

アノテーション時刻とショット長の確率分布に基づく 粒度可変型ビデオシーン検索

6Z-08

吹野 直紀[†] 角谷 和俊[‡] 田中 克己[‡]
[†]京都大学工学部 [‡]京都大学大学院情報学研究科

1 はじめに

今後、蓄積型 TV やモバイル端末への動画配信の普及により、ビデオの中の特定のシーンを取り出したいという要求が増えると考えられる。通常それを実現するにはビデオ内容にメタデータとしてラベルを付けなければならない。キーワード検索に用いるラベルは通常、ある区間に対して付く。ラベルが付く区間の決め方に関しては、ビデオをまず小さな区間に区切った後ラベルを各々の区間に付けていく方法と、ラベルごとにそれが形容する区間を決める方法があるが、前者はサッカー映像等シーンの区切れ目が曖昧な映像には向かず、後者はアノテーションの際に時間がかかり、生中継映像に対してリアルタイムにラベルを付ける際等に困難が伴う。

そこで、検索時に使う同義語・意味的上位下位関係等を記録したシソーラスに、ショット粒度の情報を付加する事によって、区間ではなく時間に対し断片的といったラベルを元にキーワード検索を実現する方法を提案する。

2 キーワードごとのショット粒度

例えば、サッカーの映像のある時点においてアノテーターが「カウンター」というラベルを付けた場合、その 3 秒後もカウンター攻撃中である可能性は高い。しかし、「スループス」というラベルを付けた場合、3 秒後もまだスループスが続けている可能性は低い。つまり、キーワードごとにそれが形容するショットの予想される時間的な幅が異なる。この幅をキーワードごとのショット粒度と呼ぶ。

このショット粒度は、確率分布で表現する事ができる。図 1 は、カウンター、スループス、ゴールそれぞれのラベルが付いた時点前後の、各々のラベルが適している確率の分布の例である。縦軸に確率、横軸にそのラベルが付いた時点をととした時間をとっている。分布の時間的な幅、ラベルが付いた時点の前方と後方どちらに広く分布しているか、等に違いがある。確率分布を正確に出すには、多くのサンプルから統計を取る必要がある。

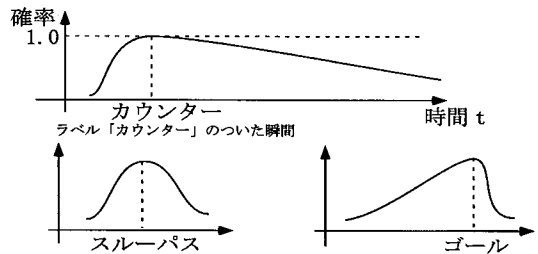


図 1: ラベル別の確率分布

3 ショット粒度情報を用いた検索

キーワードがいくつか入力された時、「全ての単語をできるだけ高い関連性を持って含む区間」を取り出したいとする。ラベルが指す区間を、そのラベルの確率分布曲線がある確率 θ を上回っている部分であると定義すると、「全ての単語を含む区間」はそれぞれのラベルが指す区間の集合に対してペアワイズグループ結合 [1] を行う事で全て取り出す事ができる。その結果図 2 の例では 4 つの区間が出る。その中には当然ふさわしく無い区間が多く含まれているので、その中からより適切なものを取り出す為の尺度として関連度というものについて考える。

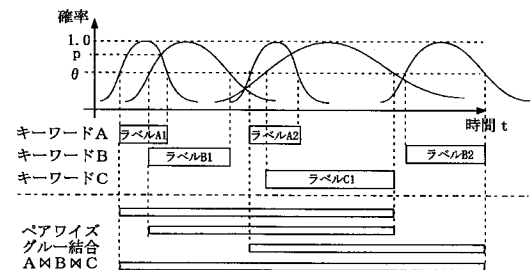


図 2: ペアワイズグループ結合

3.1 ラベル間の関連度

2 つのラベルの関連性について、次の傾向について考える。

- 同じ粒度のラベルなら、時間的に近い位置にある 2 ラベルは関連性が高い。

Video Scene Retrieval with Variable Granularity by Annotation-time/Shot-length Probability Distribution

Naoki Fukino[†] Kazutoshi Sumiya[‡], Katsumi Tanaka[‡]

[†]Faculty of Engineering, Kyoto University

[‡]Graduate School of Infomatics, Kyoto University

- 2ラベル間の時間差が一定とすると、2ラベルそれぞれの粒度が大きい方が関連性は高い。

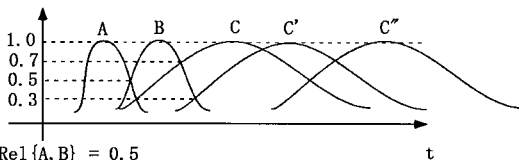
そこで、ラベル A がある時点 t を形容している確率を $P_A(t)$ 、ラベル B の同様の確率を $P_B(t)$ として、 $f(t) = \min\{P_A(t), P_B(t)\}$ なる $f(t)$ の最大値を 2 単語の関連性の尺度として採用する。図 2 の例ではラベル A 1 とラベル B 1 の関連度は p となる。この値は上記の両性質を満たす。以下、「ラベル A とラベル B の関連度が p である」という事実を $Rel\{A, B\} = p$ と表す。

3.2 3つ以上のラベルへの拡大

3つのラベル A, B, C 間の関連度に関して考える。まず、例えばラベル C が A と B ととも関連度が 0 の場合、この区間は A と B の間にしか関連が無く、不適切である。しかし、A と C に関連が無かったとしても、A と B、B と C に関連があれば、この区間は「A B C」というクエリーに従っているとと言える。このような性質を持つ関連度を出すため、次のような式を定義する。

$$Rel\{A, B, C\} = \frac{Rel\{A, B\} + Rel\{A, C\}}{2} \times \frac{Rel\{B, C\} + Rel\{B, A\}}{2} \times \frac{Rel\{C, A\} + Rel\{C, B\}}{2}$$

具体例としては、図 3 のようになる。4つ以上の場合も、それぞれのラベルに関して他のラベルとの関連度の平均を取り、その平均をそれぞれ掛ける計算をする。



$$\begin{aligned} Rel\{A, B\} &= 0.5 \\ Rel\{A, C\} &= 0.3, \quad Rel\{B, C\} = 0.7 \\ Rel\{A, C'\} &= 0.0, \quad Rel\{B, C'\} = 0.3 \\ Rel\{A, C''\} &= 0.0, \quad Rel\{B, C''\} = 0.0 \\ Rel\{A, B, C\} &= (0.5+0.3)/2 * (0.7+0.5)/2 * (0.3+0.7)/2 \\ &= 0.4 * 0.6 * 0.5 \\ &= 0.12 \\ Rel\{A, B, C'\} &= (0.5+0.0)/2 * (0.3+0.5)/2 * (0.0+0.3)/2 \\ &= 0.015 \\ Rel\{A, B, C''\} &= 0 \end{aligned}$$

図 3: 3つのラベルの関連度

関連度は、より多くのラベルの確率分布が重なっている場所が多い方が高くなる傾向にある。当然、全てのラベルが重なっているケースが最も関連度は高い。つまり、関連度の高いものから順に検索結果を吟味していくと、「 $A \wedge B \wedge C$ 」を満たす区間から優先的に取り出され、「 $A \wedge B \wedge C'$ 」等の少し緩い条件に適合するものが現れ、「 $A \wedge B \wedge C''$ 」の条件しか満たさないものが最後に現れる。これは、「A B C」というクエリーを初め「 $A \wedge B \wedge C$ 」と捉えて検索を行い、徐々にクエリーを変換して条件を緩めながら「 $A \wedge B \wedge C'$ 」という検索まで行っていく操作に近い。スコアが 0

の区間に関しては、各ラベルの確率分布曲線を時間軸方向に引き伸ばすことにより各ラベル間の距離が近い順に優先順位を付ける事が出来る。

3.3 ラベルの順序

ビデオデータは時系列データであるので、ラベルには順序が存在する。ユーザーが順序を意識してクエリーを「A B」と決めている場合、ラベル B の後ラベル A となっている区間の関連度にある値 w ($0 < w < 1$) を掛ける事で順序が違う区間のランキングを下げる。3つ以上のラベルの場合、全ての2ラベルの組み合わせについて順序を調べる必要がある。

3.4 周辺ラベルによるショット粒度制御

ラベルによっては、同時に同じシーンを形容する可能性が無いものがある。例えば、サッカーにおいてカウンター途中にコーナーキックが行われる事はない。コーナーキックというラベルが付いた場合、その時点でカウンターは終わっている事になる。そこから先のカウンターの確率は 0 なので、このラベルの確率分布は図 4 のように変える事ができる。同時に起こらないラベルの情報をシソーラスに持っておき、このような変換を行う事で多くの不適切な区間を除くことができる。

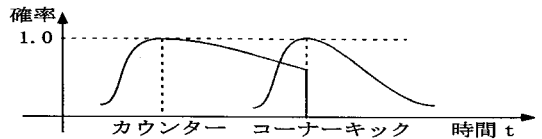


図 4: 粒度の変形

4 おわりに

リアルタイムにアノテーションされる事を想定して、区間でなく時刻に対して付いた断片的なラベルを基にショット粒度情報を用いて検索を行う方式について提案した。今後プロトタイプを実装しシステムの有効性について検証する。また、現時点ではタイミングのずれを除くアノテーションの間違いや冗長性が無いという前提だが、リアルタイムに豊富なラベルを付けるにはグループによるアノテーションを行う必要があるので、その場合間違いや冗長性を考慮しない訳にはいかない。この点の解消法についても検討する。

謝辞

本研究の一部は、NHK と京都大学との民間共同研究、文科省科研費 (課題番号 12680416, 13224054)、日本学術振興会未来開拓学術研究推進事業 (JSPS-RFTF97P00501) による。

参考文献

- [1] プラダンスジット, 田島 敬史, 田中 克己: ビデオデータ検索のための区間グループ操作と解のフィルタリング, 情報処理学会論文誌・データベース (Jan, 1999)