

## 時系列アクティブ探索法における 音響信号の高速 AND/OR 探索

柏野 邦夫      村瀬 洋

NTTコミュニケーション科学基礎研究所  
〒243-0198 神奈川県厚木市森の里若宮 3-1

kunio@ca-sun1.brl.ntt.co.jp, murase@eye.brl.ntt.co.jp

あらまし 先に我々は、音響信号の高速探索アルゴリズムである時系列アクティブ探索法を提案した。時系列アクティブ探索法は、ベクトル量子化符号のヒストグラムに基づいて探索を行う方法であり、我々は既に、探索の高速性と高い精度とを両立できることを報告してきた。本稿では、より柔軟な音響探索を実現するために、時系列アクティブ探索法による音響探索において、AND探索とOR探索を効率的に行う方法を提案する。すなわち、複数の参照信号や入力信号が与えられたときに、それらに関して単に探索を繰り返す場合に比べて、照合計算回数を削減できる場合があることを示す。さらに、その削減効果を実験的に確認する。

キーワード 音響探索, 音響検索, 枝刈り, 時系列, アクティブ探索

## Quick AND/OR Search of Audio Signals Using Time-Series Active Search

Kunio Kashino and Hiroshi Murase

NTT Communication Science Laboratories  
3-1 Morinosato-Wakamiya, Atsugi-shi,  
243-0198, Kanagawa, Japan.

Abstract In our previous work, we developed a quick audio search algorithm termed Time-Series Active Search. The feature of the method is that it employs vector quantization of signals and the histogram modeling of acoustic features. We have already shown that the method enables extremely fast and accurate audio search. In the present work, we investigate the methods to eliminate redundant matching in AND/OR search with multiple reference signals or input signals. Our goal here is to keep the number of matching calculations minimum necessary. We also present the experimental results that show the effectiveness of the proposed methods.

key words audio search, audio retrieval, pruning, time-series, active search

## 1 まえがき

UNIX 環境における標準的なツールに grep (および egrep, fgrep) というものがある。これらは、文字列探索、すなわち、ある特定の文字列が、別の長い文字列<sup>1</sup>中に含まれているかどうか (含まれているとすればどこか) の検査を行うコマンドである。これらは非常に便利であり、UNIX 環境において日常欠かすことのできないものである。

grep の仲間が便利である理由は、高速性と柔軟性にある。すなわち、これらのコマンドでは、高速な文字列探索アルゴリズムが用いられているために、普段の使用であまり待たされることがない。また、正規表現と呼ばれる形式によって探索文字列のバリエーションを一度に指定することができる<sup>2</sup>。

ところで、音に関して grep のようなコマンドを作ることは可能であろうか。信号自身や信号から抽出した特徴の相関に基づいて信号検出 [1] を行う方法では、長時間の信号に適用した場合膨大な処理時間がかかるという問題があるので、何らかの高速な方法が必要である。対象が音声 (スピーチ) の場合には、予め音声認識によって全ての発話をテキストに変換した上で、grep と同様の文字列探索を行うことが考えられるが、ここでは音声に限らず一般音を対象として考える。まず思いつくのは、音響信号をベクトル量子化などの方法で記号列に変換し、文字列探索に帰着する方法である。しかしこの方法は、アナログ系の雑音や歪みなど、信号の変形に影響を受けやすいという欠点がある。

先に我々は、音響信号の高速探索アルゴリズムである「時系列アクティブ探索法」を提案した [2, 3]。この方法は、音響信号のスペクトルを分類 (ベクトル量子化) し、各分類 (量子化符号) ごとの出現回数を数えてヒストグラムを作り、そのヒストグラムどうしの照合によって探索を進めることを特徴とする。ヒストグラムは累積特徴であるために、信号の変形による悪影響を受けにくい。また、ヒストグラムどうしの照合は、特徴どうしを直接照合するのに比べて照合自体の計算量が少ない上、時間軸方向で照合が不要な区間を求めて探索をスキップすることで、無駄な照合を省くことができる。これらの効果により、実用上十分な探索精度を保ったまま、スペクトル特徴の直接照合に比

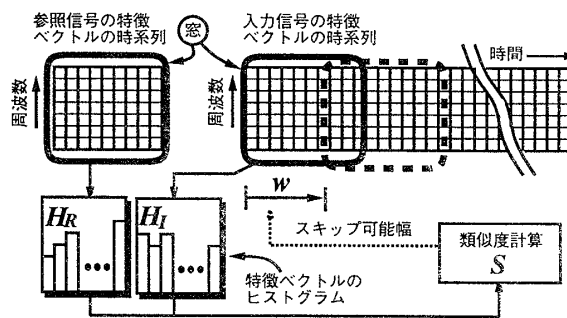


図 1: 時系列アクティブ探索法の概要

べて数百倍の探索速度が実現されている [3, 4]。

さて、時系列アクティブ探索法は、音響探索における高速な手法であるが、これを用いて grep のように便利なコマンドを実現するためには、残る条件として探索の柔軟性が必要である。すなわち、正規表現に相当するような一定の約束に従って探索目的の音を記述し、一度に探索を行えることが望ましい。このような柔軟な探索を実現するための基本となるのは、AND 探索と OR 探索である<sup>3</sup>。そこで本稿では、時系列アクティブ探索法を用いた音響探索において、AND 探索と OR 探索を効率的に行う方法を提案する。

以下 2. で、時系列アクティブ探索法について要約したのち、3. で OR 探索、4. で AND 探索について、照合回数の削減法を議論する。また、5. で、複数の長時間信号に対する探索を並行して行う場合について議論する。6. をむすびとする。

## 2 時系列アクティブ探索法

まず時系列アクティブ探索法について要約しておく。処理の流れを図 1 に示す。まず参照信号 (探索のキーとなる短時間の信号) と入力信号 (長時間の信号) からそれぞれ特徴ベクトルを抽出する。次に、参照信号と入力信号の双方に対して同じ長さの時間窓をかけ、窓内の特徴ベクトルを分類してヒストグラムを作る。そして、ヒストグラムどうしの類似度が、予め設定した値 (これを探索閾値と呼ぶ) を越えるかどうかで、参照信号の有無を判定する。このとき、類似度の値と設定値とから、探索を時間方向にスキップできる時間幅 (スキップ可能幅) を求めることができるので、その分だけ入力信号に対する窓をずらして探索を進める。

<sup>3</sup> AND 探索とは、複数の条件を共に満たす部分を探索すること、OR 探索とは、複数の条件のいずれかを満たす部分を探索することである。

<sup>1</sup> grep などのコマンドでは、探索は行単位で行われる。

<sup>2</sup> 指定の柔軟性は egrep > grep > fgrep である。

特徴ベクトルの作成方法や分類の仕方は各種考えられるが、我々は計算量の少なさを重視して、帯域フィルタバンクによって特徴ベクトルを作成し、特徴ベクトルの各要素の値をいくつかのビン (bin) に分割することによって分類している。この分類は、ベクトル量子化の簡単な場合に相当する。

また、ヒストグラムどうしの類似度についても、ユークリッド距離など各種のノルムに基づいて類似度を定義する方法など、様々なものが考えられる [5, 6]。これらのうちで、我々は特にヒストグラム重なり率に着目している。ヒストグラム重なり率  $S_{IR}$  は、次のように定義される。

$$S_{IR} = S(H_I, H_R) = \frac{1}{D} \sum_{l=1}^L \min(h_{Il}, h_{Rl}) \quad (1)$$

ここで  $H_I$  と  $H_R$  は、それぞれ入力信号と参照信号に対するヒストグラムであり、 $h_{Il}, h_{Rl}$  はそれぞれの  $l$  番目のビンに含まれる度数である。また  $L$  はヒストグラムのビンの数、 $D$  はヒストグラムの総度数 (参照信号から導かれた特徴ベクトルの総数) である。

このとき、スキップ可能幅  $w$  は、類似度の上限値に関する考察から、次式で求められる。

$$w = \begin{cases} \text{floor}(D(\theta - S)) + 1 & (S < \theta \text{ のとき}) \\ 1 & (\text{上記以外}) \end{cases} \quad (2)$$

ただし  $w$  の単位は特徴ベクトルの個数であり、 $\text{floor}(\cdot)$  は切り下げを表し、 $\theta$  は探索閾値である。類似度が  $\theta$  を越える箇所については全探索を行う (時間窓を一位ずつずらす) こととしている。

式 (1) において、各種の定義が考えられる中でヒストグラム重なり率による類似度の定義を用いる理由は、(1) 類似度計算が簡単であること、(2) 時間窓の周辺における類似度の上限値が簡単な計算によって求められること、および (3) 既に画像の物体認識などに適用され、好ましい結果が得られていること [7, 8]、の 3 点である。特に、類似度計算は探索実行中に行う必要があるため、上記 (1) および (2) のように計算が簡単であるというのは実際上極めて重要な条件である。本手法の場合、式 (1) と式 (2) から明らかなように、ある位置の時間窓に対するヒストグラムが与えられてから次の時間窓の位置を得る (スキップ可能幅を求める) ために必要な計算は、最大で「 $L+1$  回の大小判定と  $L+1$  回の加算と 1 回の減算と 1 回の切捨て」のみとなっている<sup>4</sup>。

<sup>4</sup>もし、類似度が探索閾値以下である場合にも類似度値を算出す

### 3 複数参照信号の OR 探索

#### 3.1 考察

音響探索の応用分野としてまず考えられるのは、テレビ放送やラジオ放送を蓄積したデータに対する特定のコマーシャル (CM) [9] や楽曲等の出現回数のカウント、インターネットにおける音響信号の探索エンジン等である。実は、これらの応用では、同時に数多くの参照信号に対して探索を行いたい場合が多い (OR 探索)。例えば、放送データに対する CM のカウントを行う場合、仮に同一商品の CM に着目したとしても、わずかに異なる複数の類似 CM が同時期に放送されているのが通例である。また楽曲使用回数の場合にも、同時に複数の楽曲の複数の箇所について探索を行いたいことが多い。

同一の入力信号に対して複数の参照信号を OR 探索するには、参照信号の数だけ探索を繰り返せばよいことは言うまでもないが、以下に議論するように、単に探索を繰り返した場合よりも照合計算回数を削減出来る場合がある [10, 11]。

図 2 に示すように、2 つの参照信号  $R_1, R_2$  からそれぞれヒストグラム  $H_{R1}, H_{R2}$  が作成され、入力信号  $I$  からヒストグラム  $H_I$  が作成されていたとする。ただし  $H_{R1}, H_{R2}$  の総度数は等しいとする。いま、 $H_I$  と  $H_{R1}$  について式 (1) を計算して類似度  $S_{IR1}$  を得たとしよう。このとき、もし  $H_{R1}$  と  $H_{R2}$  がよく似ていたとすると、 $S_{IR2}$  は、 $H_I$  と  $H_{R2}$  について式 (1) を計算しなくても、ある範囲内にとどまると考えられる。いま我々は、 $H_I$  と  $H_{R2}$  を具体的に照合することなく  $S_{IR2}$  の上限値を得ることに興味がある。そこで、 $H_{R1}$  と  $H_{R2}$  の類似度を  $S_{R1R2}$  とすると、 $S_{IR2}$  が上限を与えるのは、 $H_{R1}$  と  $H_{R2}$  の不一致部分がすべて  $S_{IR2}$  を増加させる方向に働いていたときである。すなわち、

$$S_{IR2} \leq S_{IR1} + (1 - S_{R1R2}) \quad (3)$$

である。ただし、もともと  $S_{IR2} \leq 1$  であるから、式 (3) は、 $S_{IR1} < S_{R1R2}$  のときだけ有効である。

一方、 $H_{R1}$  と  $H_I$  がよく似ていたとすると、 $S_{IR2}$  が上限を与えるのは、 $H_{R1}$  と  $H_I$  の不一致部分がすべて  $S_{IR2}$  を増加させる方向に働いていたときである。すなわち、

必要があるならば、これらの計算に加えて 1 回の除算が必要である。

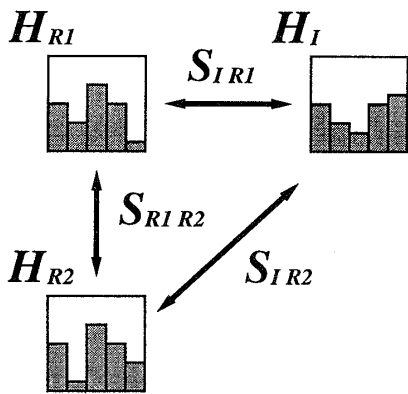


図 2: 類似度どうしの関係の説明図

$$S_{IR2} \leq S_{R1R2} + (1 - S_{IR1}) \quad (4)$$

である。ただし、もともと  $S_{IR2} \leq 1$  であるから、式 (4) は、 $S_{IR1} > S_{R1R2}$  のときだけ有効である。

式 (3) と式 (4) をまとめると、次式が得られる。

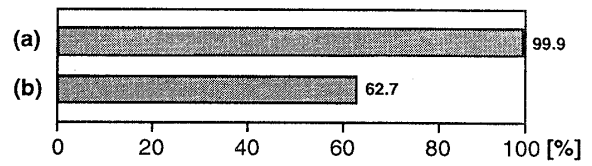
$$S_{IR2} \leq 1 - |S_{IR1} - S_{R1R2}| \quad (5)$$

すなわち、 $S_{IR1}$  と  $S_{R1R2}$  とがかけ離れた値であればあるほど  $S_{IR2}$  の上限値が制約される。式 (2) から、上限値が小さいほどスキップ可能幅が大きくなる。

以上のことから、複数の参照信号について OR 探索を行う方法は、以下ようになる。

1. 前処理として、参照信号どうしの類似度をすべての組合せについて計算しておく。
2. 現在位置を入力信号の最初に位置付ける。(ここから探索過程)
3. スキップ可能位置が現在位置に最も近い参照信号を一つ選択し、現在位置をそのスキップ可能位置とする。
4. 選択した参照信号と、現在位置の入力信号とを照合し、類似度を求める。
5. 得られた類似度をもとに、全ての参照信号に対するスキップ可能幅を更新する
6. 3. に戻る

これによって、探索過程における照合回数を、参照信号を別々に照合した場合以下とすることができる。



(a) 参照信号間の類似度が 0.11, 0.22, 0.23 の場合  
(b) 参照信号間の類似度が 0.72, 0.75, 0.88 の場合

図 3: 複数参照信号の OR 探索における照合回数比

### 3.2 実験

提案法により、複数の参照信号を別々に探索した場合に比べてどの程度照合計算回数が削減できるかを事例で調べるため、実験を行った。実験では、6 時間分のテレビ放送の音響信号を入力信号とし、15 秒間の CM を参照信号とした。提案法は、精度に関しては参照信号を別々に照合した場合と同一なので、照合回数のみを比較する。また、類似度が探索閾値を上回る場合については必ず全探索 (スキップを行わない探索) を行うので、ここでは、類似度が探索閾値以下であった照合回数が、参照信号を別々に照合した場合に比べてどのくらいの割合であるか (本稿ではこれを照合回数比という) を調べる。探索のパラメータは、サンプリング周波数=11.025 kHz、特徴次元数=7、周波数分析のフレーム長=60 ms、フレームシフト=10 ms、各特徴次元におけるビン数=3、探索閾値 $\theta=0.8$ とした。

図 3 に実験結果を示す。図 3 (a) は、無作為に選択した異なる商品の CM 3 本を参照信号とした場合、(b) は、音が極めて類似している同一商品の CM 3 本を参照信号とした場合の結果である。参照信号間の類似度は、(a) の場合 0.11, 0.22, 0.23、(b) の場合 0.72, 0.75, 0.88 であった。(a) の場合には、ほとんど照合回数が削減されていないが (照合回数比 99.9%)、(b) の場合には、閾値以下の類似度を与えた照合回数が 2/3 以下となっている (照合回数比 62.7%)。(a) の場合にほとんど効果がないのは、式 (5) によって与えられる類似度の上限値が大きいため、実質的にスキップ可能幅を増大できる場合が少ないことによると考えられる。

このように、提案法は、参照信号どうしの類似度が高い場合に有効である。しかしながら、

1. 事前の照合が必要なので ( $N$  個の参照信号に対して  $N(N-1)/2$  回)、 $N$  が大きく、参照信号どうしの類似度が低い場合には、かえって全体としての照合計算回数が増える場合がある

2. 全体としての照合計算回数が減少する場合であっても、照合のたびに全ての参照信号に対するスキップ可能幅を更新する操作が必要なので、全体としての計算コストが減少するとは限らない

という点には注意が必要である。

## 4 複数参照信号のAND探索

### 4.1 考察

複数参照信号のAND探索とは、例えば、30秒間の信号区間のうちで、ある部分が参照信号  $R_1$  に一致し、かつ、ある部分が参照信号  $R_2$  に一致するような区間を求めるような探索のことである。

参照信号のAND探索も、応用上重要な意味がある。時系列アクティブ探索法では、ヒストグラム（累積特徴）によって照合を行っているために、多少の信号の変形に影響されにくい反面、信号のわずかな違いを区別しにくいという性質がある。例えば、30秒間のCMで、末尾の2秒間だけが異なるといったようなバージョン違いのCMが、実際の放送では数多く存在しているが、30秒間にわたって累積した特徴を用いて2秒間の違いを区別することは難しい。このような場合、例えば初めの28秒間と末尾の2秒間を別々の参照信号とし、両者が一致するという条件でAND探索を行えば、正確な区別が可能となる。

ここでは、複数の参照信号  $R_1, \dots, R_N$  を、開始時刻にある時間差をつけて時間軸上に配置したとき、入力信号の中で、 $R_k$  のいずれに対しても探索閾値以上の類似度をもつ区間を求めることを考える。探索される区間の長さ  $t_d$  は、

$$t_d = \max_i(t_{off}(R_i)) - \min_j(t_{on}(R_j)) \quad (6)$$

である。ここで、 $t_{off}(R_i)$  は  $R_i$  の終了時刻、 $t_{on}(R_j)$  は  $R_j$  の開始時刻である。 $t_d$  の時間区間における全体の類似度  $S$  を以下のように定義すると、問題は、 $S$  が探索閾値を上回る箇所を探索することである。

$$S = \min_i(S_i) \quad (7)$$

ここで、 $S_i$  は、 $i$  番目の参照信号と、入力信号の対応する区間との類似度である。

ここでは、時間窓を移動するごとに  $t_d$  の時間区間における全体の類似度を計算する必要がある場合と、類似度の計算は必要ない場合とに分けて考える。

まず、全体の類似度を計算する必要がある場合、 $i$  番目の参照信号におけるスキップ可能幅  $w_i$  は、式 (2) から

$$w_i = \begin{cases} \text{floor}(D_i(\theta - S_i)) + 1 & \text{if } S_i < \theta, \\ 1 & \text{otherwise.} \end{cases} \quad (8)$$

となる。ここで  $D_i$  は  $i$  番目の参照信号における特徴ベクトルの総度数である。このとき、全体時間窓（長さ  $t_d$ ）に対するスキップ可能幅  $w$  について考える。式 (7) より、 $S_i$  のうちの一つでも  $\theta$  以下であれば、 $S$  も必ず  $\theta$  以下となるので、 $w_i$  のうちの  $i$  についての最大幅まで時間窓を移動させても、移動中に  $S$  が  $\theta$  を越えることはない。このことから、 $w$  は

$$w = \max_i(w_i) \quad (9)$$

と求めることができる。すなわち、類似度  $S$  を求めながら探索を行う場合の時間窓のスキップ可能幅は、式 (9) で求められる  $w$  となる。このとき、照合計算回数は、それぞれの参照信号について個別に照合を行う場合以下となることは明らかである。

次に、全体の類似度が閾値以下となる箇所においては全体の類似度を計算する必要がないという場合には、次のように探索を進めることができる。

1. 参照信号を、継続時間が長い順にソートする。
2. ある全体時間窓の位置に対して、長い参照信号から順に照合を行い、類似度  $S_i$  を求める。
3. 探索閾値  $\theta$  に満たない  $S_i$  が発見された時点で、直ちに全体時間窓をスキップする。その場合のスキップ可能幅は、それまでに計算されている  $w_i$  のうちの最大値とする。

ここで、継続時間が長い参照信号から順に照合を行うのは、長い参照信号の方が、平均的なスキップ可能幅が大きいと期待されるからである。

### 4.2 実験

参照信号のAND探索を行った場合の照合計算回数を、実験により確かめた。前節の実験と同様に、6時間分のテレビ放送の音響信号を入力信号とし、ある15秒間のCMを探索対象とした。このCMを、重複がないように均等な長さに8分割して8つの参照信号とし、これらのAND探索を行うことにした。探索のパラメータは前節の実験と同様とし、探索閾値  $\theta=0.7$  とした。

図4に照合回数比の測定結果を示す。本実験において照合回数比は、8つの参照信号のそれぞれについて個別に探索を行った場合の照合回数の和に対する、照合回数の比のことである。図4(a)は、時間窓を移動するごとに全体の類似度計算を行った場合、(b)は、探索閾値を越える箇所のみで類似度計算を行った場合の結果である。両者の場合とも、個別に探索を行った場合に比べ照合回数が削減されていることが示されている。特に(b)の場合に削減効果が大い。1つの参照信号だけの照合回数が平均12.5%(1/8)であることから、(b)の場合には、1つめの参照信号の照合で類似度が探索閾値に達せずスキップが行われる場合が大半を占めることが確かめられた。なお、(a)の場合と(b)の場合とでは、探索結果は同一である。

一方、図5と図6は、AND探索の効果を示したものである。図5は、上記15秒間のCM全体を1つの参照信号として探索した場合、図6は、CMを8分割して8つの参照信号のAND探索を行った場合の類似度のパターンを示している。8分割AND探索の方が探索閾値に対するマージンが大きくなっており、一部のみが異なる音響信号の区別などに効果があると考えられる。なお、図6では、類似度が探索閾値に達しない箇所においても類似度を求めて表示しているので、表示に必要な照合計算回数は図4(a)の場合である。

## 5 複数入力信号のAND探索

### 5.1 考察

前節までに、同一の入力信号に対して、複数の参照信号をOR探索やAND探索する場合について議論したが、この他に、複数の入力信号に対して、それぞれ対応する参照信号についてのAND探索を行いたい場合もある。例えば、音と映像についてのマルチメディア探索で、音と映像の双方が参照信号と類似している

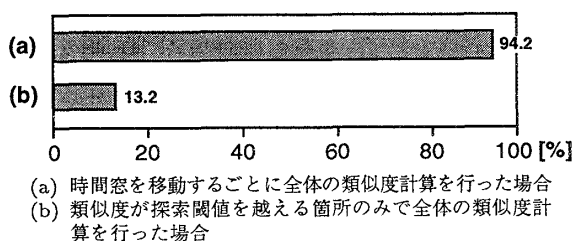


図4: 複数参照信号のAND探索における照合回数比

図5と図6は、テレビ放送の録音を入力信号とし、あるCM(15秒間)を参照信号とした探索結果。縦軸は類似度(0から1)、横軸は時刻(1998年1月22日)。○印が探索された場所を示す。これらの探索結果は正しい(探索もれも余分な出力もない)。探索閾値(破線)は0.7。

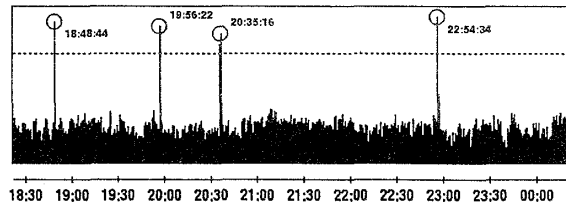


図5: 探索結果(単独探索の場合)

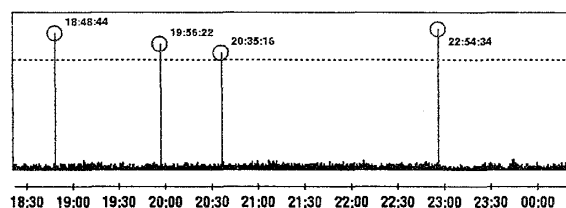


図6: 探索結果(8分割AND探索の場合)

箇所を探索したい場合がこれに相当する。

この場合は、前節の、参照信号のAND探索の場合と同様な議論が成り立つ。すなわち、各入力信号についての類似度  $S_i$  のうちの最小値を、全ての入力信号についての類似度  $S$  とすればよい。そして、探索途中における類似度を求めることが必要な場合には、各入力信号について求めたスキップ可能幅  $w_i$  を実時間に換算し、実時間上で最も大きいスキップ幅を採用することによって、照合回数を、それぞれの入力信号について別々に探索する場合以下にすることができる。これを、音と映像の並行探索を例として、図7に示す。また、探索閾値を越えた箇所においてのみ類似度を計算すれば良い場合には、前節と同様に、全ての入力信号に対する参照信号のうちで、実時間が長い順に探索を行い、探索閾値に満たない場合には直ちに時間窓のスキップを行う方法をとれば良い。

### 5.2 実験

入力信号に関するAND探索を行った場合の照合計算回数を、音と映像の並行探索を題材として実験により確かめた。前節の実験と同様に、6時間分のテレビ放送の音響信号および映像を入力信号とし、ある15秒間のCMの音響信号および映像を参照信号とした。音響探索に関するパラメータは前節の場合と同じとした。

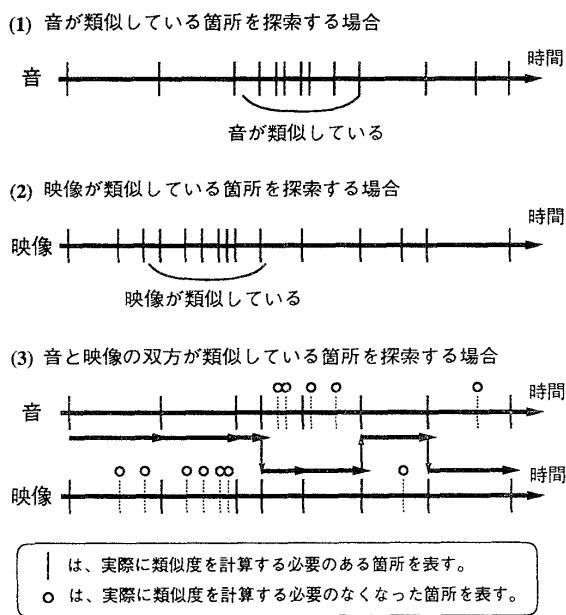


図 7: 入力信号に関する AND 探索（音と映像の並行探索）におけるスキップ可能幅の例

時系列アクティブ探索法は、映像から特徴ベクトルの時系列を抽出することにより、映像の探索にも音の場合と同様に適用することができる。本節の実験では、映像の特徴抽出は、毎秒 29.97 フレームの周波数でキャプチャした 32 × 24 サイズの映像の各フレームを 8 つの部分に分け、それぞれの部分における平均 RGB 値を算出して 3 次元特徴ベクトルとする方法を用いた。1 フレームの画像から 8 つの特徴ベクトルが生成されるので、これを単に時間軸方向に並べて特徴ベクトルの系列とした。すなわち映像探索のパラメータは、サンプリング周波数=29.97 Hz、特徴次元数=3、各特徴次元におけるピン数=8、探索閾値  $\theta=0.7$  とした。なお本実験では、音響信号・映像とも、参照信号を重複がないように等しい長さに 2 分割し、参照信号の AND 探索も併せて行った。

図 8 に照合回数比の測定結果を示す。本実験において照合回数比は、音響信号と映像のそれぞれについて、時間窓を移動するごとに全体の類似度計算を行いながら個別に探索を行った場合の照合回数の和に対する、照合回数の比のことである。図 8 (a) は、時間窓を移動するごとに全体の類似度計算を行った場合である。また (b) と (c) は、いずれも探索閾値を越える箇所のみで類似度計算を行った場合の結果であり、(b) は音響信号と映像のそれぞれについて個別に探索を行っ

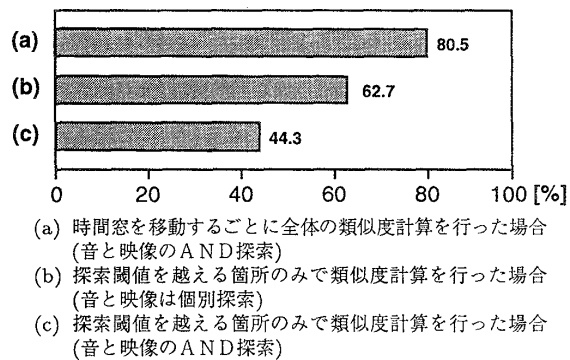


図 8: 複数入力信号の AND 探索における照合回数比

た場合の照合回数の和、(c) は入力信号の AND 探索を行った場合の照合回数を表している。探索閾値を越えない箇所において類似度計算を行った場合・行わなかった場合とも、個別に探索を行った場合に比べ、入力信号の AND 探索により照合回数が削減されていることが示されている。

## 6 むすび

本稿では、音響探索において grep のような柔軟な探索を実現するための基礎的な検討として、時系列アクティブ探索法における OR 探索と AND 探索の効率的な計算方法について考察した。参照信号の OR 探索では、事前に参照信号どうしの照合を行っておくことにより、類似した参照信号について探索を行った場合の照合計算回数が、別々に探索を行った場合以下にできることを示した。また参照信号や入力信号の AND 探索では、時間的に長い参照信号から優先的に照合を行い、探索閾値に満たない参照信号が発見された時点で時間窓の移動を行う方法により、効率的な探索が行えることを示した。また、各場合について、照合回数の削減効果を実験により確認した。

本稿で検討したような OR 探索と AND 探索を組み合わせることにより、いくつかのパターンを並行して精度良く探索を行うことや、マルチメディア探索を行うことが可能になると考えられる。今後は、高速探索という特性を保ったままで、さらに信号の変形やバリエーションを許容することのできる探索手法について検討を進める予定である。

## 謝辞

ご指導頂く NTT コミュニケーション科学基礎研究所の東倉洋一所長、石井健一郎部長、萩田紀博部長、およびメディア認識研究グループの諸氏に感謝する。

## 参考文献

- [1] J. C. Hancock and P. A. Wintz: "Signal Detection Theory", McGraw-Hill (1966).
- [2] G. Smith, H. Murase and K. Kashino: "Quick audio retrieval using active search", *Proc. of ICASSP-98*, Vol.6, pp.3777-3780 (May 1998).
- [3] 柏野 邦夫, ガビンスミス, 村瀬 洋: "ヒストグラム特徴系列に基づく長時間音響信号の高速探索", 音響講論集 98 秋, pp.561-562 (Sep.1998).
- [4] K. Kashino, G. Smith and H. Murase: "Time-series active search for quick retrieval of audio and video", *Proc. of ICASSP-99*, Vol.6, pp.2993-2996 (Mar.1999).
- [5] 杉山 雅英: "ビデオデータにおける指定人物の検出と追跡", 映像情報メディア学会技術報告, Vol.22, No.66, pp.7-12 (Nov.1998).
- [6] 杉山 雅英: "セグメントの高速探索法", 信学技報, SP98-141, pp.39-45 (Feb.1999).
- [7] M. J. Swain and D. H. Ballard: "Color indexing", *Int. J. Computer Vision*, Vol.7, No.1, pp.11-32 (1991).
- [8] V. V. Vinod and H. Murase: "Focused color intersection with efficient searching for object extraction", *Pattern Recognition*, Vol.30, No.10, pp.1787-1797 (1997).
- [9] R. Lienhart, C. Kuhmunch and W. Effelsberg: "On the detection and recognition of television commercials", *Proc. IEEE Conf. on Multimedia Computing and Systems*, pp.509-516 (1997).
- [10] 村瀬 洋, V. V. Vinod: "局所色情報を用いた高速物体検索 - アクティブ探索法 -", 信学論 D-II, Vol.J81-DII, No.9, pp.2035-2042 (1998).
- [11] 柏野 邦夫, 村瀬 洋: "複数の音響信号の高速探索", 音響講論集 99 春, pp.447-448 (Mar.1999).