

セグメントの高速探索法

杉山 雅英

会津大学 コンピュータ理工学部
〒965-8580 会津若松市一箕町

あらまし 音声・画像の時系列データとして与えられるビデオデータにおいて指定された人物の登場時刻・登場位置を高速に検出・追跡するための基礎検討を行なっている。今後のマルチメディア処理技術において、時空間における指定対象の検出及び追跡技術は重要な要素技術の一つとなる。文字列探索の手法は検索キーと検索対象（ファイル）の文字の完全一致を要求条件とする検索であり、幾つかの高速アルゴリズムが知られている。しかしながら音声などの特徴パターンを検索キーとする検索においては、パターンが自然に持つゆらぎや雑音の影響があるので文字列に対する完全一致探索を適用することはできない。完全一致探索を緩めた部分文字列一致探索や歪み許容探索が必要となる。本論文では、パターンの持つゆらぎなどの歪みの特性とそれに応じた探索手法に関する関連付けについて述べる。

キーワード マルチメディア処理, 音響セグメント, 探索手法

Fast Segment Search Algorithms

M. Sugiyama

School of Computer Science and Engineering
The University of Aizu
Aizu-Wakamatsu, Fukushima 965-8580

Abstract Multimedia database management and retrieving are on a world-wide demand and is a core in a search engine of a huge multimedia database. Current search engines are fundamentally based on character based matching techniques and several fast searching algorithms for characters are well-known, however, searching on a multimedia database must handle fuzzy and robust matching because multimedia information including sounds and images has naturally noise and distortion. This paper describes various fast search algorithms considering characteristics of noise and distortion.

Key words: Multimedia processing, Acoustic segment, Searching methods

1 まえがき

音声・画像の時系列データとして与えられるビデオデータにおいて指定された人物の登場時刻・登場位置を高速に検出・追跡するための基礎検討を行なっている [1, 2, 3]。今後のマルチメディア処理技術において、時空間における指定対象の検出及び追跡技術は重要な要素技術の一つとなる。現在構築されつつある電子図書館のような大規模なマルチメディアデータベースやネットワーク上に分散するビデオデータなどにおいて指定の人物の検出・追跡技術はデータ検索エンジンに組み込まれて使用される。インターネットにおいて現在実現されている検索エンジンの検索機能はテキスト（文字列）に基づくが、将来は特定の人物、特定の言葉（キーワード）を検索する必要性が高まる。特に検索の高速化技術は検索エンジンの処理の高速化に直接関係し、ユーザーインタフェースにも影響を与える。

ビデオデータは音声・画像の時系列データとして与えられる。時空間系列において指定の特徴を持つ系列を効率的に検索する手法について検討する。文字列探索の手法は検索キーと検索対象（ファイル）の文字の完全一致を要求条件とする検索であり、これらについては既に幾つかの高速アルゴリズムが知られている。

音声などの特徴パターンを検索キーとする検索においては、パターンが自然に持つゆらぎや雑音の影響があるので文字列に対する完全一致探索を適用することはできない。従って完全一致探索を緩める部分文字列一致探索や文字間に歪みを導入する歪み許容探索や堅牢な特徴量や統計量に変換した歪み許容探索が要求されることになる。本研究では、パターンの持つゆらぎなどの歪みの特性とそれに応じた探索手法に関する関連付けを検討する。特にパターン歪みの幾つかの特性に対する高速化アルゴリズムを提案する。

関連する研究について述べる。米国では CMU などを中心としてビデオの画像及び音声を解析し、内容の情報検索の研究が進められている。Xerox Palo Alto 研究所、米国ベル研究所、IBM 研究所などの研究機関から音声・画像の区分化、indexing などの研究が発表されている [4, 5, 6]。一方国内においては龍谷大学においてニュース音声を対象としてアナウンサー発話区間の検出とアナウンサー音声区間の音声認識結果に基づくニュース話題検出の研究が行なわれている [7, 8]。また NTT 基礎研究所において音声・画像パターンの高速検索手法について研究されている [9, 10]。動画データにおけるシーンセグメンテーション処理の研究動向 [11]、音楽と音声をういたビデオデータ indexing の研究 [12] が報告されている。本研究はその高速化手法を包含し一般化・体系化すること、また話者区間の高速検索・区分化への応用を目的としている。

セグメントはフレーム列で定義され、膨大なフレーム

列から指定のセグメントの探索する手法について検討する。全数探索を省略し高速化する方法について幾つかの方法について述べ、音響信号からの話者セグメントの探索への応用について述べる。

2 セグメントの高速探索の分類

該当するセグメントを検索する場合には、探索に先だつてある程度の区間の分割を行なうことにより処理の高速化を図ることが可能である [14, 15]。文献 [13] において与えられた音声特徴に該当する音声ファイルを高速に探索する手法について検討した。一方、文献 [10] で該当するセグメントを高速に検出するための Active search 法の音響領域における適用について述べている。パターン探索を以下のように分類できる。

1. 全数探索と等価な手法
2. 計算の簡単化に基づく手法

第 1 の方法は全数探索と等価でありながら探索の性質を用いて計算量を省くことを目的としている。第 2 の方法は全数探索ではなく性能を劣化させることなく計算量の削減を目的としている。

一方、特徴量に基づく探索を行なう場合、検索モデル（検索キー）と入力セグメントの特徴が完全に一致することは期待できないので、歪みを許す探索とならざるを得ない。探索においては探索対象と与えられたキー（もしくは与えられたパターンから生成されるキー）との照合を行なうが、その照合においてパターンのゆれ・歪みを許容するかどうかによって以下のように分類される。

1. 歪み零探索（一致検索、文字列パターン照合）
2. 歪み ϵ 探索
3. 歪み許容探索

第 1 の方法は、文字列パターン照合であり歪みを許容しない歪み零探索であり、Knuth-Morris-Pratt 法（失敗回数に基づく方法）、Boyer-Moore 法（不一致文字法）、Rabin-Karp 法（ハッシュに基づく方法）などの高速化手法が知られている ([16, 17] を参照)。第 2 の方法である、文献 [10] で述べられた方法はパターンの若干の歪みを許容でき、歪みを許容するがその歪みは大きくないことを仮定する方法であり、歪みが零に近いという意味で歪み ϵ 探索と呼ぶことにする。第 3 の方法である文献 [13] で述べた方法は歪みを許す探索である。

3 セグメント探索の定式化

ベクトルの時系列 $V = v_1 v_2 \dots v_M$ に対して連続したベクトルをセグメントと呼ぶ。 V を長さ L のセグメント S の時系列とみなし、指定のセグメントを探索する。ここで、以下の処理過程からなっている。

1. 信号の標本化
2. 信号ベクトル変換 (特徴抽出)
3. ベクトル文字変換 (ベクトル量子化)
4. セグメント間の差異の定量化

第1の標本化においては探索の効率化のためにダウンサンプリングなども含まれる。第2の信号ベクトル変換は音声などの特徴抽出である。高速化のためには音声認識のための特徴抽出と一致させる必要はない。第3のベクトル文字変換ではベクトル量子化法を用いるが、2分木ベクトル量子化などを用いて高速化を図る。第4のセグメント間の差異の定量化は、歪みの定式化によって種々の手法が導出できる。

3.1 文字列探索法 (歪み探索)

歪みを許さない探索 (一致検索) として文字列探索の手法である Knuth-Morris-Pratt 法、Boyer-Moore 法、Rabin-Karp 法などが知られている。連続量であっても VQ を用いて離散量に、即ち、文字に変換することが可能である。ここで雑音やパターンのゆらぎが存在する時の文字列のロバスト性が保障されるのであれば、また文字列に誤りが発生しないという前提があるのであれば、全数探索の計算量が $O(ML)$ であるのに対して文字列探索手法は計算量が $O(M+L)$ で与えられるので非常に効率的である。

3.2 最長共通部分文字列探索法

文字列を扱う点では上で述べた手法と同様であるが探索すべき文字列の完全一致を求めるのではなく文字列間の最長共通部分文字列を求めることを目的としたアルゴリズムが LCS (Longest Common Sequence) 法 [18] として知られている。文字列に発生する誤りが極わずかである場合に有効な手段である。以下では基本 LCS 法について述べる。2つの文字列 A, B

$$\begin{cases} A = a_1 a_2 \dots a_m \\ B = b_1 b_2 \dots b_n \end{cases}$$

に対して部分列 $A(1:i), B(1:j)$ に対する LCS の値を $g(i, j)$ と表す時、図1に示すように以下の漸化式が成り

立つ。

$$g(i, j) = \begin{cases} g(i-1, j-1) + 1 & (a_i = b_j) \\ \max(g(i-1, j), g(i, j-1)) & (a_i \neq b_j) \end{cases} \quad (1)$$

初期化は以下のように境界部分に対して行なう。

$$g(i, 0) = 0, \quad g(0, j) = 0$$

対応する文字 a_i, b_j に対して以下のように一致行列を定義する。

$$d(i, j) = \begin{cases} 1 & (a_i = b_j) \\ 0 & (a_i \neq b_j) \end{cases}$$

LCS 法は一致行列をもとに動的計画法を用いて大局一致度を算出する類似な手法となる。通常の動的計画法では1文字に任意個数の文字を対応可能であるが LCS 法では文字の対応は1文字となる。この LCS 探索の高速アルゴリズムとしてハント・シマンズキ法が知られている [18](p.114)。LCS 法を用いてセグメント探索に関して部分文字列一致探索を行なうことが可能である。

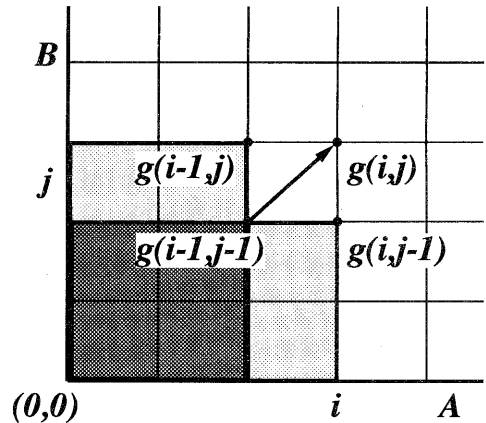


図1: LCS 漸化式の導出

4 Active 探索

4.1 類似度を用いた Active 探索法の導出

文献 [9, 10] では分布間の以下の類似度 S を用いて Active 探索法と呼ぶセグメント高速探索法を導出した。

$$S(p, p') = \sum_{m=1}^M \min(p_m, p'_m)$$

ここで、

$$0 \leq p_m, p'_m \leq 1, \quad \sum_m p_m = \sum_m p'_m = 1$$

であるので

$$0 \leq S(\mathbf{p}, \mathbf{p}') \leq 1$$

$$S(\mathbf{p}, \mathbf{p}') = 1 \iff \mathbf{p} = \mathbf{p}'$$

が成り立つ。

4.2 距離を用いた Active 探索法の導出

分布間の類似度を用いてセグメント高速探索法である Active 探索法を導出しているが、以下では幾何学的で理解しやすく一般化が容易である距離 (ノルム) を用いた導出法について述べる。

時系列上での長さ (フレーム, コマ) L のセグメント S に対して VQ 符号帳 $\mathbf{V} = \{v_1, v_2, \dots, v_M\}$ のベクトルの出現確率を $\mathbf{p} = (p_1, p_2, \dots, p_M)$ とし、出現度数 (頻度) を $\mathbf{h} = (h_1, h_2, \dots, h_M)$ とする¹。

$$\mathbf{p} = \frac{1}{L} \mathbf{h} \quad (p_m = \frac{h_m}{L} \quad (m = 1, 2, \dots, M)) \quad (2)$$

二つの確率分布間 \mathbf{p}, \mathbf{p}' の距離を確率ベクトル間のノルムで定義する。

$$d(\mathbf{p}, \mathbf{p}') = \|\mathbf{p} - \mathbf{p}'\|$$

ここでベクトル $\mathbf{x} = (x_m)$ のノルムとして以下を考える。

$$\begin{cases} \|\mathbf{x}\|_0 = \max_{1 \leq m \leq M} |x_m| \\ \|\mathbf{x}\|_1 = \sum_{m=1}^M |x_m| \\ \|\mathbf{x}\|_2 = \sqrt{\sum_{m=1}^M |x_m|^2} \end{cases} \quad (3)$$

これらのノルムには以下の大小関係が成り立つ。

$$\|\mathbf{x}\|_0 \leq \|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq M \|\mathbf{x}\|_0$$

右に 1 フレーム移動したセグメント S' の出現確率、度数を各々 \mathbf{p}', \mathbf{h}' とする。 S' は S の左端のフレームが削除され、新たに右端にフレームが追加されるので、削除された左端のフレームに対応する番号 m の度数 h_m が 1 だけ減少し、逆に追加された右端のフレームに対応する度数が 1 だけ増加することになる。減少と増加の符号ベクトルの番号が一致する場合には度数分布は変化しない。従って各々のノルムに対して以下の不等式を満たす。

$$\begin{cases} \|\mathbf{h} - \mathbf{h}'\|_0 \leq 1 \\ \|\mathbf{h} - \mathbf{h}'\|_1 \leq 2 \\ \|\mathbf{h} - \mathbf{h}'\|_2 \leq \sqrt{2} \end{cases} \quad (4)$$

一般的にセグメントの移動幅が N フレーム ($N < L$) の時には以下の不等式となる。

$$\begin{cases} \|\mathbf{h} - \mathbf{h}'\|_0 \leq N \\ \|\mathbf{h} - \mathbf{h}'\|_1 \leq 2N \\ \|\mathbf{h} - \mathbf{h}'\|_2 \leq \sqrt{2N} \end{cases} \quad (5)$$

¹VQ 符号帳を用いることは本質的ではない。

従って $d(\mathbf{p}, \mathbf{p}') = \frac{1}{L} \|\mathbf{h}' - \mathbf{h}\|$ であるのでその出現確率ベクトル間の距離は以下のように評価される。

$$\begin{cases} d_0(\mathbf{p}, \mathbf{p}') \leq N/L \\ d_1(\mathbf{p}, \mathbf{p}') \leq 2N/L \\ d_2(\mathbf{p}, \mathbf{p}') \leq \sqrt{2N}/L \end{cases} \quad (6)$$

一般に出現確率ベクトル間の距離が距離の公理 (対称性及び三角不等式) を満たす時、以下の不等式が得られる。

$$\begin{cases} d(\mathbf{p}, \mathbf{q}) + d(\mathbf{p}, \mathbf{p}') \geq d(\mathbf{p}', \mathbf{q}) \\ d(\mathbf{p}', \mathbf{q}) + d(\mathbf{p}, \mathbf{p}') \geq d(\mathbf{p}, \mathbf{q}) \end{cases} \quad (7)$$

これから次の不等式を得られる。

$$|d(\mathbf{p}, \mathbf{q}) - d(\mathbf{p}', \mathbf{q})| \leq d(\mathbf{p}, \mathbf{p}') \quad (8)$$

ベクトル \mathbf{x} を中心とする半径 r の球の領域を $B(\mathbf{x}, r)$ と表す。図 2 に示すように \mathbf{p}' が $B(\mathbf{p}, d(\mathbf{p}, \mathbf{p}'))$ の球上に存在するので、 \mathbf{p} と \mathbf{p}' が与えられた正数 T の $B(\mathbf{q}, T)$ に対して同一の領域 (球の内側もしくは外側) に属するためには、以下の不等式が十分である²。

$$d(\mathbf{p}, \mathbf{p}') \leq |T - D| \quad (9)$$

ここで $D = d(\mathbf{p}, \mathbf{q})$ である。この不等式と式 (6) から以下を得る。

$$\begin{cases} N_0/L \leq |T - D| \\ 2N_1/L \leq |T - D| \\ \sqrt{2N_2}/L \leq |T - D| \end{cases} \quad (10)$$

これより各々のノルムに対して移動幅 N_i ($i = 0, 1, 2$) に関する不等式を得る。

$$\begin{cases} N_0 \leq L|T - D| \\ N_1 \leq L|T - D|/2 \\ N_2 \leq L^2|T - D|^2/2 \end{cases} \quad (11)$$

N は L 以下であるので、 N_i の最大値は以下のように得られる。

$$\hat{N}_i = \max(L, N_i)$$

この式 (11) はフレーム (コマ) を幾つかの領域 (K) に分割して出現頻度・確率を算出する場合にも同様に成り立つ。 K 領域に分割するので式 (2) は以下のように変形される。

$$\mathbf{p} = \frac{1}{KL} \mathbf{h} \quad (12)$$

さらに式 (5) は以下で評価されるので、式 (6) は変わらない³。

$$\begin{cases} \|\mathbf{h} - \mathbf{h}'\|_0 \leq NK \\ \|\mathbf{h} - \mathbf{h}'\|_1 \leq 2NK \\ \|\mathbf{h} - \mathbf{h}'\|_2 \leq \sqrt{2NK} \end{cases} \quad (13)$$

² T は閾値に対応する

³ $\|\cdot\|_2$ に対しては分母に \sqrt{K} が残る。

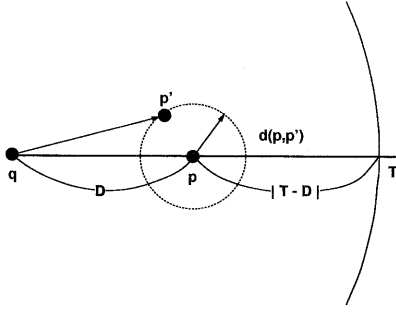


図 2: 球の包含関係

4.3 Fuzzy Active 探索法

以上の議論は VQ を用いた出現度数ではなく以下の式 (14), (15) で定義される Fuzzy VQ を用いた fuzzy membership 関数に拡張できる。これを Fuzzy 出現度数などと呼ぶことにする。Fuzzy VQ を用いる場合においても式 (5) が成り立つことを示せば良い。

入力ベクトル x に対する fuzzy membership 関数 w_l から算出される。式中で N_k は x の VQ 符号帳 V_N の k 近傍の符号番号の集合である。 F は fuzziness, k は近傍数、 d_l は v_l と x との距離であり、 M はフレーム数である。

$$p_l = \frac{K_l}{M}, \quad K_l = \sum_t w_{tl}, \quad (14)$$

$$w_l = \begin{cases} 1 / \left(\sum_{j=1}^k \left(\frac{d_l}{d_j} \right)^{\frac{1}{F-1}} \right), & l \in N_k, \\ 0 & \text{それ以外.} \end{cases} \quad (15)$$

右に 1 フレーム移動したセグメント S' の Fuzzy 出現確率、Fuzzy 度数を各々 p', h' とする。 S' は S の左端のフレームが削除され、新たに右端にフレームが追加される。削除される左端および追加される右端のフレームに対応する fuzzy membership 関数を $w_m^{(+)}, w_m^{(-)}$ ($m = 1, 2, \dots, M$) とする。この時 $h' = (h'_m)$ は以下の式で与えられる。

$$h'_m = h_m + w_m^{(+)} - w_m^{(-)} \quad (m = 1, 2, \dots, M) \quad (16)$$

ベクトルで表せば以下のようなになる。

$$h' = h + w^{(+)} - w^{(-)}. \quad (17)$$

ここで、 w', w は fuzzy membership 関数であるので、確率分布としての制約条件を満たす。

$$\sum_{m=1}^M w_m^{(+)} = \sum_{m=1}^M w_m^{(-)} = 1$$

従って以下の不等式を満たす。

$$\begin{cases} \|h - h'\|_0 \leq 1 \\ \|h - h'\|_1 \leq 2 \\ \|h - h'\|_2 \leq \sqrt{2} \end{cases} \quad (18)$$

この方法を Fuzzy Active 探索法と呼ぶことにする。通常の VQ を用いた出現確率を算出する方法は Fuzzy VQ を用いて出現確率を算出する方法において式 (14) の近傍数を $k = 1$ とする場合に相当する。従って、Fuzzy Active 探索法は Active 探索法の拡張となっている。

4.4 セグメント伸長決定への Fuzzy Active 探索法の適用

セグメント移動によるセグメント探索法について述べたが、これはセグメント伸長にも適用できる。簡単化するために $\|*\|_1$ の場合に限定して述べる。

セグメントを N フレーム伸長させるときの Fuzzy 頻度はフレーム数が N だけ増加するので、以下の式で評価される。

$$\|h - h'\|_1 = N \quad (19)$$

Fuzzy 出現確率は以下の式で与えられる。

$$p = \frac{1}{L} h, \quad p' = \frac{1}{L+N} h' \quad (20)$$

また各々のノルムは以下の値となる。

$$\|h\|_1 = L, \quad \|h'\|_1 = L + N$$

式 (6) は以下で評価される。

$$\begin{aligned} d_1(p, p') &= \left\| \frac{1}{L} h - \frac{1}{L+N} h' \right\|_1 \quad (21) \\ &= \frac{1}{L(L+N)} \sum_m |(L+N)h_m - Lh'_m| \\ &\leq \frac{1}{L(L+N)} \sum_m L|h_m - h'_m| + \frac{1}{L(L+N)} \sum_m N h_m \\ &= \frac{1}{L(L+N)} LN + \frac{1}{L(L+N)} NL = \frac{2N}{L+N} \end{aligned}$$

これより、式 (9) に代入して $|T - D| < 2$ の条件の元で伸長幅に対する以下の不等式を得る。

$$N \leq \frac{|T - D|}{2 - |T - D|} L \quad (22)$$

x が確率ベクトルである時は以下の等式が成り立つ。

$$\|x\|_1 = \sum_{m=1}^M |x_m| = 1$$

従って閾値 $T \leq 1$ となり、以下の不等式を得る。

$$|T - D| < 2$$

5 話者セグメントの高速探索

話者セグメントの探索を行なうためには音声セグメントの長さ(継続時間長)が一定以上であることが必要である。そこでその長さを L とする。探索する話者の VQ 符号帳を V とし、探索対象区間 X を以下で表す。

$$X = x_1 x_2 x_3 \dots x_M$$

X の長さ L のセグメント(部分区間) $X(t+1:t+L)$ に対して探索は以下の二つの方法が考えられる。

1. 分布間の差異に基づく方法 (VQ histogram 法)
2. VQ 歪み法

第1の方法では部分区間から VQ histogram を作成しあらかじめ与えられた histogram との間の歪みが閾値 T 以下である時、そのセグメントはその話者の発話であるとする。この手法は第4章で述べた方法が適用できる。第2の方法では VQ 歪みが閾値 T 以下である時、そのセグメントはその話者の発話であるとする。

$$D(X(t+1:t+L), V) = \frac{1}{L} \sum_{i=1}^L d(x_{t+i}, V)$$

$$D(X(t+1:t+L), V) < T$$

第2の方法に対して高速化のために以下の2つの方法がある。

5.1 記号列探索的手法

基本的な考え方は図3に示すように

$$D(X(t+1:t+L), V) = \frac{1}{L} \sum_{i=1}^L d(x_{t+i}, V)$$

に対して終端からの部分和を計算し閾値と比較し、これを越える場合には和の計算を打ち切り探索開始位置を移動する。

$$d(n) = \sum_{i=1}^n d(x_{t+L-i}, V)$$

記号列探索において用いられるように、一致しない探索が多く発生する場合には後端から探索を開始し不一致条件を検出しそれを満たす範囲で移動を行なう。

この方法が有効であるのは L が大きく、開始する後端の早い探索において不一致条件を満たすことである。

具体的な手順を以下に示す。

Step 0 開始位置の設定

$$t = 1$$

Step 1 初期設定

$$d = 0, i = L$$

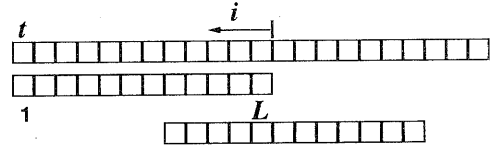


図3: 記号列探索的手法の概念

Step 2 部分和の計算

$$-- i; d += d(x_{t+i}, V)$$

Step 3 条件判定

1. 以下の不等式を満たす時、Step 4 に。

$$d \geq LT$$

2. そうでなければ $i > 0$ であれば Step 2 に。
3. セグメント検出成功であり

$$t ++$$

として Step 1 に。

Step 4 探索位置 t の移動

$$t = t + L - i + 1$$

Step 1 に。

Step 3 の条件判定の3において検出成功の場合、探索開始位置の決定について式(23)で述べる Active 探索を用いることも可能である。

5.2 Active 探索的手法

$D(t) = D(X(t+1:t+L), V)$ とおく時、 $D(t)$ と $D(t-1)$ との関係を用いる方法である。

$$LD(t) = LD(t-1) + d(x_{t+L}, V) - d(x_t, V)$$

$$|D(t) - D(t-1)| \leq \frac{1}{L} |d(x_{t+L}, V) - d(x_t, V)| \leq \frac{\delta}{L}$$

ここで δ は以下で定義される。

$$\delta = \max_i d(x_i, V) = \max_i \min_n d(x_i, v_n)$$

従って一般に N フレーム移動する時には以下の不等式を得る。

$$|D(t+N) - D(t)| \leq \frac{\delta N}{L}$$

従って以下の不等式を満たせば良いことになる。

$$\frac{\delta N}{L} \leq |T - D(t)| \quad N \leq \frac{L|T - D(t)|}{\delta} \quad (23)$$

6 むすび

音声・画像の時系列データとして与えられるビデオデータにおいて指定された人物の登場時刻・登場位置を高速に検出・追跡するためのセグメントの高速探索について述べた。パターンが自然に持つゆらぎや雑音の影響を考慮した完全一致探索を緩めた部分文字列一致探索や歪み許容探索について述べた。今後は実際の音声データを用いて提案した手法の評価を行なう。

謝辞

本研究は会津大学研究プロジェクトの研究の一貫として行なわれたものである。会津大学 野口正一学長をはじめとして本研究プロジェクトを支援して下さる学内の皆様に感謝いたします。また有益な助言をいただくプロジェクトのメンバーの林隆史助教授、R. Huang 講師、J. Ma 講師、また日頃有益な討論をいただくヒューマンインタフェース学講座の諸氏に感謝します。

参考文献

- [1] T.Asano, M.Sugiyama, Object Location and Tracking in Video Data, SPECOM98 (Oct. 1998).
- [2] 杉山, ビデオデータにおける指定人物の検出と追跡 — 音声部分を用いた処理 —, 映像情報メディア学会ヒューマンインフォメーション研究会, Vol.22, No.66, pp.7-12(1998-11).
- [3] T.Asano, M.Sugiyama, Segmentation and Classification of Auditory Scenes in Time Domain, Proc. of IWHIT98, pp.13-18 (1998-11).
- [4] J.S.Boreczky, L.D.Wilcox, A Hidden Markov Model Framework for Video Segmentation Using Audio and Image Features, ICASSP98 Vol.6, p.3741 (1998).
- [5] G.Potamianos, H.P.Graf, Discriminative Training of HMM Stream Exponents for Audio-Visual Speech Recognition, ICASSP98 Vol.6, p.3733 (1998).
- [6] R.Mohan, Video Sequence Matching, ICASSP98 Vol.6, p.3697 (1998).
- [7] 西田, 有木, 自動学習による話者セグメンテーション, 音声研資, SP97-57, pp.1-6 (1997-11).
- [8] 鷹尾, 緒方, 有木, ニュース音声に対するトピックセグメンテーションの検討, 音学講論, 1-R-20, pp.157-158 (1998-09).
- [9] G.A.Smith, H.Murase, K.Kashino, Quick Audio Retrieval Using Active Search, ICASSP98 Vol.6, p.3777 (1998).
- [10] 柏野, G.A.Smith, 村瀬, 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, 音学講論, 2-9-24, pp.561-562 (1998-09).
- [11] Tsuhan Chen, The Past, Present, and Future of Image and Multidimensional Signal Processing, IEEE Signal Processing Magazine, Vol.15, No.2, pp.21-58 (March 1998).
- [12] K.Minami, A.Akutsu, H.Hamada, Y.Tonomura, Video Handling with Music and Speech Detection, Special Issue on Multimedia and Music in IEEE Multimedia, pp.17-25 (1998).
- [13] 杉山, 音声特徴 key による音声検索の高速化, 音響学会講演論文集, 3-P-19, pp.199-200 (1995-03).
- [14] 浅野, 杉山, 音響シーンセグメンテーション法の検討, 2-Q-12 (1998-03).
- [15] 杉山, 音響シーンセグメントの接続の検討, 音学講論, 1-R-19 (1998-9).
- [16] R.Sedgewick, Algorithms in C, Addison Wesley.
- [17] 大森, 他, アルゴリズムの基礎, 共立出版.
- [18] 疋田, C で書くアルゴリズム, サイエンス社.