

ビデオ・データベースにおけるトップダウン処理に基づく 効率的な類似シーン検索に関する考察

牛尼 剛聡[†] 小坂 健[‡] 渡邊 豊英[‡]

[†]九州芸術工科大学芸術工学部 芸術情報設計学科

〒 815-8540 福岡市南区塩原 4-9-1

[‡]名古屋大学大学院 工学研究科 情報工学専攻

〒 464-8603 名古屋市千種区不老町

ushiana@kyushu-id.ac.jp {kosaka,watanabe}@watanabe.nuie.nagoya-u.ac.jp

本稿では、イベント系列としてビデオの内容が記述されたビデオ・データベースにおいて効率的な類似シーン検索を実現する手法について考察する。従来、記号系列の類似度を評価するために DP マッチングなどの手法が利用されてきたが、個々の要素の対応づけをボトムアップに処理して最適なマッチングを見つけるために処理コストが大きいためシーン検索に用いるのは現実的でない。一般にデータベースの類似検索では、与えられた例との類似度の高いオブジェクトを発見することが目的であり、全てのオブジェクトに対して正確な類似度を計算する必要はない。我々は、利用者が与えた例を抽象化して概念階層を自動生成し、上位の概念からトップダウンに処理を行い、類似度が低いオブジェクトに対する詳細な計算を省略することにより、効率的に類似したシーンを検索するための展望を述べる。

キーワード：ビデオ・データベース，シーン検索，類似検索，抽象化，is-a 階層

A Study for Top-down Method for Efficient Similarity Search of Scenes on Video Databases

Taketoshi USHIAMA[†], Ken KOSAKA[‡] and Toyohide WATANABE[‡]

[†]Department of Art and Information Design, Kyushu Institute of Design

4-9-1 Shiobaru, Minami-ku, Fukuoka, 815-8540 Japan

[‡]Department of Information Engineering, Graduate School of Engineering, Nagoya University

Furo-cho, Chikusa-ku, Nagoya, 464-8603 Japan

ushiana@kyushu-id.ac.jp {kosaka,watanabe}@watanabe.nuie.nagoya-u.ac.jp

In this paper, we introduce an approach for efficient Searching similar scene on the video databases whose contents are described with sequence of events. Up to now, similarity between two sequences of events is estimated with bottom-up approach such as DP-matching method, but this approach is too costly for applying on the scene retrievals on video databases. In our approach, estimation of similarity is processed using an automatically generated is-a hierarchy in a top-down manner.

keyword: video database, scene retrieval, similarity, abstraction, is-a hierarchy

1 はじめに

ビデオ・データは対象世界の動的な活動を人間にとって直感的に分かりやすく表現可能であるため、報道、娯楽、教育等様々な分野で対象世界の動的な側面の記録・伝達に利用されている。近年、大量のビデオを計算機上で効率的に管理し、効果的に利用するためにビデオ・データベースに関する研究が活発に行われている。ビデオは静止画像の時系列であり、その部分系列を考えることができるが、意味的なまとまりがある部分系列はシーンと呼ばれる。ビデオ・データを利用するには特定のシーンのみを必要とする場合があるため、ビデオ・データベースから特定のシーンを検索する機構が重要である。

これまでに、シーン検索に関して多くの研究がなされている。従来の一般的なシーン検索手法では、利用者がキーボードから文字列を入力したり、GUI上のメニューを選択することにより検索条件を指定する必要がある。しかし、一般の利用者が検索条件を正確に記述することは困難である。また、検索条件の指定に熟練した利用者であっても、検索条件を正確に指定するためには多大な労力が必要である。さらに、利用者がビデオ・データベース中にどのようなデータが含まれているかに関して具体的な知識を持たない場合には、考えられる検索条件の数が膨大になり、利用者は要求する検索結果が得られるまで、様々な検索条件で検索を繰り返す必要がある。また、利用者が予想しなかったデータがデータベース中に含まれており、それが利用者の検索要求と一致する場合も考えられる。

今後、インターネットやケーブルテレビ回線を通じて一般の利用者がビデオ・データベースにアクセスする環境が整備されていくと予想される。こうした環境では、より簡単な操作でシーンを検索するための機構が必要である。こうした背景の下、我々はデータベースをブラウジングして利用者の検索要求に類似したシーンを指定し、指定されたシーンを例として類似したシーンを検索するシステムの開発を行っている。我々はこれまでに概略表現の生成に関する手法 [1] を提案しており、本稿では、利用者が例となるシーンが発見できたことを前提として、利用者が例として指定したシーンに類似するシーンをビデオ・データベースから検索する手法を提案する。

我々は利用者の視点に基づいて多種多様なシーンを検索するための内容記述・検索モデルとしてイベント-アクティビティ・モデルを採用している [2]。このモデルでは対象世界の動的な事象を継続時間のないイベントと継続時

間のあるアクティビティに分類し、アクティビティをイベントの系列として表現する。ビデオ中のシーンの内容はアクティビティに対応すると考えられるため、フレームにイベントを対応づけておくことにより、系列パターンのマッチングにより部分系列を同定することでシーン検索を実現する。このアプローチでは検索結果となるシーンを動的に構成することができるため、スポーツ中継などのように発生するイベントの種類が限定可能であり、利用者の視点によって様々なアクティビティが想定可能な対象に対して、シーン単位のインデキシングよりも利用者の視点を適切に反映可能である。本稿ではイベント-アクティビティ・モデルに基づいて内容が記述されたビデオ・データベースを対象とする。

イベント-アクティビティ・モデルでは、ビデオの内容はイベント系列として表現されるため、類似検索のためには系列データの類似性を評価する必要がある。系列データの類似性の評価のために、系列を構成する要素間の類似性に基づいて系列の全体の類似性を評価する必要がある。しかし、系列パターンのマッチングは比較対象となるパターンの長さが一致しない場合には要素間の対応関係を一意に決定することができない。この問題に対処するために、系列パターンのマッチングには DP マッチング [3] を用いる手法が提案されている。DP マッチングではパターンを非線型に圧縮し最適なマッチングを行う手法であり、音声処理、特に単語認識において広く利用されている。音声認識以外にもドラマ映像・音声・シナリオ文書の対応づけに用いる手法 [4] 等様々な対象に対して適応されている。イベント系列の類似性の判断のために DP マッチングを適用することは可能である。しかし、DP マッチングは対象となる系列の最適なマッチングをボトムアップに処理していくため、ビデオ・データベース上のシーン検索のようにマッチングの対象が膨大である場合には処理コストが大きくなり現実的でない。

大量のデータを対象とするデータベースの類似検索においては、利用者は類似度が高いオブジェクトに対して興味がある。したがって、類似度が高いオブジェクトに関しては詳細な順序付けが必要となるが、類似度が低いオブジェクトに対しては詳細な順序付けは不必要である。本稿で提案する手法では、例示によって与えられる条件を抽象化し、抽象化された条件のマッチングによってオブジェクトの大局的な特徴をチェックする。このチェックによって、高い類似度が期待されるオブジェクト集合を限定し、徐々

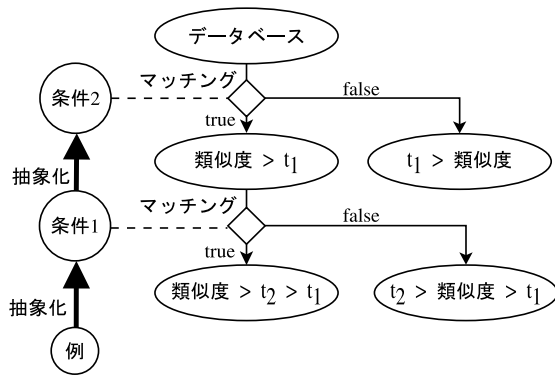


図 1: 提案手法の概念図

に詳細な（抽象度が低い）条件とのマッチングを行い正確な類似度を計算する．本手法の概念図を図 1 に示す．また，本手法は，あらかじめ用意された概念階層の類似性に関する知識を利用せず，データベース内に格納されている統計的な情報を利用して類似性を評価する．したがって，データベース内に格納されているデータの種類の依存しない一般的な類似検索手法である．

2 関連する研究

QBIC では，シーンから代表フレームと呼ばれるパノラマ画像を作成し，静止画像レベルで類似検索を行う手法を採用している [5]．これらはビデオの持つ時間的な特徴が空間的な特徴として表現されてしまうため，時間的な特徴を正確に評価することが困難である．長坂ら [6] はフレーム系列の特徴をランレングス符号化することにより，映像中の類似区間を実時間で検索する手法を提案している．この手法は各フレームの表装的な特徴の類似性を対象としており，意味内容の類似性は考慮していない．

3 アプローチ

3.1 概念の is-a 階層と類似性

類似検索をするためには類似性の定義を明確化する必要がある．我々は概念の is-a 階層に基づいて類似性を定義する．

概念とは，対象となるオブジェクト集合上で意味的なまとまりを持つ部分集合を規定する基準である．概念によって規定された部分集合は外延 (extention) と呼ばれる．また，外延を持つ性質は内包 (connotation) と呼ばれる．内包は外延の要素となるオブジェクトを規定する条件の集合であると考えられることができる．

いま，概念 C の外延を $\text{ext}(C)$ と表記し，内包を $\text{con}(C)$ と表記する．対象とするオブジェクト集合を \mathbf{O} とするとき， $\text{ext}(C) \subseteq \mathbf{O}$ を満足する．また，オブジェ

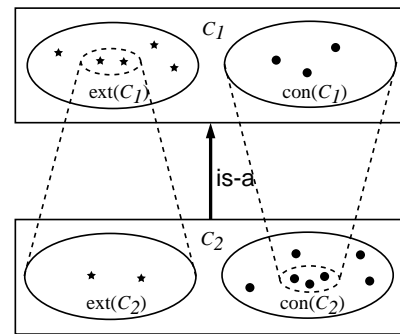


図 2: is-a 関連の例

クトに対して適用し得る条件の集合を Θ とするとき， $\text{con}(C) \subseteq \Theta$ を満足する．これらの表記に基づいて概念集合 \mathbf{C} 上の is-a 関係を以下のように定義する．

【定義:is-a 関係】 $C_1, C_2 \in \mathbf{C}$ に対して以下の二つの条件を共に満足するとき， C_2 is-a C_1 と表記する．

1. $\text{ext}(C_2) \subseteq \text{ext}(C_1)$
2. $\text{con}(C_1) \subseteq \text{con}(C_2)$

図 2 に is-a 関係を持つ概念の例を示す．星印がオブジェクトを表し，黒点が条件を表す．オブジェクトを囲む直線の楕円は外延を表し，条件を囲む楕円は内包を表している．破線の楕円は外延または内包の部分集合を表している． C_2 is-a C_1 であるとき， C_1 を C_2 の上位概念と呼び， C_2 を C_1 の下位概念と呼ぶ．下位概念の外延は上位概念の外延の部分集合である．また，下位概念の内包は上位概念の内包と新たな条件集合との集合和をとったものである．定義より，is-a 関係は順序関係を満足し，同一の上位概念を持つ概念の集合を概念階層と呼ぶ．いま， $C_0 \in \mathbf{C}$ が $\text{ext}(C_0) = \mathbf{O}$ かつ $\text{con}(C_0) = \emptyset$ を満たすとする． C_0 は対象とするオブジェクト集合全体を表す概念である．このとき，概念集合 \mathbf{C} は C_0 を最上位概念とする単一の概念階層となる．下位概念を持たない概念を原子概念と呼ぶ．

2つのオブジェクトが類似していると考えられるとき，類似しているという観点から，それらは一つの意味的なまとまりをもつ部分集合の要素であると考えられることができる．したがって，それらのオブジェクトを含む概念を想定することができる．概念階層に基づいて類似関係を定義する．

【定義:類似関係】概念階層 $(\mathbf{C}, \text{is-a})$ において，ある原子概念 C_n から is-a 関係をたどって最上位階層 C_0 まで到達する分岐のない経路を考える．この経路が通過する概念の要素から構成され， $C_i \text{ is-a } C_{i-1} (1 \leq i \leq n)$ を満足する系列 C_n, C_{n-1}, \dots, C_0 を概念系列と呼ぶ．概念系列における

C_i の外延のうち、いずれの下位概念の外延にもなり得ないオブジェクトの集合 E_i とする．形式的には E_i は次のように再帰的に定義できる．

1. $E_n = \text{ext}(C_n)$
2. $E_i = \text{ext}(C_i) - \bigcup_{k=i+1}^n E_k$

ここで、 E_0, \dots, E_n は対象とするオブジェクト集合の排他分割になっている． E_i の要素を概念 C_i のインスタンスと呼ぶ．このとき、オブジェクト間の類似関係を以下の条件を満足するオブジェクト集合上の全順序関係 \preceq_{sim} として定義する．

$i \leq j$ を満足する場合、 $\forall o_1 \in \text{ext}(C_i), \forall o_2 \in \text{ext}(C_j)$ に対して $o_1 \preceq_{\text{sim}} o_2$ である．

ここで定義した類似関係は概念階層上の一つの経路に対して定義される．同一の概念階層であっても、経路の選び方が異なればオブジェクト間の類似関係は異なる．

類似関係に基づいて、原始概念の外延に対する類似の程度を比較することができる．たとえば、オブジェクト o_2 と o_3 に対して、 o_1 に類似する程度を比較することを考える．いま、 $o_3 \preceq_{\text{sim}} o_2 \preceq_{\text{sim}} o_1$ であるとすると、 o_2 の方が o_3 よりも o_1 に類似していると判断できる．

3.2 類似検索と処理コスト

類似検索では、指定されたオブジェクトに対してデータベース中の類似性が高いオブジェクトを検索結果とする．このとき、検索結果は類似の程度によって順序付けされることが期待される．概念の is-a 階層に基づいて類似検索を実現するためには、is-a 階層上で先頭要素が例示したオブジェクトをインスタンスとする原子概念となっている概念系列を決定する必要がある．検索処理は、データベース内のすべてのオブジェクトを概念系列を構成する概念のインスタンスとして割り当てることによって行われる．

いま、概念を規定する任意の条件（特性） $\theta \in \Theta$ に対して、オブジェクトがその条件を満足するかを判断するための処理コストが同一であると仮定する．このとき、例に対する類似度が高いインスタンスをもつ概念は、その上位概念と比較して内包の要素数が多いため、条件を処理するコストが大きくなる．

is-a 階層においては下位概念の外延は上位概念の外延の部分集合である．したがって、上位概念の外延が決定されている場合には、下位概念の外延を決定する際に必要な条件とのマッチングは上位概念の外延に対して実行すれば十分である．さらに、上位概念の内包は下位概念の内包の部分集合であることから、下位概念の外延を決定するために

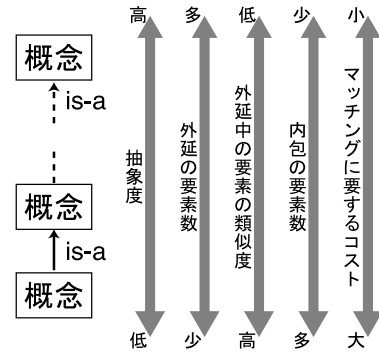


図 3: 概念階層の特徴

は、下位概念によって新たに付加された条件のマッチングを行えば十分である．これらのことから、概念系列上の上位概念から下位概念の順序でマッチングを進めていくことにより、効率的に概念のインスタンスを決定することができる．概念階層の特徴を表す概念図を図 3 に示す．

3.3 概念系列に基づく効率的な類似検索

本稿で提案する類似検索では利用者が提示した例にしたがって動的に概念系列を生成する．与えられた例に対して、概念系列は多数存在し、類似関係は概念系列に基づいて定義されるため、適切な概念系列を生成することが必要である．

いま、対象とするオブジェクト集合間で計算可能な類似度が定義されているとする．このとき、概念系列を用いて類似検索を行うためには、概念のインスタンスに対して、計算される類似度が概念系列において定義される類似関係を満足しなければならない．また、概念系列に基づいて類似性を判断する場合には、上位概念における条件を満足しなかったオブジェクトに対して、下位概念における条件のチェックは行われない．したがって、上位概念における条件を満足しなかったオブジェクトが下位概念において新たに付加される条件を全て満足した場合でも、上位概念が与える類似度を越えることがあってはならない．たとえば、オブジェクト o_1 が所属する原始概念が持つ条件の集合を Θ_{o_1} とし、 o_1 に対するオブジェクト o_2 の類似度を

$$\text{類似度} = \frac{\Theta_{o_1} \text{ 内で } o_2 \text{ が満足する条件の数}}{\Theta_{o_1} \text{ の要素数}}$$

と定義した場合を考える．このとき、概念系列中の任意の概念 C_i に対して、 C_i において満足される条件の総数は、下位概念によって満足される可能性がある条件の総数よりも大きくなければならない．したがって、形式的には、以下の条件が満足される必要がある．

$$\text{con}(C_i) - \text{con}(C_{i-1}) < \text{con}(C_a) - \text{con}(C_i)$$

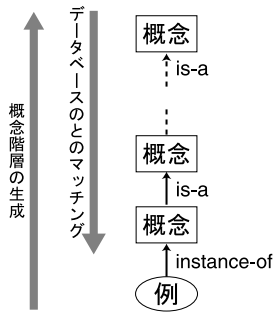


図 4: 概念階層の生成とマッチング

ここで、 $\text{con}(C_a)$ は原始概念を表す。

上記の条件は、上位概念において処理すべき条件の数が規定されることを示している。ここでは、条件に重みが一律であることを仮定したが、条件によって類似度に関わる重みが異なる場合がある。上位概念では、多数のオブジェクトとの条件マッチングが必要となるので、処理コストを小さくするために条件の数をできるだけ少なくしたい。それぞれの条件に重みづけがなされている場合、重みが多い条件を上位概念でチェックするようにすることにより、上位概念に割り当てられる条件の総数を減少させることができる。

下位概念から上位概念を生成する操作を汎化と呼ぶ。一つ概念に対する汎化には様々なものを考えることができる。以下の条件を満足するような汎化を抽象化と呼ぶ。

【定義：抽象化の条件】生成された上位概念における条件を満足しなかったオブジェクトが下位概念において新たに付加される条件を全て満足した場合でも、上位概念が与える類似度を越えない。

本手法では利用者が与えた例に対して抽象化を満足するような汎化を行い、上位概念からトップダウンに検索処理を行う。

利用者が与える例が有する特性には、利用者が恣意的に含めた部分と、偶然含まれた部分とが存在する。効率的な処理を実現するためには、偶然含まれたと予想される部分をチェックするオブジェクトの総数を減少させることが望ましい。また、偶然含まれたと予想される部分は検索における重要度が低いため、ある一定の類似度が予想される対象の順序付けをするために用いれば良いと考える。したがって、偶然含まれたと予想する特性をチェックする条件を内包から削除することにより、効率的な処理を実現する。

4 類似シーン検索

本稿で提案する類似シーン検索手法は、イベント・アクティビティ・モデルに基づいて内容が記述されたビデオ・データベースを対象としている。本節では、まず、イベント・アクティビティ・モデルの簡単な説明を行い、次に前節のアプローチの適用方法について述べる。

4.1 イベント・アクティビティ・モデル

実体は、対象世界上で区別可能で継続的に存在するモノであり、文字列として表現される。イベントは、フレーム間に観測可能な特徴的な実体の特性の変化および実体間の関係の変化を表す。イベントは、イベント識別子 eid と実体 en_i ($1 \leq i \leq n$) の n 項の対 $(eid, [en_1, \dots, en_n])$ として表現される。実体とイベントはそれぞれ型に所属する。たとえば、野球の試合においては、実体として”投手”、”ボール”などを考えることができる。また、イベント型 throw に所属するイベント ($\#6, [”投手”, ”ボール”]$) は、投手がボールを投げることを表す。

コンテキストは特殊記号 $\mathcal{L}, \$$ とイベント ev_i ($1 \leq i \leq n$) から構成される系列 $\mathcal{L}, ev_1, \dots, ev_n, \$$ である。ここで、 \mathcal{L} は系列の先頭を、 $\$$ は系列の最後尾を表す。ビデオ中で発生したイベントの並びを表すコンテキストが本モデルにおけるビデオの内容表現となる。それぞれのイベントはインデックス関数によって動画像中のフレームに一意に対応づけられ、イベントによってフレームを同定可能である。

イベントは対応づけられたフレームを同定するために一意に区別される必要があるが、イベントのマッチングを行うためには、イベントを分類しなければならない。この分類のために、イベント記号を導入する。イベント記号はイベントの種類を表すための記号であり、 $e = (eid, t_{ev}, arg)$ に対して同一の型と同一の引数を持つイベントのみにマッチするパターンを e の原子イベント記号と呼び (t_{ev}, arg) と表記する。それぞれのイベントに対して原子イベント記号は一意に定まる。例えば、イベント型 throw のインスタンス ($\#6, [”投手”, ”ボール”]$) に対応する原始イベント・パターンは $(\text{throw}, [”投手”, ”ボール”])$ である。本稿では、イベントは系列中の順序によって識別されるものとし、イベントとその原子イベント記号を区別しない。

4.2 系列の類似度

イベント・アクティビティ・モデルでは、シーンの内容はイベント系列として表現される。したがって、類似シーン検索は、利用者が指定したシーンの内容を表すイベント系列に類似したコンテキストの部分系列を検索することに

よって実現される。

いま、イベント系列 $x = x_1, \dots, x_n$ に対する、イベント系列 $y = y_1, \dots, y_m$ の類似性を考える。系列 x は、全順序集合 (E_x, \leq_{e_x}) として表現できる。ここで、 E_x はイベント系列 x を構成する全ての要素からなる集合 $E_x = \{x_1, \dots, x_n\}$ であり、全順序関係 \leq_{e_x} は $x_i \leq_{e_x} x_j (i \leq j)$ として定義される。類似性を判断するためには、系列間の要素を対応づける必要がある。いま、 x のそれぞれの要素に対して、 y 中の要素の 1 個または 0 個の要素が対応づけられるとする。このとき、イベント系列 x 中の i 番目の要素 x_i に対応づけられる y の要素を $\text{assign}(x_i)$ と表記する。 x_i に割り当てる要素が 0 個である場合は $\text{assign}(x_i) = \epsilon$ となる。この対応づけにおいては、それぞれの系列での順序関係が保持されるように以下の条件が満足されるものとする。

【定義：対応づけの条件】系列 x を表す全順序集合を $E_x = (\{x_1, \dots, x_n\}, \leq_{e_x})$ 、系列 y を表す全順序集合を $E_y = (\{y_1, \dots, y_m\}, \leq_{e_y})$ とする。このとき、系列 x の任意の 2 つの要素 x_i, x_j に対する対応づけ $\text{assign} : E_x \rightarrow E_y$ は、以下の条件を満足する。

$$x_i \leq_{e_x} x_j \iff \text{assign}(x_i) \leq_{e_y} \text{assign}(x_j)$$

なお、 $\forall y \in E_y$ に対して、 $\epsilon \leq_{e_y} y$ および $y \leq_{e_y} \epsilon$ が成り立つものとする。

いま、2 個のイベント e_1, e_2 の類似度を $s(e_1, e_2)$ と表記し、 $0 \leq s(e_1, e_2) \leq 1$ を満足するものとする。さらに、ある対応づけにおける x_i の類似度 $s(x_i, \text{assign}(x_i))$ を s_i と表記し、これに基づいて系列間の類似度を以下のように定義する。

【定義：系列の類似度（対応づけが与えられる場合）】系列 x に対する系列 y の対応づけ assign が与えられている場合、これら 2 系列間の類似度を以下のように定義する。

$$S(x, y, \text{assign}) = \frac{\sum w_i s_i}{\sum w_i}$$

ここで、 w_i は系列 x の i 番目の要素の類似度の重みを表す。

上記の類似度において、分母は対応づけに対する重みの和であり、分子は重みづけられた要素の類似度の和である。要素の類似度が全て 1 の場合に系列の類似度は最大値 1 をとり、いずれの条件も満足しない系列に対しては最小値 0 をとる。なお、ここでは簡単化のために、 y において対応づけがなされなかった要素が類似度に与える影響は無視している。

つぎに、系列の類似度を、対応づけが与えられた際の類似度の最大値として定義する。

【定義：系列の類似度】系列 x に対する系列 y の類似度は

$$S(x, y) = \max_{\text{assign}} (S(x, y, \text{assign}))$$

である。

4.3 概念の抽象化

類似検索のために概念階層を生成する目的は、効率的に類似度の計算を行うことである。効率化のために、あらかじめ類似度が低いことが判明したオブジェクトに関してはその詳細な値を計算する必要がないことを前提とする。

系列間の類似度の分母は定数であり、これを W とする。また、分子を構成する個々の要素に対する類似度は独立に求めることができる。そこで、例として与えられた系列 x を構成する要素を 2 つの集合 E_1, E_2 に分解した場合、それぞれの集合における要素の重みの和をそれぞれ W_{E_1}, W_{E_2} とする。これらは、類似度の定義により $W_{E_1} + W_{E_2} = W$ を満足する。いま、 $\frac{W_{E_2}}{W} \leq \frac{W_{E_1}}{W}$ となるように、 E_1, E_2 を構成したとする。このとき、ある系列 y の E_1 の要素に対する対応づけの類似度が全て 0 であった場合、 E_2 の対応づけの類似度が全て 1 であったとしても、系列の類似度は W_{E_2} を越えることはない。したがって、 E_1 の対応づけに関する類似度の計算をすれば、類似度が W_{E_2} 以下であるオブジェクトの一部を特定できる。

この考察に基づいて、抽象化の条件を満足する概念系列を生成することを考える。まず、系列間の比較に必要な対応づけの条件を θ_0 とする。つぎに、 E_1 の要素に関する類似度を計算することによって得られる条件 θ_1 を

$$\theta_1 \triangleq \frac{S((E_1, \leq_{e_x}), y)}{W} > 0$$

として定義し、 E_2 の要素に関する類似度を計算することによって得られる条件 θ_2 を

$$\theta_2 \triangleq \frac{S((E_2, \leq_{e_x}), y)}{W} > 0$$

として定義する。

いま、概念 C_1 の内包を $\text{con}(C_1) \triangleq \{\theta_0, \theta_1\}$ とし、概念 C_a の内包を $\text{con}(C_a) \triangleq \{\theta_0, \theta_1, \theta_2\}$ とする。このとき、概念 C_1 は概念 C_a を抽象化した概念になっている。

ここで、 E_2 をさらに 2 個の集合 E_{21}, E_{22} に排他分割し、それぞれの集合における要素の重みの和を $W_{E_{21}}, W_{E_{22}}$ とする。さらに、 $\frac{W_{E_{22}}}{W} \leq \frac{W_{E_{21}}}{W}$ となるよう

に, E_1, E_2 を構成し,

$$\theta_{21} \triangleq \frac{S((E_{21}, \leq_{e_x}), \mathbf{y})}{W} > 0$$

とし, 概念 C_2 の内包を $\text{con}(C_2) \triangleq \{\theta_0, \theta_1, \theta_{21}\}$ とする. このとき, C_2 は C_a を抽象化した概念になっており, C_1 は C_2 を抽象化した概念になっている. こうした処理を繰り返すことにより, 上位概念が抽象化の条件を満足するような概念系列を構成できる.

5 イベントの類似度と重みづけ

利用者が指定した例は利用者が検索したい活動と完全に一致していない場合が多い. 例の中には利用者が恣意的に含めた部分と偶然に含まれてしまった部分とが混在していると考えられる. 本節では, イベントの情報量とエントロピーに基づいてそれらの部分を判断し, イベントの類似度と重みを定義する.

5.1 イベントの情報量とエントロピー

イベント発生 of 統計的な性質を求めるために, コンテキストを全イベント集合 $\mathbf{E} = \{e_1, \dots, e_N\}$ を状態空間とする離散的な確率過程 X_m としてモデル化する. X_0, \dots, X_{n-1} の離散的な n 次元確率分布を

$$\begin{aligned} p(X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}) \\ \triangleq p(x_0, x_1, \dots, x_{n-1}) \end{aligned}$$

と表記する. また, 確率過程 X_m から発生した長さ L のイベント系列 x_0, x_1, \dots, x_L を \mathbf{x} と表記し, 部分系列 $x_i, x_{i+1}, \dots, x_j (0 < i \leq j < L)$ を \mathbf{x}_i^j と表記する. \mathbf{x} において, x_i 以外の要素を条件とする x_i の条件付き確率を

$$p(x_i | \mathbf{x}_0^{i-1}, \mathbf{x}_L^{i+1}) \triangleq \frac{p(x_0, \dots, x_L)}{p(x_0, \dots, x_{i-1}, x_{i+1}, \dots, x_L)}$$

と表記する.

いま, X_m を単純マルコフ過程であると仮定する. 単純マルコフ過程は, 条件

$$p(x_n | \mathbf{x}_0^{n-1}) = p(x_n | x_{n-1}), \quad n = 1, 2, \dots$$

を満足する確率過程である. すなわち, イベント x_n が出現する確率は直前に発生したイベント x_{n-1} へのみ依存する. 確率 $p(x_i | x_{i-1})$ は既知データにおけるイベントの並び x_{i-1}, x_i の出現頻度 $\text{freq}(x_{i-1}, x_i)$ を用いて, 以下のよう推定する.

$$p(x_i | x_{i-1}) \simeq \frac{\text{freq}(x_{i-1}, x_i)}{\sum_{x_{i-1}} \text{freq}(x_{i-1}, x_i)}$$

イベントの生起が単純マルコフ過程に従うとすることにより, コンテキストは単純マルコフ情報源としてモデル化できる. マルコフ情報源は内部状態を持つ情報源であり, イベント e_m が発生した直後の情報源の状態を S_{e_m} と表記する. このとき, 時刻 $i-1$ において状態が S_{e_m} であり, 時刻 i において状態 S_{e_n} に遷移する確率は

$$p_{e_m, e_n} \triangleq p(X_i = e_n | X_{i-1} = e_m)$$

となる.

長さ L のイベント系列 \mathbf{x} 内の要素 x_i , ($0 < i < L$) を考える. x_i 以外の要素を条件とする x_i の条件付き確率は, 単純マルコフ過程の条件から,

$$\begin{aligned} p(x_i | \mathbf{x}_0^{i-1}, \mathbf{x}_{i+1}^L) &= \frac{p(x_0, \dots, x_L)}{\sum_{x_i} p(x_0, \dots, x_L)} \\ &= \frac{p_{x_{i-1}, x_i} p_{x_i, x_{i+1}}}{\sum_{x_i} p_{x_{i-1}, x_i} p_{x_i, x_{i+1}}} \\ &= \frac{p_{x_{i-1}, x_i} p_{x_i, x_{i+1}}}{p_{x_{i-1}, x_{i+1}}^{(2)}} \end{aligned}$$

である. この確率の意味は, 分子は x_{i-1} が出現した後に x_i が出現する確率を表し, 分母は x_{i-1} が出現してから 2 ステップ後に x_{i+1} が出現する確率を表している. 上の式から, イベント系列 \mathbf{x} の要素 x_i が与える情報量は

$$I(x_i | \mathbf{x}_0^{i-1}, \mathbf{x}_{i+1}^L) = -\log \frac{p_{x_{i-1}, x_i} p_{x_i, x_{i+1}}}{p_{x_{i-1}, x_{i+1}}^{(2)}}$$

となる. また, X_i に関する不確定度は条件付きエントロピー

$$\begin{aligned} H(X_i | \mathbf{x}_0^{i-1}, \mathbf{x}_{i+1}^L) \\ = -\sum_{x_i} \frac{p_{x_{i-1}, x_i} p_{x_i, x_{i+1}}}{p_{x_{i-1}, x_{i+1}}^{(2)}} \log \frac{p_{x_{i-1}, x_i} p_{x_i, x_{i+1}}}{p_{x_{i-1}, x_{i+1}}^{(2)}} \end{aligned}$$

と与えられる.

5.2 重要度と重みづけ

系列中のイベントの情報量は, 系列中でそのイベントが出現する確率に基づいて定義され, イベント系列の中でそれぞれのイベントが担う重要度を表す尺度であると考えることができる. しかし, 情報量が高いイベントであっても, エントロピーが高い場合には, イベント発生に関する不確かさが大きくなるため, 例に対象とするイベントが偶然含まれてしまった可能性が高くなる. そこで, 情報量とエントロピーの相互関係に基づいて重要さを表す尺度として恣意度 (arbitrariness) を導入する. イベント系列 \mathbf{x} の i 番目の要素 x_i の恣意度 $\text{arb}(x_i)$ は

$$\text{arb}(x_i) = \begin{cases} \frac{2}{\pi} \tan^{-1} \frac{I_i}{H_i} & , \text{if } H_i \neq 0 \\ 1 & , \text{if } H_i = 0 \end{cases}$$

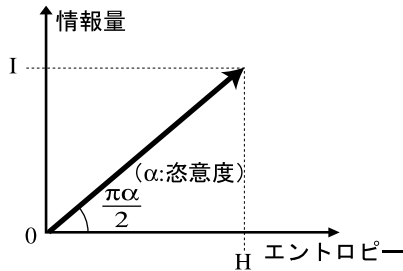


図 5: 恣意度

である．ここで， $I_i = I(x_i | x_0^{i-1}, x_{i+1}^L)$ であり， $H_i = H(X_i | x_0^{i-1}, x_{i+1}^L)$ である．恣意度の概念図を図 5 に示す．

恣意度を用いて重要度 (importance) を以下のように定義する．

【定義：重要度】イベント系列 x の i 番目の要素 x_i の重要度を

$$\text{imp}(x_i) \triangleq \text{arb}(x_i)I(x_i | x_0^{i-1}, x_{i+1}^L)$$

として定義する．

我々の類似検索手法では， x_i 重要度を x_i に対応づけの類似度の重みとして利用する．

5.3 許容度と類似度

イベント系列の中で恣意度が高いイベントは，そのイベントが表す特徴を利用者が重要視していると考えられる．したがって，恣意度が高いイベントが他の恣意度が高いイベントに置き換えられた場合には，重要視していた特徴が別の特徴に変化してしまうため，当初の系列の内容との分離が大きくなる可能性が高い．一方，恣意度が低いイベントは，そのイベントが表す特徴には利用者はあまり注目していないため，イベントが置き換えられた場合にも内容におよぼす影響は少ないと考えられる．そこで，比較元のイベント系列 x の要素 x_i に類似していると見なすことが可能なイベントの恣意度の上限を許容度 (tolerance) と呼び

$$\text{tol}(x_i) = 1 - \text{arb}(x_i)$$

として定義する．恣意度が高くなるにつれて許容度は低くなる．

類似度を以下のように定義する．

【定義：類似度】イベント系列 x の i 番目の要素 x_i とイベント e の類似度を

$$\triangleq \begin{cases} 1 & (\text{if } x_i = e) \\ 0 & (\text{else if } \text{arb}(e) > \text{tol}(x_i)) \\ 1 - \text{arb}(x_i) - \text{arb}(e) & (\text{otherwise}) \end{cases}$$

として定義する．なお，ここで $\text{arb}(e)$ は系列 x の i 番目の要素が e であったと仮定したときの恣意度である．

なお，ここで示したイベントの類似度の評価ではイベントの内部構造を考慮していない．実際には，同一の型に所属するイベントや同一の引数を有するイベントは類似度が高いと考えることができる．イベントの内部構造を含めたイベント系列の要素の類似度の評価は今後の課題である．

6 まとめと今後の課題

本稿ではビデオ・データベースにおけるシーンの類似検索を効率的に行うためのアプローチについて述べた．このアプローチでは，利用者が指定した例を抽象化することにより概念階層を動的に生成し，抽象度の高いレベルからトップダウンに検索処理を実行する．

今後，以下の方針で研究を進めていく予定である．

- 具体的な処理アルゴリズムの設計と評価．
- 本稿で定義した類似性の正当性の検証．
- DP マッチング等の処理方法との比較実験による本アプローチの効率性の検証．
- 利用者が複数の例を指定することによる明確な意図推定と適切な類似性の評価手法の開発．

参考文献

- [1] 牛尼剛聡，渡邊豊英：ビデオ・ブラウジングのための確率モデルに基づく概略表現生成手法，アドバンスト・データベース・シンポジウム 99 予稿集，pp. 39–48 (1999).
- [2] 牛尼剛聡，広部一弥，渡邊豊英：利用者の視点に基づくシーン検索のためのイベント・アクティビティ・モデル，信学論 (D-I)，Vol. J82-D-I，No. 1，pp. 256–267 (1999).
- [3] 古井貞熙：デジタル音声処理，東海大学出版会 (1985).
- [4] 柳沼良知，坂内正夫：DP マッチングを用いたドラマ映像・音声・シナリオ文書の対応づけ手法の一提案，信学論 (D-II)，Vol. J79-D-II，No. 5，pp. 747–755 (1996).
- [5] Flickner, M. et al.: Query by Image and Video Content: The QBIC System, *IEEE Computer*, Vol. 28, No. 9, pp. 23–32 (1995).
- [6] 長坂晃朗，宮武孝文：時系列フレーム特徴の圧縮符号化に基づく映像シーンの高速分類手法，信学論 (D-II)，Vol. J81-D-II，No. 8，pp. 1831–1837 (1998).