

特異画像の解析を基にした周期的動作の記述

鈴木 貴彦 吉野 和芳 川嶋 稔夫 青木 由直

giants@huie.hokudai.ac.jp

北海道大学 工学部

〒060 札幌市北区北13条西8丁目

あらまし 本稿では、人物の周期的動作の特徴的シーンに注目して、動作の記述を行う方法を提案する。特徴的シーンの定義は、「画像内で人物の体の一部分と他の部分が接触する、離れる、人物がほぼ静止して見えるシーン」である。人物の抽出は、スライス画像に対してのフィルタ処理で行う。動作の特異状態の検出では、隣接2フレームにおける画素の幾何学的配置に基づいた境界線追跡を適用する。一連の周期的動作の意味論的分割、特徴的シーンの検出では、特異状態の出現に周期性があるという性質を利用する。実験では、ラジオ体操を解析し、動作のダイジェスト化が可能なことを示す。

Description of Periodic Motion Based on Singular Image Analysis

Takahiko Suzuki, Kazuyoshi Yoshino, Toshio Kawashima and Yoshinao Aoki

giants@huie.hokudai.ac.jp

Faculty of Engineering, Hokkaido University.

Kita-13, Nishi-8, Kita-ku, Sapporo, Hokkaido, 060, Japan

Abstract We propose a method for describing human motion while concentrating on singular scenes in his periodic motion. We define the following scenes as singular ones; the scenes indicate that in an image one part on his body touches the other part or separates from it and that he seems still. Human are detected by filter process of slice-images. Singular states in motion are detected by boundary line tracing based on geometrical arrangement of pixels in 2 successive frames. A series of motion is divided into semantic groups and then singular scenes are detected using the characteristics that singular states appear periodically. In our experiment, we analyzed radio gymnastics. Finally, we describe that it is possible to digest human motion.

1 はじめに

コンピュータビジョンにおいて、人物等の対象物の動きの理解は動画像解析の重要なテーマの一つとなっている。最近では、対象物の動作を簡潔に記述する、動画像から必要とするシーンを検索する研究が盛んになってきており、今後、重要性が増すと考えられている。

我々の日常、身のまわりにおいて、人物や動物、物体は様々な動きをする。動きはランダムである非周期的な動きと規則性のある周期的運動に大別される。人物の行動を見ると、意外にも周期的動作は多く、歩行、走行、ラジオ体操、踊り等がこれに該当する。これまでに、[1], [2] 等で人物の周期的動作の解析が行われてきた。

本研究では、人物の様々な動作の中で、周期的動作に注目し、その動作のダイジェスト化を行うことを目的としている。これまでに、[3] では、人物の関節等にマークを付け、そのマークの軌跡を周期的動作と関連づけるという研究が紹介されている。一方、本研究では、動作のダイジェスト化を行うために、周期的動作を時系列画像としてファイルに格納し、動作の特異状態に基づいて動作の周期性を調べ、一連の動作を意味論的に数区間の異なる動作に分割していく。そして、その中から特徴的シーンを検出することを行なっている。この結果、特徴的シーンを列挙し、何らかの補足説明(脚注)を加えることにより、動きのダイジェスト化が可能になると考えられる。ラジオ体操を例にとると、「『手足の運動』が4回、『腕を回す運動』が4回、……」と表現できるようになる。

本研究では、人物の動作の周期性を調べ、特徴的シーンを検出する際、動きの特異状態に注目する。ここで、特異状態とは、画像内で人物の体の一部分と他の部分が接触した、離れた状態、人物がほぼ静止して見えると考えられる状態であると定義する。同じ周期的動作内では、同種の動きの特異状態(接触、静止)の出現は周期的であるという現象を利用し、一連の動作を示す時系列画像を意味論的に異なった周期をもつ動作に分割する。その後、分割された動作の中から1周期分を取り出し、再び、特異状態の情報を利用して、動作を記述する。適切に、特徴的シーンが検出されたならば、動画像の中から、我々が必要としているシーンを検索することが可

能になると考えられる。

以下、第2節では、前処理として、動画像から対象物を抽出する方法について述べる。第3節では、動きの特異状態の検出方法について述べ、第4節では、動きの周期性を見出し、動画像を意味論的に異なる動作に分割する方法について述べる。第5節では、分割された動作の中から特徴的シーンを検出する方法について述べる。第6節では、ラジオ体操のシーンをを用いて、本手法による実験結果を示し、第7節では、まとめを行い、今後の課題を述べる。

2 対象物の抽出方法 — 前処理

動画像から対象物である人物を抽出するために、本研究では、まず、動画像列(フレーム画像)をXYT 3次元 voxel 空間とみなし、複数のXT スライス画像を作成する(図1)。

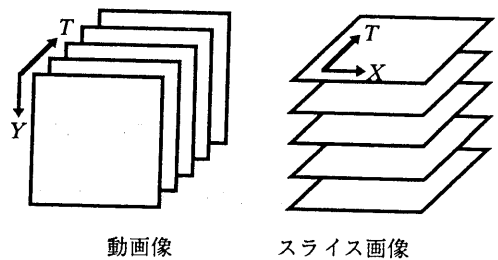


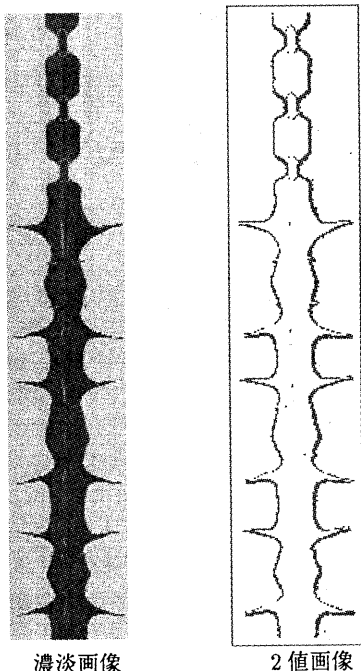
図1: スライス画像の作成

次に、各XT スライス画像にLoGフィルタ(Laplacian-of-Gaussian filter)

$$f_{LoG}(x, t, \sigma_x, \sigma_t) = -\frac{1}{2\pi\sigma_x\sigma_t} \left\{ \frac{\sigma_t^4 t^2 + \sigma_x^4 x^2}{(\sigma_x\sigma_t)^4} - \frac{\sigma_t^2 + \sigma_x^2}{(\sigma_x\sigma_t)^2} \right\} \times \exp \left\{ -\frac{1}{2} \left(\frac{t^2}{\sigma_t^2} + \frac{x^2}{\sigma_x^2} \right) \right\} \quad (1)$$

(x, t : フィルタ内の座標; σ_x, σ_t : 偏差)

を施して、人物の輪郭を抽出する。CCDカメラの位置を固定した場合、背景模様的位置は時間が経過しても変化しないため、XTスライス画像においては、この背景模様は時間軸Tに平行となって



(フィルタサイズ $X = 15, T = 9,$
 $\sigma_x = 1.2, \sigma_t = 0.9,$ 閾値 10.0)

図 2: LoG フィルタの適用例

現れる。この性質を利用して、可能な限りノイズを除去しながら、第 3 節で述べる特異状態の検出を行い易く、人物の輪郭を抽出できるような LoG フィルタを施す。その後、閾値処理を施して、 XT スライス画像を 2 値化する。

この処理を全ての XT スライス濃淡画像に対して行うことで、 XYT 3 次元 voxel 空間における人物の輪郭の抽出が可能となる。

本手法では、式 (1) で示した LoG フィルタに対して、フィルタサイズ X 方向=15, T 方向=9, 偏差 $\sigma_x = 1.2, \sigma_t = 0.9$ として、人物の輪郭抽出を行っている。

ある XT スライス画像に対して LoG フィルタを適用した例を図 2 に示す。

3 動きの特異状態の検出方法

3.1 接触する状態、分離する状態の検出方法

第 1 節でも少し述べたが、本研究では、画像内において

- 対象物の一部分が他の部分と接触する状態
- 対象物のある部分が分離する状態
- 対象物がほぼ静止していると考えられる状態

を特異状態と定義する。

接触する状態、分離する状態を XYT 3 次元 voxel 空間において時間軸 T 方向に見ると、独立して動作している部分は図 3 に示される特異領域に向かって収束、あるいは特異領域から分岐して行くと考えられる。この現象を XYZ の 3 次元物体として捉えるならば、特異領域は鞍型構造として表現される。従って、本手法では、この鞍型構造の検出を行い、対象物が接触する状態、あるいは分離する状態を定義する。

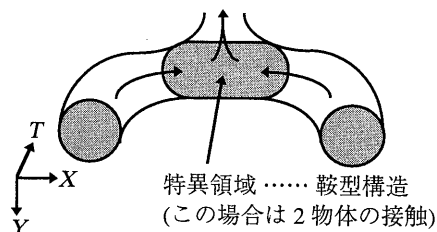


図 3: 特異領域の解説図

鞍型構造の検出には隣接する 2 フレームを使用する。 XT スライス画像から対象物が抽出された時点 (第 2 節) で、画像は 2 値画像になっている。図 4 のように現フレーム、前フレームにおける対象物を構成する画素の幾何学的配置により画素値が 1 (現フレームにおける 1-画素), 2 (前フレームにおける 1-画素), その他が 0 である画像を作成し、1-画素の連結成分において境界線追跡を行う。このとき、0-画素から 2-画素へ、2-画素から 0-画素へと変わる回数 c を数える。 $c = 4$ のとき、3 次元物体は、局所的な部分で鞍型を構成、すなわち特異領域を構成していることが知られている (濃淡画像に

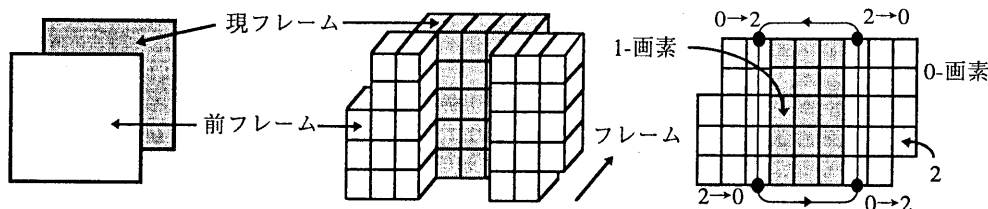


図 4: 対象物の接触の状態, 分離の状態の検出

おける画素状態の分類でも同様の定義がある) [4]。従って, 本手法では, $c=4$ となるフレーム画像を探せばよいことになる。

動画を調べた結果, 特異領域があると判定されたとき,

- (a) 特異領域を含むフレーム番号 F_i
- (b) 接触した状態であるか, あるいは分離した状態であるか? — 接触 or 分離
- (c) 境界線追跡で考慮したフレーム画像内での座標 (x, y)

の情報をファイルに保存しておく。この情報は, 第 4 節, 第 5 節の過程で使用する。

3.2 静止状態の検出方法

人物の動作を簡潔に記述する場合, 対象物がほぼ静止して見える状態にも注目する必要があると我々は考えている。複数の周期運動が連続する場合, 動作の変わり目として捉えることもできるためである。また, 第 3.1 節で述べた特異領域は存在しないが, 振り子の運動において, 速度が 0 になり, 運動方向が反対になるというシーンに注目すべきであるという考えもある [5]。静止状態を検出するために, 本手法では, 隣接する 2 フレーム (2 値画像) の差分をとる。その結果, 差が極めて小さい場合, 対象物がほぼ静止状態にあると定義する。

る状態) の検出結果 (a)~(c) を利用する。図 5 に示すように, フレーム F_A とフレーム F_B において, (b) の状態が一致し, (c) の座標値が類似しているかを調べ, フレーム内の人物の類似度を求める。類似度が低ければ, F_A と F_B の画像をさらに細かく調べる必要がないため, 計算時間の短縮へとつながる。ここに, 特異状態の情報を使うメリットが見られる。類似度を利用する考えは同じ動作が周期的に続くならば, 特異領域を有するフレームの前後数フレームにおいて, 動物体の幾何学的形状はほぼ同じであり, 特異領域に関する情報も周期的に出現するという仮説に基づいている。周期的動作では, ある特異領域を含むフレームを見た場合, その 1 周期前, 2 周期前, ..., 1 周期後, 2 周期後, ... で類似度が高くなる。これにより, 周期性が検出される。

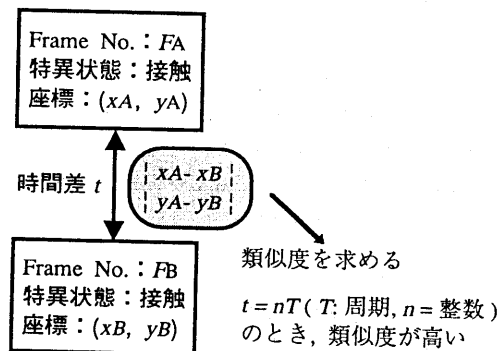


図 5: 動作の周期性の有無の検討

4 動作の周期性検出と分割

4.1 動作の周期性検出

人物の動作の周期性を調べるために, 本研究では, 第 3.1 節で得た特異状態 (接触する状態, 分離す

ただし, (a) のフレーム番号は, 第 2 節で述べた LoG フィルタの適用により, XT スライス画像の平滑化が行われるため, 実際の特異状態に相当するフレーム番号とは若干異なる。接触する状態は,

実際の接触より数フレーム前で、分離する状態は、実際より数フレーム後で検出される。(c)のフレーム内での座標についても同様のことがいえる。

4.2 動作の分割

第 4.1 節で得られたフレーム間の類似度を基に、周期的動作の周期を求め、一連の動作を意味的に異なる動作に分割する。本研究では、フレーム F_A からフレーム F_B までの区間のフレーム画像 (枚数は $F_A - F_B - 1$ 枚) において、どの位の類似性があるかを調べる。

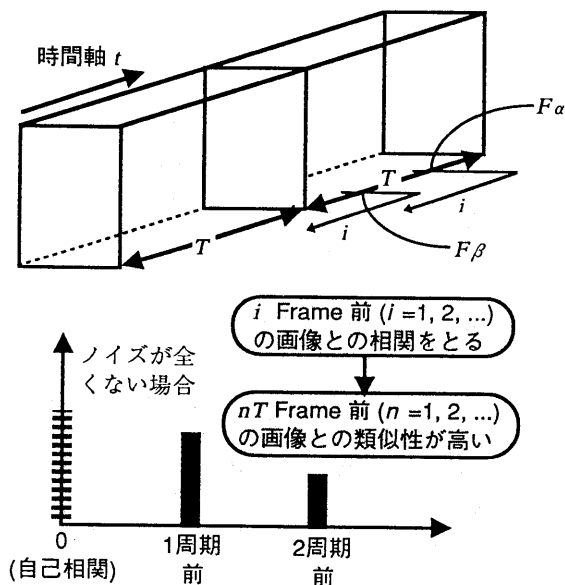


図 6: フレーム間の相関

図 6 のように、注目しているフレーム F_α (または F_β) の過去 i フレームとの相関をとる。理論的には、1 周期前、2 周期前、... のフレームのシーンとの相関が高くなる。これをヒストグラム表示すると、図 6 (下) のようになる。ピーク間の距離が動作の周期となるので、この距離とピークの本数によって、一連の動作を分割する。

5 特徴的シーンの検出

分割された動画の意味的に異なる動作の 1 周期分の中から、特徴的シーンを検出する。第 3.1 節で述べたように、画像内で対象物の一部分と他の部分が接触したシーン、離れたシーン、対象物がほぼ静止して見えるシーンを特徴的シーンとする。特異状態を含むフレームを列挙した場合、特徴的シーンは多過ぎる。しかし、本研究では、特徴的シーンは周期的に出現するので、第 4.1 節で調べた相関の度合も加味する。相関の高いフレーム画像のみを考慮し、これを列挙する。相関の低いフレーム画像は特徴的シーンの候補から除外する。これにより、特徴的シーンは適切と思われる枚数に限定される。尚、1 周期分のフレーム画像を列挙するため、周期の最初のフレーム画像、最後のフレーム画像も特徴的シーンに加える。

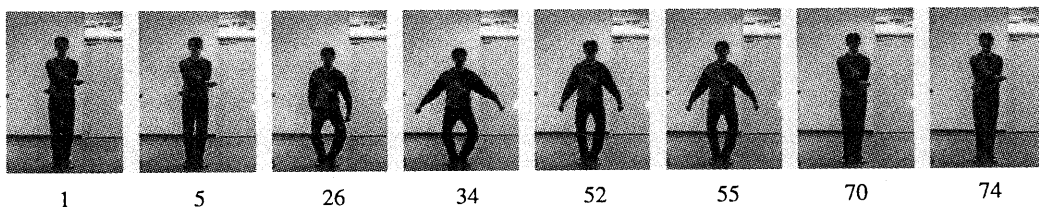
6 実験結果

6.1 準備

本研究では、実験用画像として、ラジオ体操第一の最初の方のシーンを使用した。動画は「『手足の運動』を 3 回、『腕を回す運動』を 3 回」含んでいる。実験 1 では、 140×200 のカラー画像をフレーム間隔 1/30 秒で計 755 フレーム使用した。背景は主に単色で、ノイズは比較的少ない画像である。実験 2 では、同様の動作を背景が比較的複雑で異なる場所で撮影し、実験 1 と同サイズの画像を計 655 フレーム使用した。画像のサイズを全く同じ規格にし、同じラジオ体操のシーンを採用したのは、今後の課題でもあるが、動画像中から必要としているシーンの検索を行いたいためである。2 実験とも、まず、200 枚の XT スライス画像 (濃淡画像) を作成した。次に、 LoG フィルタについての各パラメータを以下のように設定し、フィルタ処理を行なった。

- フィルタサイズ: X 方向...15 画素分, T 方向...9 画素分
- 偏差: X 方向 $\sigma_x \cdots 1.2$, T 方向 $\sigma_y \cdots 0.9$
- 2 値画像作成における閾値: ± 10.0

【前半】 — 「手足の運動」 → 反復回数 3 回



【後半】 — 「腕を回す運動」 → 反復回数 3 回

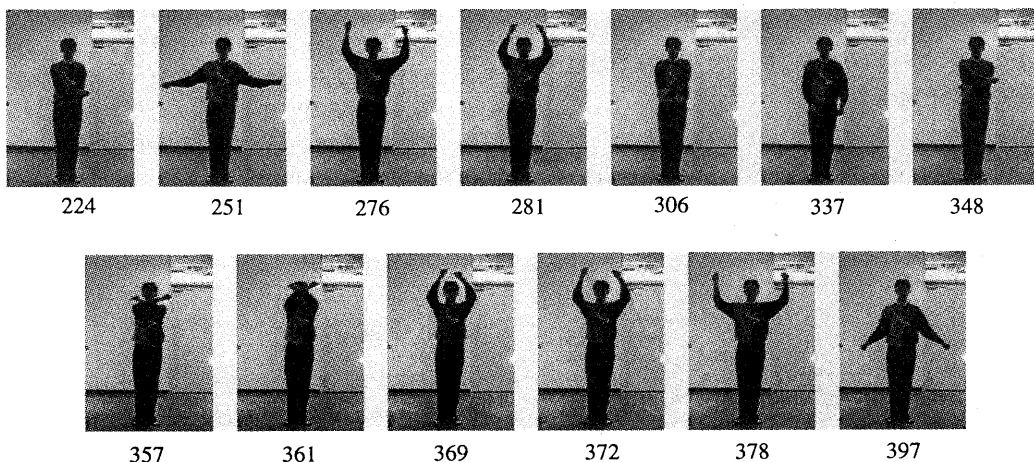


図 7: 実験 1 — ラジオ体操の特徴的シーン (背景が比較的単色)

その後、人物の輪郭抽出の結果を基に、動きの特異状態の検出を行なった。

6.2 実験結果および考察

実験 1

特異状態の検出結果を基に、フレーム間の相関を求め、ラジオ体操の周期的動作の識別、記述を行なった。

その結果、

- 『手足の運動』の周期は 73, 反復回数 3 回
- 『腕を回す運動』の周期は 176, 反復回数 3 回

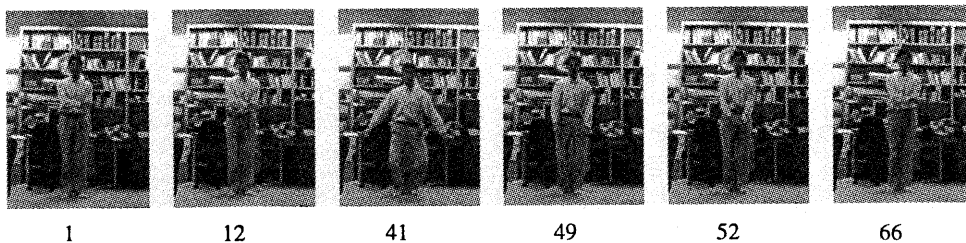
と求められた。

次に、二つの運動の各 1 周期に注目し、特徴的シーンを検出した。その結果をまとめると、図 7 のようになった。画像の下は数字は対応するフレーム番号を示している。この結果を見ると、ラジオ体操の概略を捉えることが可能になることがわかる。所々のフレームにおいて、両腕と両脚が離れているにも関わらず、本手法では、特徴的シーンとして検出されている。これは人物の抽出のために *LoG* フィルタを施した結果、*XT* スライス画像の平滑化が行われ、接触のシーン、分離のシーンが実際の動作のシーンより各々、数フレーム前、数フレーム後で検出されたことによると考えられる (第 4.1 節)。

実験 2

実験 1 で扱った画像より背景が複雑なため、特

【前半】 — 「手足の運動」 → 反復回数 3 回



【後半】 — 「腕を回す運動」 → 反復回数 3 回

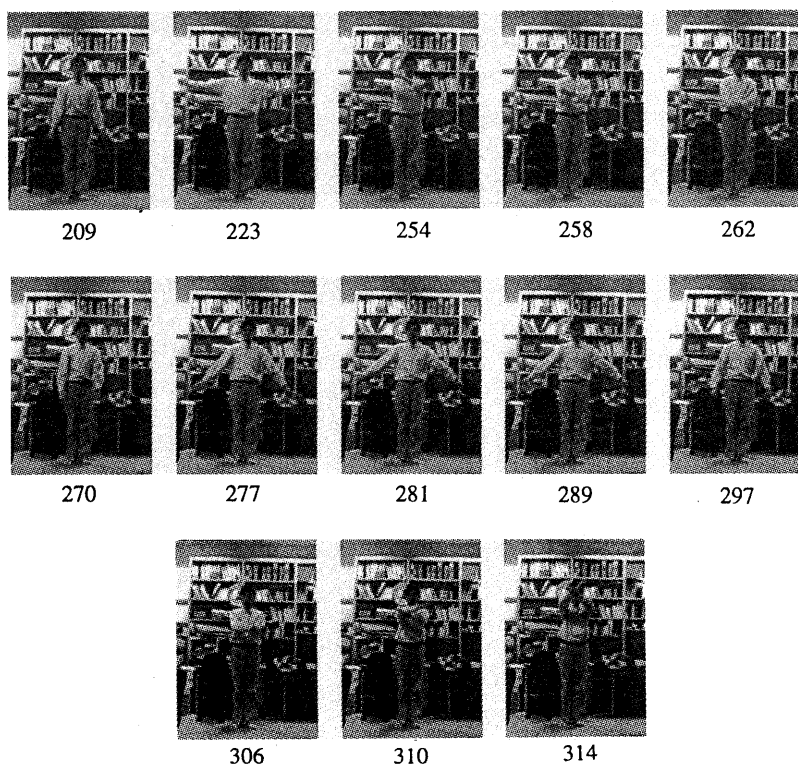


図 8: 実験 2 — ラジオ体操の特徴的シーン (背景複雑下)

異状態の検出に倍近くの時間を要した。XT スライス画像に LoG フィルタを施して 2 値化を行う際、適切に背景が除去されず、ノイズとして残ったためと考えられる。動作の周期的識別、記述を 実験 1 同様、行なった。

その結果、

- 『手足の運動』の周期は 66, 反復回数 3 回
- 『腕を回す運動』の周期は 106, 反復回数 3 回

と求められた。

特徴的シーンの検出結果は図 8 のようになった。ノイズが多いにも関わらず、特徴的シーンのある程度、適切に取り出せたと思われる。両腕と両脚が離れているシーンがあるが、実験 1 と同様の理由で特徴的シーンとして判定されている。前半の動作については、実験 1 と同様、要求していたシーンが検出された。しかし、後半の動作においては、頭上で両腕が交差するシーンが欠落している。この原因は LoG フィルタを用いて XT スライス画像の 2 値化を行なったとき、本棚の本が除去されず、ノイズとして残ってしまったためと考えられる。また、後半の動作については、周期の測定を誤ったため、動作が途中で切れるという結果を導き出したと考えられる。

7 おわりに

7.1 本稿のまとめ

本稿では、人物の周期的動作の特徴的シーンをを用いて、動作の記述を行う方法を提案した。本手法により、ほぼ理想的な特徴的シーンが検出された。また、特徴的シーンの列挙により、動作のダイジェスト化が可能であることを示した。

7.2 今後の課題

今後の課題として、

- 動画の中から必要としているシーンを検索する場合、背景の複雑さに依存せず、どのように人物の動作のみを理解し、抽出するか?
- テレビのラジオ体操のように、途中でカメラの撮影方向が変わった場合、どのように動作を理解し、シーンの検索へ結びつけるか?

- 周期的動作の中に非周期的動作が含まれる場合、非周期的動作をどのように検出、理解するか?

などが挙げられる。

本研究の一部は文部省科学研究費補助金 課題番号 07207201 の補助のもとに行われた。

参考文献

- [1] S.A.Niyogi and E.H.Adelson : “Analyzing and Recognizing Walking Figures in XYT”, Proc. IEEE CVPR '94, pp.469-474 (1994.6)
- [2] R.Polana and R.Nelson : “Detecting Activities”, Proc. IEEE CVPR '93, pp.2-7 (1993.6)
- [3] Lee Campbell and Aaron Bobick : “Recognition of Human Body Motion Using Phase Space Constraints”, M.I.T Media Lab. Perceptual Computing Section Technical Report No. 309 Abbreviated version appears in ICCV '95
- [4] 鳥脇 純一郎 : “ディジタル信号処理シリーズ — 画像解析のためのディジタル信号処理 [II]”, 昭光堂
- [5] 鈴木 貴彦, 吉野 和芳, 川嶋 稔夫, 青木 由直 : “芯線構造に基づく動作の理解”, 平成 6 年度 電気関係学会北海道支部連合大会 講演論文集, pp.442 (1994.10)
- [6] 鈴木 貴彦, 吉野 和芳, 川嶋 稔夫, 青木 由直 : “特異画像の解析による周期的動作の識別”, 1995 年 電子情報通信学会総合大会 講演論文集 情報・システム (2) D-626, pp.352 (1995.3)