0 は、登場人物 B と登場人物 C との間における関係性値を距離に応じた値に増加させる ( 0 . 1    0 . 3 )。そして、情報処理装置 1 0 0 は、更新後の関係性マトリックスに基づいてダイジェスト動画の構成やコミック表示の構成を変化させる。

#### 【 0 1 4 3 】

以上、関係性値の編集操作に用いるユーザインターフェースの構成例、及び関係性値の編集に伴う情報処理装置 1 0 0 の動作について説明した。

#### 【 0 1 4 4 】

[ 2 - 4 : 変形例 ( テンプレートを利用した動画要約手法 ) ]

これまで、関係性値を利用した動画要約手法について説明してきた。ここでは、登場物間の関係性に関する属性情報が得られる場合、或いは、動画の各区間におけるカメラワークやエフェクトなどの情報が得られる場合に、これらの情報を用いてストーリー性のある要約結果が得られるようにする手法について説明する。この手法においては、図 3 1 に示すようなテンプレートが利用される。

#### 【 0 1 4 5 】

( 2 - 4 - 1 : テンプレートの構成 )

まず、図 3 1 を参照しながら、テンプレートの構成について説明する。図 3 1 に示すように、テンプレートには、関係性の種類を示す属性情報が含まれる。また、テンプレートには、シーンの個数などの情報が含まれていてもよい。例えば、友人 ( 2 シーン )    他人 ( 1 シーン )    本人 ( 5 シーン )    敵 ( 1 シーン )    恋人 ( 3 シーン ) などといった情報

10

20

30

40

50

がテンプレートに含まれていてもよい。また、カメラワークやエフェクトなどの情報がメタデータとして得られている場合には、このメタデータも動画の要約処理に利用される。

【0146】

例えば、図31の例において、登場人物Aが選択された場合、情報処理装置100は、登場人物Aと友人関係にある登場人物が登場するシーンのうち、メタデータがパンを示すシーンを選択する。次いで、情報処理装置100は、登場人物Aと他人関係にある登場人物が登場するシーンのうち、メタデータがチルトを示すシーンを選択する。次いで、情報処理装置100は、登場人物Aが登場するシーンのうち、メタデータがズームを示すシーンを選択する。次いで、情報処理装置100は、登場人物Aと敵対関係にある登場人物が登場するシーンのうち、メタデータがズームアウトを示すシーンを選択する。次いで、情報処理装置100は、登場人物Aと恋人関係にある登場人物が登場するシーンのうち、メタデータがフェードアウトを示すシーンを選択する。

10

【0147】

一方、図31の例において、登場人物Bが選択された場合、情報処理装置100は、登場人物Bと友人関係にある登場人物が登場するシーンのうち、メタデータがパンを示すシーンを選択する。次いで、情報処理装置100は、登場人物Bと他人関係にある登場人物が登場するシーンのうち、メタデータがチルトを示すシーンを選択する。次いで、情報処理装置100は、登場人物Bが登場するシーンのうち、メタデータがズームを示すシーンを選択する。次いで、情報処理装置100は、登場人物Bと敵対関係にある登場人物が登場するシーンのうち、メタデータがズームアウトを示すシーンを選択する。次いで、情報処理装置100は、登場人物Bと恋人関係にある登場人物が登場するシーンのうち、メタデータがフェードアウトを示すシーンを選択する。

20

【0148】

このように、テンプレートを利用することにより、ストーリー性のあるダイジェスト動画を生成することが可能になる。また、関係性値を利用した動画要約手法と同様に、選択する登場物に応じてダイジェスト動画の構成が動的に変更される。

【0149】

(2-4-2: 要約処理の流れ)

ここで、図32を参照しながら、テンプレートを利用した動画要約手法に係る情報処理装置100の動作について説明する。図32は、テンプレートを利用した動画要約手法に係る情報処理装置100の動作について説明するための説明図である。

30

【0150】

図32に示すように、まず、情報処理装置100は、テンプレートを取得する(S151)。例えば、情報処理装置100は、筐体内に設けられた記憶装置(非図示)、ネットワークに接続された記憶装置(非図示)、或いは、テンプレートを提供するサービスなどからテンプレートを取得する。次いで、情報処理装置100は、 $n = 1 \sim N$ について、テンプレートの $n$ 番目に記載の属性に適合する区間を検出し、第 $n$ の区間に設定する(S152)。 $n = 1 \sim N$ についてステップS152の処理を実行した情報処理装置100は、処理をステップS153に進める。処理をステップS153に進めた情報処理装置100は、抽出したシーンの情報として第1～第 $N$ の区間を出力し(S153)、要約処理に係る一連の処理を終了する。

40

【0151】

以上、テンプレートを利用した動画要約手法に係る情報処理装置100の動作について説明した。

【0152】

(2-4-3: テンプレートの自動生成方法)

上記の説明においてはテンプレートを外部から取得するものとしたが、機械学習などを利用してテンプレートを自動生成する方法も考えられる。そこで、図33を参照しながら、テンプレートの自動生成方法について簡単に説明する。

【0153】

50

図 3 3 に示すように、まず、情報処理装置 1 0 0 は、予め用意した複数の作品（動画や静止画群など）を解析し、各作品に含まれる典型的な関係性の時系列変化を抽出する（S 1 6 1）。次いで、情報処理装置 1 0 0 は、ステップ S 1 6 1 で抽出した関係性の時系列変化を学習データとして利用し、機械学習により特徴的な関係性の時系列変化を抽出してテンプレートを生成する（S 1 6 2）。次いで、情報処理装置 1 0 0 は、ステップ S 1 6 2 で生成したテンプレートを出力し（S 1 6 3）、テンプレートの自動生成に係る一連の処理を終了する。例えば、情報処理装置 1 0 0 は、筐体内に設けられた記憶装置（非図示）又はネットワークに接続された記憶装置（非図示）にテンプレートを格納したり、テンプレートを提供するサービスなどにテンプレートを提供したりする。

【 0 1 5 4 】

以上、テンプレートの自動生成方法について説明した。

【 0 1 5 5 】

< 3 : ハードウェア構成例 >

上記の情報処理装置 1 0 0 が有する各構成要素の機能は、例えば、図 3 4 に示すハードウェア構成を用いて実現することが可能である。つまり、当該各構成要素の機能は、コンピュータプログラムを用いて図 3 4 に示すハードウェアを制御することにより実現される。なお、このハードウェアの形態は任意であり、例えば、パーソナルコンピュータ、携帯電話、PHS、PDA等の携帯情報端末、ゲーム機、又は種々の情報家電がこれに含まれる。但し、上記のPHSは、Personal Handy-phone Systemの略である。また、上記のPDAは、Personal Digital Assistantの略である。

【 0 1 5 6 】

図 3 4 に示すように、このハードウェアは、主に、CPU 9 0 2 と、ROM 9 0 4 と、RAM 9 0 6 と、ホストバス 9 0 8 と、ブリッジ 9 1 0 と、を有する。さらに、このハードウェアは、外部バス 9 1 2 と、インターフェース 9 1 4 と、入力部 9 1 6 と、出力部 9 1 8 と、記憶部 9 2 0 と、ドライブ 9 2 2 と、接続ポート 9 2 4 と、通信部 9 2 6 と、を有する。但し、上記のCPUは、Central Processing Unitの略である。また、上記のROMは、Read Only Memoryの略である。そして、上記のRAMは、Random Access Memoryの略である。

【 0 1 5 7 】

CPU 9 0 2 は、例えば、演算処理装置又は制御装置として機能し、ROM 9 0 4、RAM 9 0 6、記憶部 9 2 0、又はリムーバブル記録媒体 9 2 8 に記録された各種プログラムに基づいて各構成要素の動作全般又はその一部を制御する。ROM 9 0 4 は、CPU 9 0 2 に読み込まれるプログラムや演算に用いるデータ等を格納する手段である。RAM 9 0 6 には、例えば、CPU 9 0 2 に読み込まれるプログラムや、そのプログラムを実行する際に適宜変化する各種パラメータ等が一時的又は永続的に格納される。

【 0 1 5 8 】

これらの構成要素は、例えば、高速なデータ伝送が可能なホストバス 9 0 8 を介して相互に接続される。一方、ホストバス 9 0 8 は、例えば、ブリッジ 9 1 0 を介して比較的データ伝送速度が低速な外部バス 9 1 2 に接続される。また、入力部 9 1 6 としては、例えば、マウス、キーボード、タッチパネル、ボタン、スイッチ、及びレバー等が用いられる。さらに、入力部 9 1 6 としては、赤外線やその他の電波を利用して制御信号を送信することが可能なリモートコントローラ（以下、リモコン）が用いられることもある。

【 0 1 5 9 】

出力部 9 1 8 としては、例えば、CRT、LCD、PDP、又はELD等のディスプレイ装置、スピーカ、ヘッドホン等のオーディオ出力装置、プリンタ、携帯電話、又はファクシミリ等、取得した情報を利用者に対して視覚的又は聴覚的に通知することが可能な装置である。但し、上記のCRTは、Cathode Ray Tubeの略である。また、上記のLCDは、Liquid Crystal Displayの略である。そして、上記のPDPは、Plasma Display Panelの略である。さらに、上記

10

20

30

40

50

の E L D は、E l e c t r o - L u m i n e s c e n c e D i s p l a y の略である。

【 0 1 6 0 】

記憶部 9 2 0 は、各種のデータを格納するための装置である。記憶部 9 2 0 としては、例えば、ハードディスクドライブ ( H D D ) 等の磁気記憶デバイス、半導体記憶デバイス、光記憶デバイス、又は光磁気記憶デバイス等が用いられる。但し、上記の H D D は、H a r d D i s k D r i v e の略である。

【 0 1 6 1 】

ドライブ 9 2 2 は、例えば、磁気ディスク、光ディスク、光磁気ディスク、又は半導体メモリ等のリムーバブル記録媒体 9 2 8 に記録された情報を読み出し、又はリムーバブル記録媒体 9 2 8 に情報を書き込む装置である。リムーバブル記録媒体 9 2 8 は、例えば、D V D メディア、B l u - r a y メディア、H D D V D メディア、各種の半導体記憶メディア等である。もちろん、リムーバブル記録媒体 9 2 8 は、例えば、非接触型 I C チップを搭載した I C カード、又は電子機器等であってもよい。但し、上記の I C は、I n t e g r a t e d C i r c u i t の略である。

【 0 1 6 2 】

接続ポート 9 2 4 は、例えば、U S B ポート、I E E E 1 3 9 4 ポート、S C S I、R S - 2 3 2 C ポート、又は光オーディオ端子等のような外部接続機器 9 3 0 を接続するためのポートである。外部接続機器 9 3 0 は、例えば、プリンタ、携帯音楽プレーヤ、デジタルカメラ、デジタルビデオカメラ、又は I C レコーダ等である。但し、上記の U S B は、U n i v e r s a l S e r i a l B u s の略である。また、上記の S C S I は、S m a l l C o m p u t e r S y s t e m I n t e r f a c e の略である。

【 0 1 6 3 】

通信部 9 2 6 は、ネットワーク 9 3 2 に接続するための通信デバイスであり、例えば、有線又は無線 L A N、B l u e t o o t h ( 登録商標 )、又は W U S B 用の通信カード、光通信用のルータ、A D S L 用のルータ、又は各種通信用のモデム等である。また、通信部 9 2 6 に接続されるネットワーク 9 3 2 は、有線又は無線により接続されたネットワークにより構成され、例えば、インターネット、家庭内 L A N、赤外線通信、可視光通信、放送、又は衛星通信等である。但し、上記の L A N は、L o c a l A r e a N e t w o r k の略である。また、上記の W U S B は、W i r e l e s s U S B の略である。そして、上記の A D S L は、A s y m m e t r i c D i g i t a l S u b s c r i b e r L i n e の略である。

【 0 1 6 4 】

< 4 : まとめ >

最後に、本実施形態の技術的思想について簡単に纏める。以下に記載する技術的思想は、例えば、P C、携帯電話、ゲーム機、情報端末、情報家電、カーナビゲーションシステムなど、種々の情報処理装置に対して適用することができる。

【 0 1 6 5 】

上記の情報処理装置の機能構成は、以下のように表現することができる。例えば、下記 ( 1 ) に記載の情報処理装置は、関係性情報と区間メタデータとに基づき、ユーザが選択した登場物と関係性のある登場物が登場する区間を動画中から検出し、その区間に対応する画像を出力する機能を有している。つまり、この情報処理装置は、ユーザが選択した登場物に関係する登場物が登場するシーンを繋ぎ合わせてダイジェスト映像を出力したり、抽出したシーンに対応する画像を並べてダイジェストシーンの俯瞰表示を実現したりすることができる。もちろん、ユーザが選択する登場物が変われば、これらのダイジェスト映像や俯瞰表示の構成も、その登場物に適した構成に自動変更される。

【 0 1 6 6 】

上記のように、本実施形態に係る動画要約手法を用いると、人物や物品などの登場物を選択するだけで、その登場物に関するダイジェスト映像や俯瞰表示の内容を自動的に生成することが可能になる。例えば、運動会の模様を撮影した動画を用意し、自分の子供を選択すると、自分の子供に関するダイジェスト映像が自動的に生成される。このダイジェス

10

20

30

40

50

ト映像には、自分の子供はもちろんのこと、友人や先生など、自分の子供に関係する人物が登場するため、そのダイジェスト映像から人間関係を中心とした学校生活の雰囲気を把握することも可能になる。また、結婚式などのイベントで撮影した動画を用意し、新郎・新婦を選択すると、新郎・新婦を中心とし、友人や親しい親類などが登場する編集されたダイジェスト映像が自動生成される。

#### 【0167】

このように、本実施形態に係る動画要約手法を実現可能な構成を有する下記の情報処理装置を用いれば、専門家の手を借りずとも、人間関係を考慮した優れたダイジェスト映像を容易に生成することが可能になる。もちろん、ダイジェスト映像に代えて、各シーンに対応する画像を並べて表示する俯瞰表示なども実現することができる。また、本実施形態に係る動画要約手法は、人物の関係性だけでなく、人物と物品との関係性、或いは、物品と物品との関係性についても同様に扱うことができる。例えば、映画を収録した動画を用意し、登場するサングラスを選択すると、そのサングラスと関係性のある俳優の登場シーンをまとめたダイジェスト映像を生成することなども可能になる。このようなダイジェスト映像は、例えば、サングラスのコマーシャル映像として利用することもできよう。このように、本実施形態に係る動画要約手法は、応用範囲が広く様々な場面で利用できる。

#### 【0168】

##### (1)

動画に登場する各登場物間の関係性を示す関係性情報と、前記動画の中で各登場物が登場する区間を示す区間メタデータとを利用し、ユーザにより選択された登場物との間に所定の関係性を有する登場物が登場する区間を検出する区間検出部と、

前記区間検出部により検出された区間に対応する画像を出力する画像出力部と、  
を備える、

情報処理装置。

#### 【0169】

##### (2)

前記関係性情報は、前記関係性の強さを示す関係性値を含み、

前記区間検出部は、前記ユーザにより選択された登場物との間における関係性値の総和が最大となるように所定数の登場物を選択し、選択した登場物が登場する所定数の区間を検出する、

上記(1)に記載の情報処理装置。

#### 【0170】

##### (3)

前記関係性情報は、前記関係性の強さを示す関係性値を含み、

前記区間検出部は、

前記ユーザにより選択された第1の登場物が登場する第1の区間を検出する処理と、

$n = 2 \sim N$  ( $N \geq 2$ ) に関し、第  $n - 1$  の区間の次に再生される第  $n$  の区間に登場する第  $n$  の登場物を選択する処理と、

前記第  $n - 1$  の登場物と前記第  $n$  の登場物との間における関係性値を  $n = 2 \sim N$  に関し  
て積算し、前記関係性値の積算値を算出する処理と、

前記関係性値の積算値が最大となる第2～第  $N$  の登場物を選択する処理と、

選択した前記第  $n$  の登場物 ( $n = 2 \sim N$ ) が登場する第  $n$  の区間を検出する処理と、  
を実行し、前記第1～第  $N$  の区間を検出結果として出力する、

上記(1)に記載の情報処理装置。

#### 【0171】

##### (4)

前記関係性情報は、前記関係性の強さを示す関係性値を含み、

前記区間検出部は、

前記ユーザにより選択された登場物が登場する第1の区間を検出する処理と、

$n = 2 \sim N$  ( $N \geq 2$ ) に関し、第  $n - 1$  の区間以降の各区間について、各区間内に登場

する全ての登場物と、前記第  $n - 1$  の区間に登場する全ての登場物との間における関係性値の総和が最大になる区間を検出して第  $n$  の区間に設定する処理と、  
を実行し、前記第 1 ～ 第  $N$  の区間を検出結果として出力する、

上記 ( 1 ) に記載の情報処理装置。

【 0 1 7 2 】

( 5 )

前記関係性情報は、前記関係性の強さを示す関係性値を含み、

前記区間検出部は、

前記ユーザにより選択された登場物が登場する第 1 の区間を検出する処理と、

$n = 2 \sim N$  (  $N \geq 2$  ) に関し、第  $n - 1$  の区間以降の各区間について、各区間内に登場する全ての登場物と、前記ユーザにより選択された登場物との間における関係性値の総和を算出し、算出した関係性値の総和が前記第  $n - 1$  の区間に関する関係性値の総和に最も近くなる区間を検出して第  $n$  の区間に設定する処理と、

を実行し、前記第 1 ～ 第  $N$  の区間を検出結果として出力する、

上記 ( 1 ) に記載の情報処理装置。

【 0 1 7 3 】

( 6 )

前記関係性情報は、前記関係性の強さを示す関係性値を含み、

前記区間検出部は、前記ユーザにより選択された登場物との間における関係性値の総和が最小となるように所定数の登場物を選択し、選択した登場物が登場する所定数の区間を検出する、

上記 ( 1 ) に記載の情報処理装置。

【 0 1 7 4 】

( 7 )

前記区間検出部は、前記関係性情報に含まれる少なくとも 1 つの関係性値が変更された場合に、変更後の前記関係性情報に基づいて前記区間を再検出し、

前記画像出力部は、前記区間検出部による再検出の結果に応じて出力内容を更新する、

上記 ( 2 ) ～ ( 6 ) のいずれか 1 項に記載の情報処理装置。

【 0 1 7 5 】

( 8 )

前記ユーザにより選択された登場物に対応する第 1 のオブジェクトを操作領域の中心に表示し、当該登場物との間における関係性値の大きさに応じた距離だけ前記第 1 のオブジェクトから離れた位置に、一又は複数の他の登場物にそれぞれ対応する一又は複数の第 2 のオブジェクトを表示するオブジェクト表示部と、

前記操作領域において前記第 2 のオブジェクトの位置が変更された場合に、変更後の当該第 2 のオブジェクトの位置に応じた関係性値を算出し、算出した関係性値により前記関係性情報を更新する関係性情報更新部と、

をさらに備える、

上記 ( 7 ) に記載の情報処理装置。

【 0 1 7 6 】

( 9 )

同じ作品に登場する頻度が高いほど、或いは、共に登場する区間の長さが長いほど、前記関係性値が大きくなるように、前記区間メタデータを利用して各登場物の組に関する関係性値を算出して前記関係性情報を生成する関係性情報生成部をさらに備える、

上記 ( 2 ) ～ ( 8 ) に記載の情報処理装置。

【 0 1 7 7 】

( 1 0 )

前記関係性情報は、前記関係性の種類を示した第 1 ～ 第  $N - 1$  の種類情報を含み、

前記区間検出部は、

前記ユーザにより選択された登場物が登場する第 1 の区間を検出する処理と、

$n = 1 \sim N - 1$  に関し、前記ユーザにより選択された登場物との間において第  $n$  の種類情報が示す種類の関係性を有する第  $n$  の登場物を検出し、第  $n$  の区間以降の区間の中から当該第  $n$  の登場物が登場する区間を検出して第  $n + 1$  の区間に設定する処理と、

を実行し、前記第 1 ～ 第  $N$  の区間を検出結果として出力する、

上記 ( 1 ) に記載の情報処理装置。

【 0 1 7 8 】

( 1 1 )

予め用意された複数の動画における特徴的な関係性の時系列変化を検出し、検出した特徴的な関係性の時系列変化に適合するように前記第 1 ～ 第  $N - 1$  の種類情報を設定する種類情報設定部をさらに備える、

上記 ( 1 0 ) に記載の情報処理装置。

【 0 1 7 9 】

( 1 2 )

前記第 1 ～ 第  $N - 1$  の種類情報には、それぞれカメラワーク又はエフェクトの種類を示す補助情報が対応付けられており、

前記区間検出部は、

前記ユーザにより選択された登場物が登場する第 1 の区間を検出する処理と、

$n = 1 \sim N - 1$  に関し、前記ユーザにより選択された登場物との間において第  $n$  の種類情報が示す種類の関係性を有する第  $n$  の登場物を検出し、第  $n$  の区間以降の区間の中から当該第  $n$  の登場物が登場する区間であり、かつ、前記補助情報に適合する区間を検出して

第  $n + 1$  の区間に設定する処理と、

を実行し、前記第 1 ～ 第  $N$  の区間を検出結果として出力する、

上記 ( 1 0 ) 又は ( 1 1 ) に記載の情報処理装置。

【 0 1 8 0 】

( 1 3 )

動画に登場する各登場物間の関係性を示す関係性情報と、前記動画の中で各登場物が登場する区間を示す区間メタデータとを利用し、ユーザにより選択された登場物との間に所定の関係性を有する登場物が登場する区間を検出するステップと、

検出された区間に対応する画像を出力するステップと、

を含む、

動画要約方法。

【 0 1 8 1 】

( 1 4 )

動画に登場する各登場物間の関係性を示す関係性情報と、前記動画の中で各登場物が登場する区間を示す区間メタデータとを利用し、ユーザにより選択された登場物との間に所定の関係性を有する登場物が登場する区間を検出する区間検出機能と、

前記区間検出機能により検出された区間に対応する画像を出力する画像出力機能と、をコンピュータに実現させるためのプログラム。

【 0 1 8 2 】

( 備考 )

上記のシーン抽出部 1 0 4 は、区間検出部の一例である。上記の出力部 1 0 5 は、画像出力部の一例である。上記の UI 表示部 1 0 6 は、オブジェクト表示部の一例である。上記の関係性情報生成部 1 0 2 は、関係性情報更新部の一例である。上記の動画解析部 1 1 1 は、種類情報設定部の一例である。

【 0 1 8 3 】

以上、添付図面を参照しながら本技術に係る好適な実施形態について説明したが、本技術はここで開示した構成例に限定されないことは言うまでもない。当業者であれば、特許請求の範囲に記載された範疇内において、各種の変更例又は修正例に想到し得ることは明らかであり、それらについても当然に本技術の技術的範囲に属するものと了解される。

【 符号の説明 】

10

20

30

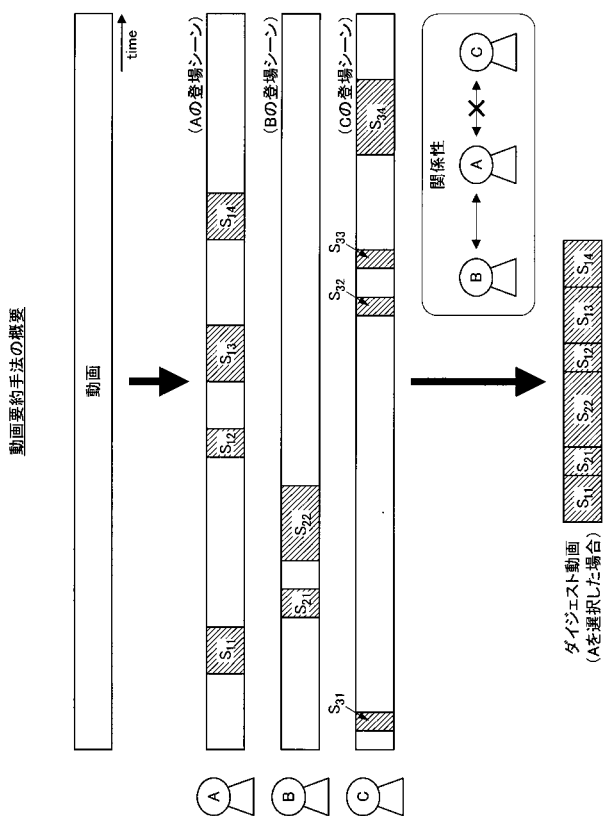
40

50

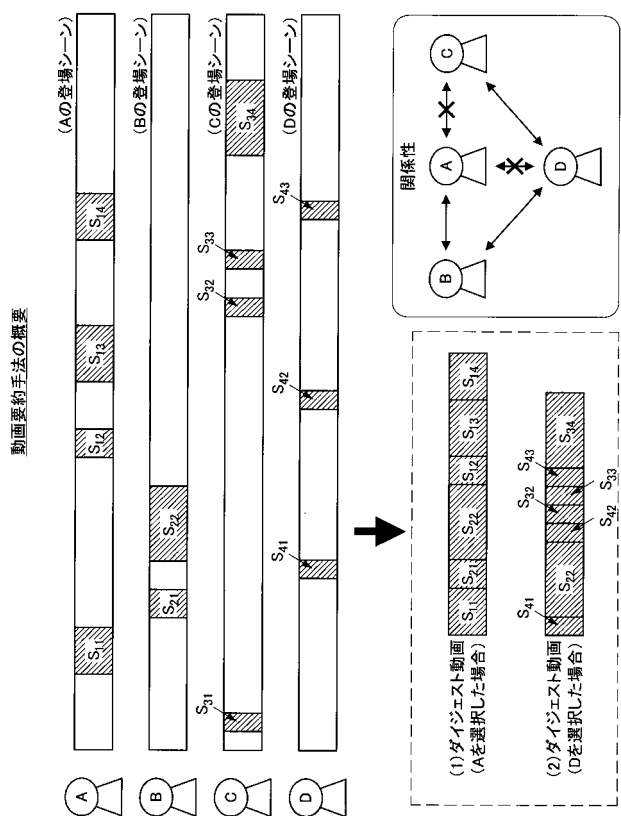
## 【 0 1 8 4 】

1 0 0	情報処理装置
1 0 1	メタデータ取得部
1 0 2	関係性情報生成部
1 0 3	入力部
1 0 4	シーン抽出部
1 0 5	出力部
1 0 6	UI表示部
1 0 7	動画取得部
1 1 1	動画解析部
1 2 1	関係性情報取得部

【 図 1 】



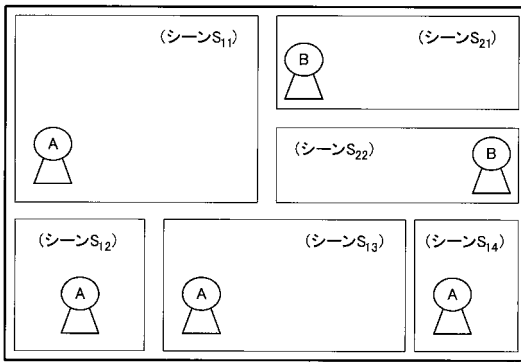
【 図 2 】





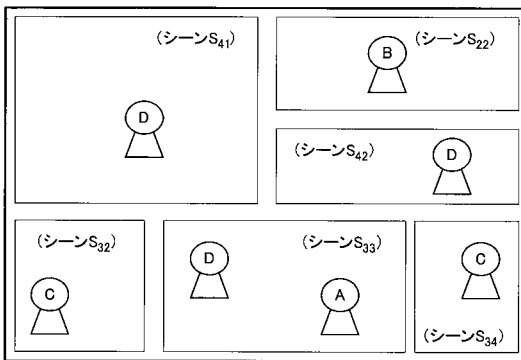
【図 3】

コミック表示 (Aを選択した場合)



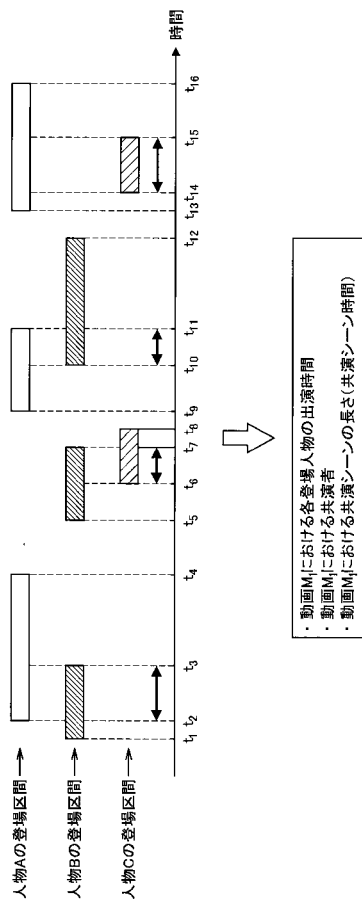
【図 4】

コミック表示 (Dを選択した場合)



【図 6】

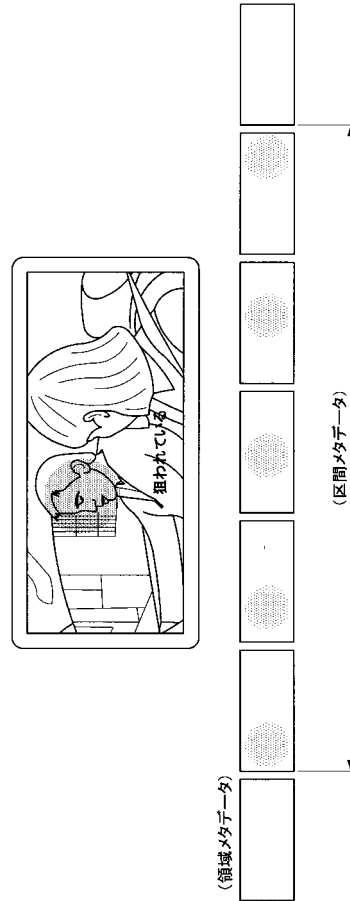
動画M<sub>i</sub>に関する区間メタデータの構成例  
(区間メタデータから得られる情報について)



【図 5】

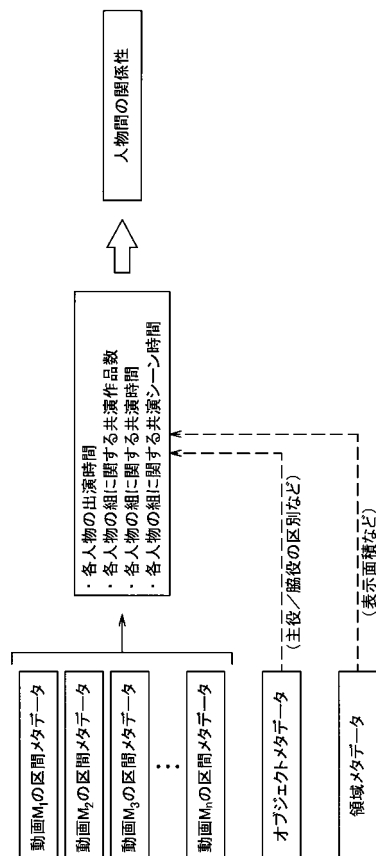
[動画タイムラインメタデータの概要]

- ・領域メタデータ: 動画フレーム内に登場する人物や物品などの位置 (領域)
- ・区間メタデータ: 人物や物品などが登場する区間
- ・オブジェクトメタデータ: 登場する人物や物品などの関連情報



【図 7】

関係性の評価方法



【図 8】

関係性マトリックスの構成

	人物A	人物B	人物C
人物A		人物A,Bの 関連性値	人物A,Cの 関連性値
人物B	人物A,Bの 関連性値		人物B,Cの 関連性値
人物C	人物A,Cの 関連性値	人物B,Cの 関連性値	

【図 9】

共演作品数に基づく関係性マトリックスの計算方法

(共演関係)

- ・ 動画M<sub>1</sub>: A, B, C
- ・ 動画M<sub>2</sub>: A, C
- ・ 動画M<sub>3</sub>: B
- ・ 動画M<sub>4</sub>: A
- ・ 動画M<sub>5</sub>: A, B
- ・ 動画M<sub>6</sub>: A, B



(関係性値の算出結果: 共演作品数)

	人物A	人物B	人物C
人物A	(5)	3	2
人物B	3	(4)	1
人物C	2	1	(2)

【図 10】

共演作品数に基づく関係性マトリックスの計算方法(重み付き)

(共演関係)

- ・ 動画M<sub>1</sub>: A(主役), B(主役), C(脇役)
- ・ 動画M<sub>2</sub>: A(主役), C(主役)
- ・ 動画M<sub>3</sub>: B(主役)
- ・ 動画M<sub>4</sub>: A(脇役)
- ・ 動画M<sub>5</sub>: A(脇役), B(脇役)
- ・ 動画M<sub>6</sub>: A(主役), B(脇役)



- ・ X(主役) & Y(主役)の動画→スコア=4
- ・ X(主役) & Y(脇役)の動画→スコア=2
- ・ X(脇役) & Y(脇役)の動画→スコア=1

(関係性値の算出結果: スコアの合計値)

	人物A	人物B	人物C
人物A	(14)	7	6
人物B	7	(10)	2
人物C	6	2	(5)

【図 11】

出演時間に基づく関係性マトリックスの計算方法

(共演関係: 括弧内は出演時間)

- ・ 動画M<sub>1</sub>: A(40), B(30), C(10)
- ・ 動画M<sub>2</sub>: A(30), C(40)
- ・ 動画M<sub>3</sub>: B(20)
- ・ 動画M<sub>4</sub>: A(10)
- ・ 動画M<sub>5</sub>: A(5), B(10)
- ・ 動画M<sub>6</sub>: A(40), B(5)



・ X(t1) & Y(t2)の動画→スコア=t1 × t2

(関係性値の算出結果: スコアの合計値)

	人物A	人物B	人物C
人物A	(2665)	1450	1600
人物B	1450	(1425)	300
人物C	1600	300	(1700)

【図 12】

共演シーン時間に基づく関係性マトリックスの計算方法

(共演関係: X,Yの共演シーン時間を[X,Y]と表記)

- ・ 動画M<sub>1</sub>: [A,A]=40, [B,B]=30, [C,C]=10, [A,B]=20, [A,C]=5, [B,C]=5
- ・ 動画M<sub>2</sub>: [A,A]=30, [C,C]=40, [A,C]=20
- ・ 動画M<sub>3</sub>: [B,B]=20
- ・ 動画M<sub>4</sub>: [A,A]=10
- ・ 動画M<sub>5</sub>: [A,A]=5, [B,B]=10, [A,B]=5
- ・ 動画M<sub>6</sub>: [A,A]=40, [B,B]=5, [A,B]=5



(関係性値の算出結果: 共演シーン時間の合計値)

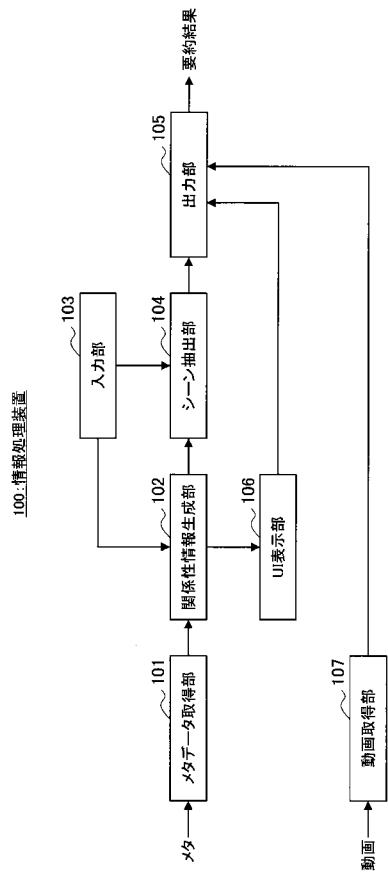
	人物A	人物B	人物C
人物A	(125)	30	25
人物B	30	(45)	5
人物C	25	5	(50)

【図 13】

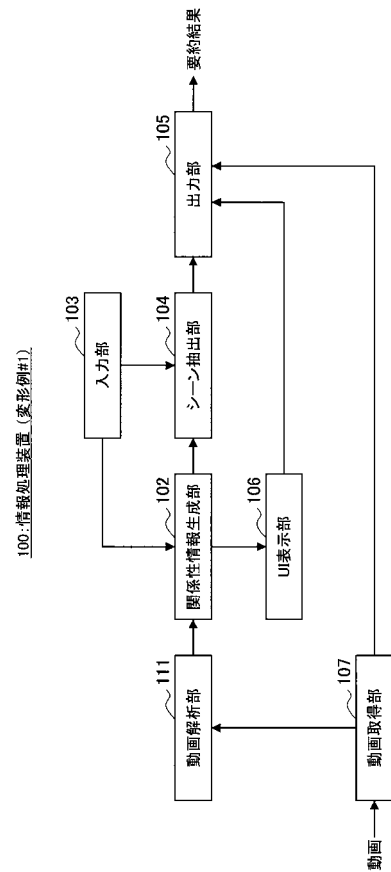
関係性マトリックスの例(非対象、属性付き)

	人物A	人物B	人物C	人物D
人物A	本人(1.0)	家族(0.9)	友人(0.8)	他人(0.2)
人物B	家族(0.9)	本人(1.0)	他人(0.1)	恋人(0.9)
人物C	敵(0.0)	他人(0.2)	本人(1.0)	家族(0.9)
人物D	敵(0.0)	友人(0.7)	家族(0.9)	本人(1.0)

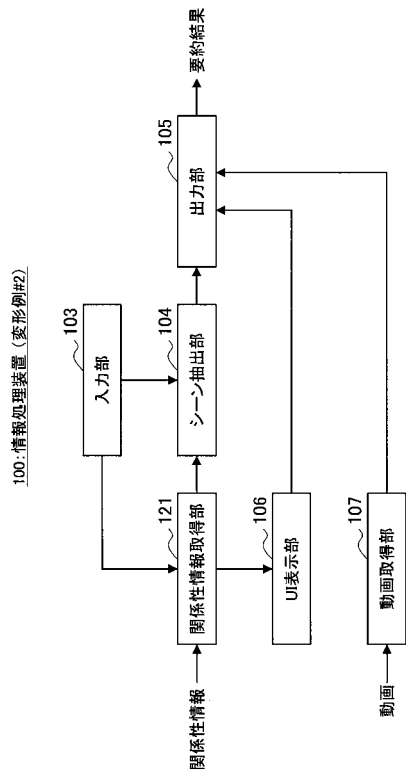
【図 1 4】



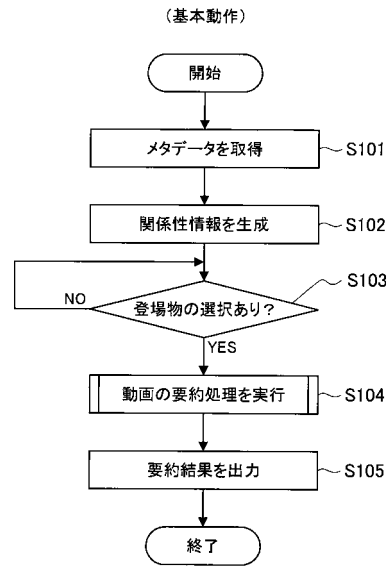
【図 1 5】



【図 1 6】

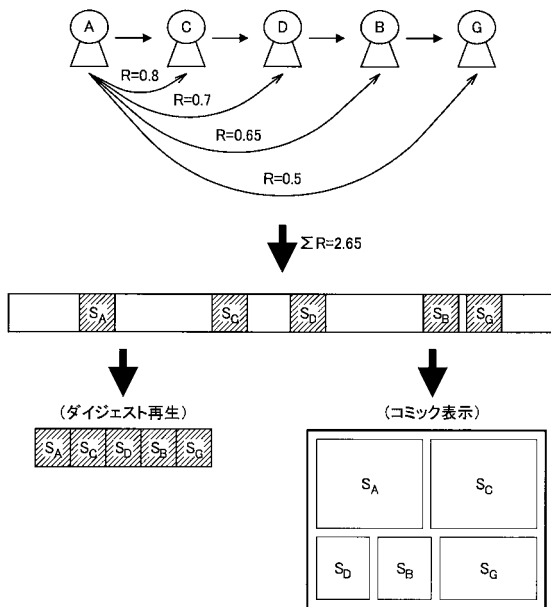


【図 1 7】



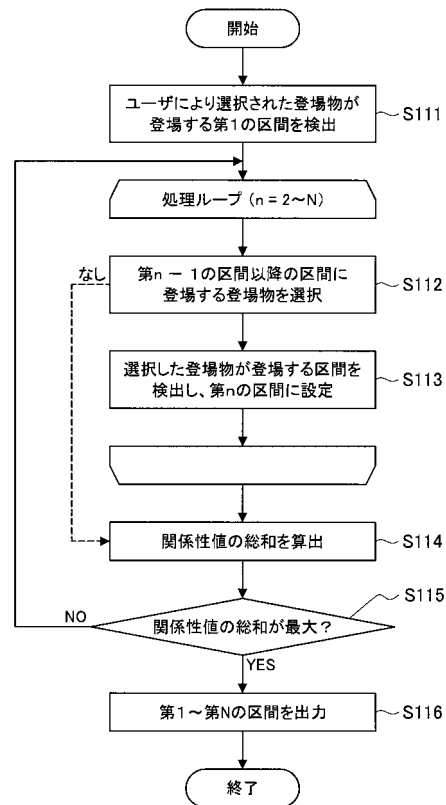
【図 18】

(要約処理#1の概要)



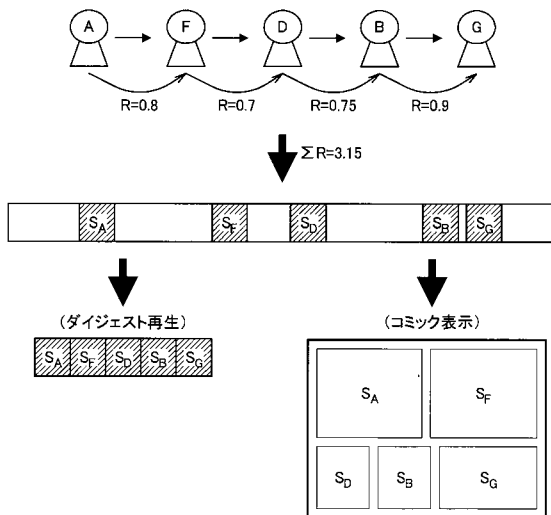
【図 19】

(要約処理#1)



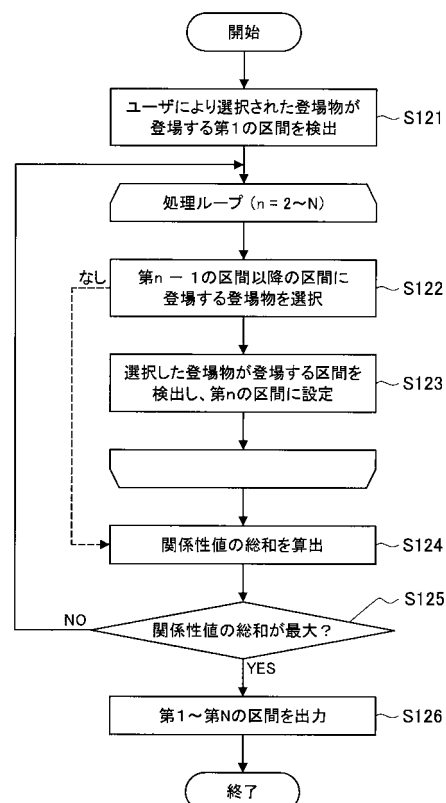
【図 20】

(要約処理#2の概要)

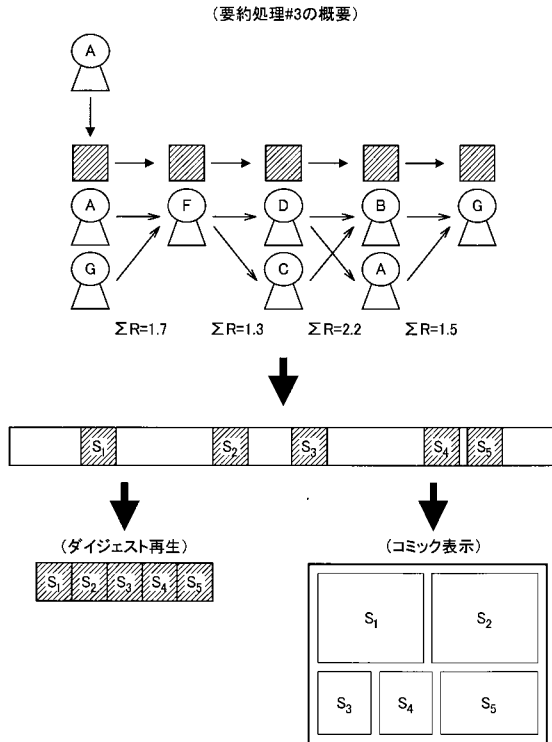


【図 21】

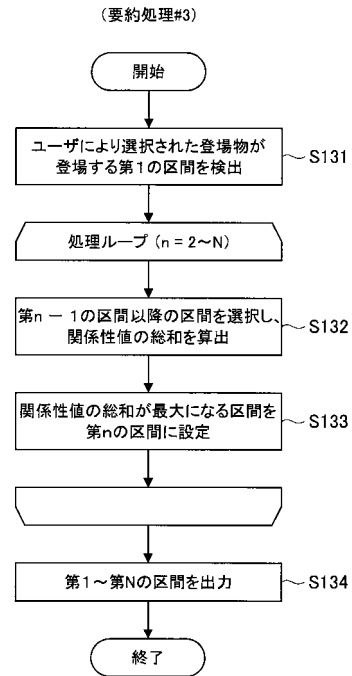
(要約処理#2)



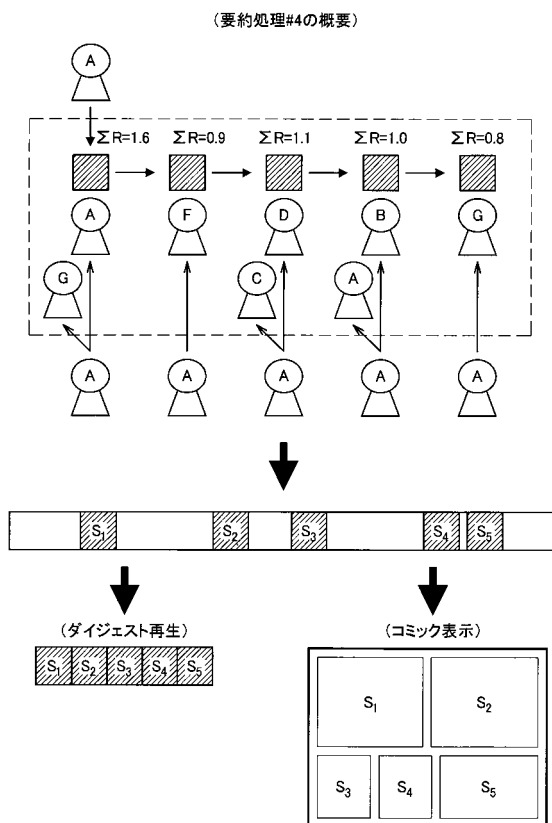
【図 2 2】



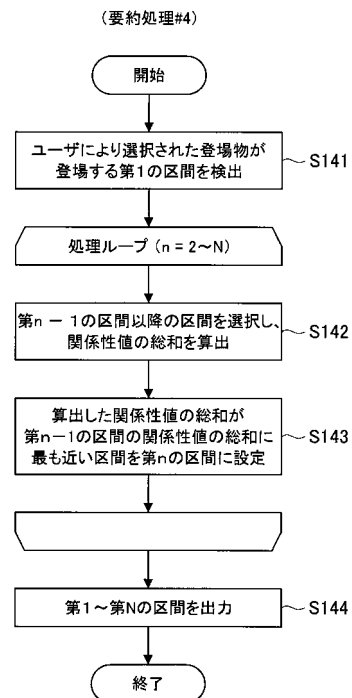
【図 2 3】



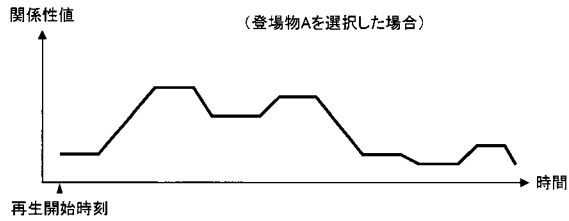
【図 2 4】



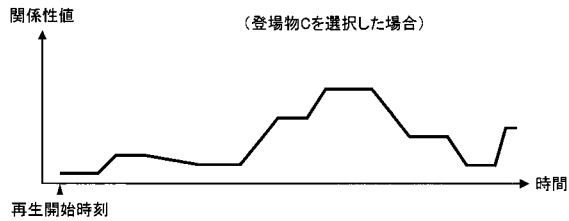
【図 2 5】



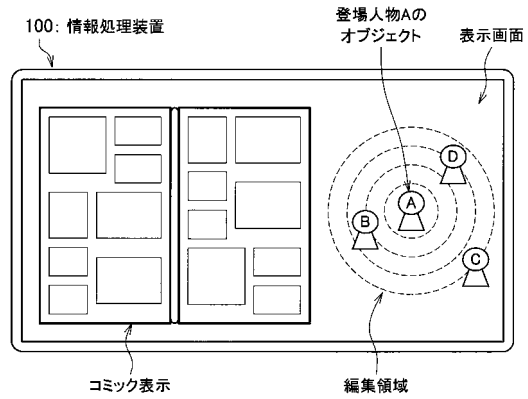
【図 26】



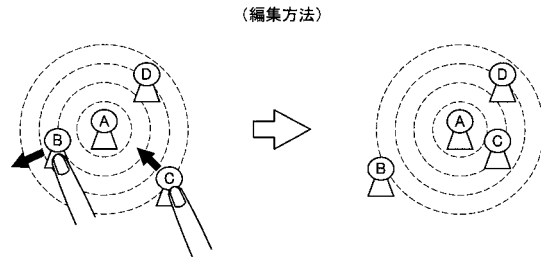
【図 27】



【図 28】



【図 29】



【図 30】

関係性マトリックス(編集前)

	人物A	人物B	人物C	人物D
人物A	1.0	0.8	0.2	0.5
人物B	0.8	1.0	0.1	0.0
人物C	0.2	0.1	1.0	0.1
人物D	0.5	0.0	0.1	1.0



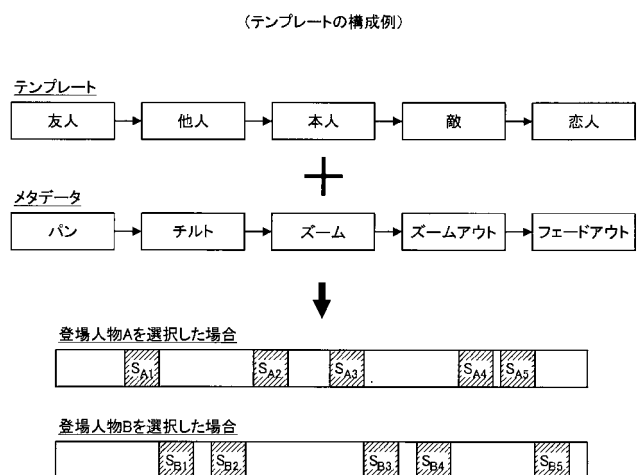
関係性マトリックス(編集後)

	人物A	人物B	人物C	人物D
人物A	1.0	0.1	0.7	0.5
人物B	0.1	1.0	0.1	0.0
人物C	0.7	0.3	1.0	0.1
人物D	0.5	0.0	0.3	1.0

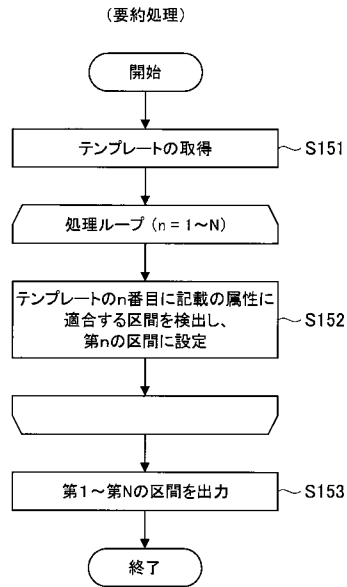


□要約結果の動的な変化  
 → ダイジェストの構成を変更  
 → 強調される登場人物が人物Bから人物Cへ

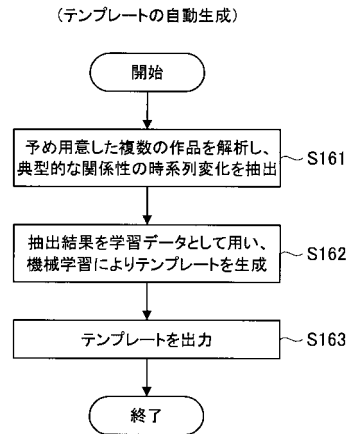
【図 31】



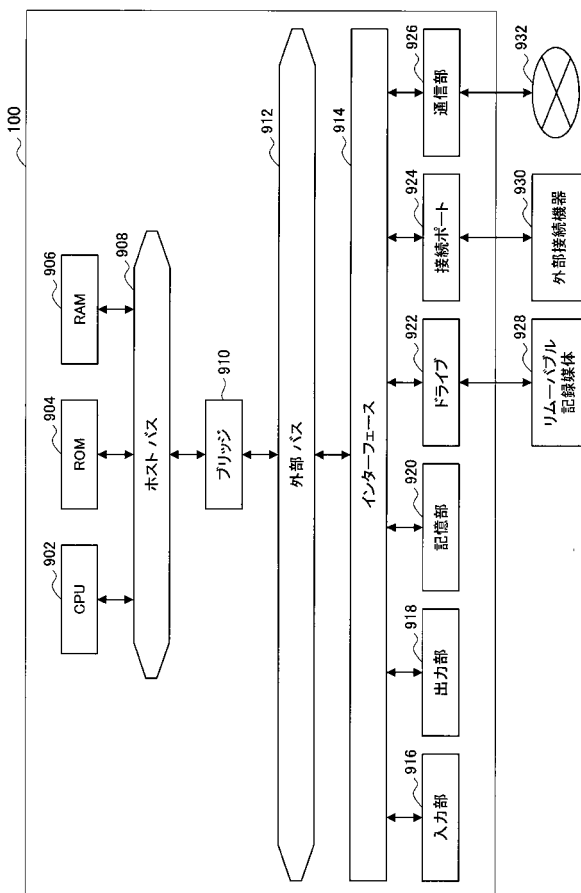
【 図 3 2 】



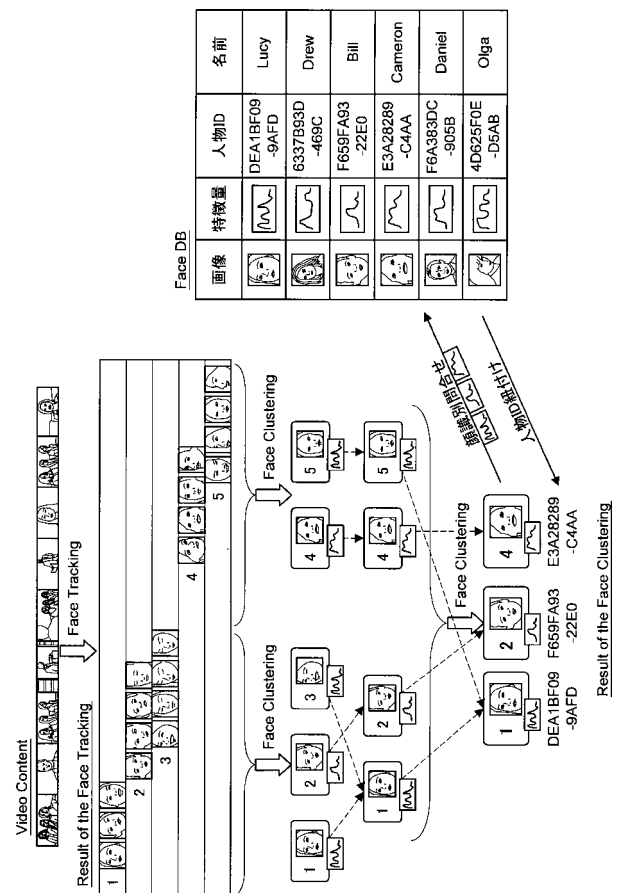
【 図 3 3 】



【 図 3 4 】



【 図 3 5 】



---

フロントページの続き

- (72)発明者 倉田 雅友  
東京都港区港南 1 丁目 7 番 1 号 ソニー株式会社内
- (72)発明者 村田 誠  
東京都港区港南 1 丁目 7 番 1 号 ソニー株式会社内
- (72)発明者 澁谷 直樹  
東京都港区港南 1 丁目 7 番 1 号 ソニー株式会社内