

# 移動物体の軌跡に基づく最小数のキーフレーム選択手法

北澤 仁志<sup>†</sup> 小野澤 晃<sup>††</sup>

デジタルカメラの高性能化、低価格化、ネットワークの高速化により多くの固定モニタカメラが用いられようになってきた。しかし、これらを見るには多くの時間を要する。多量のモニタ映像を有効に活用するには重要な部分や注目物体が写っている部分だけを抽出する技術が必要となる。固定カメラ映像の場合は基本的に重要な部分は移動物体がある場面に限られるが、この部分を見るだけでもまだ長時間を要する。本研究では固定モニタカメラ映像を対象とし、移動物体を抽出した後すべての移動物体を含み、かつ、移動軌跡の特徴が反映されるような最小数のキーフレームを抽出する手法を提案する。

## Minimum Number of Keyframe Selection Method Based on Trajectories of Moving Objects

HITOSHI KITAZAWA<sup>†</sup> and AKIRA ONOZAWA<sup>††</sup>

Due to the development of the broad-band networks and cheap video camera, fixed monitoring cameras are increasing rapidly. Since it takes lots of time to survey the captured video images, it is inevitable to develop techniques for extracting only "important" scenes from them. In the case of monitoring fixed camera images, the "important" scenes can be defined as those with moving objects, which may still require long time to be surveyed. In this paper, in order to survey the video images in a short time, novel techniques for extracting the minimum number of key frames which reflect the feature of objects' trajectories are proposed.

### 1. はじめに

デジタルカメラの高性能化、低価格化、ネットワークの高速化により多くの固定モニタカメラが用いられ多量の映像が容易に手に入るようになってきた。しかし、これを見る人の時間には限りがあり、モニタ映像を有効に活用するには重要な部分や注目物体が写っている部分だけを抽出する技術が必要となる。

映像の要約やキーフレームの抽出に関しては多くの研究がなされてきた<sup>1)~4)</sup>。これらは主として映画やニュース番組、スポーツ番組など一度人手で編集された映像を対象とするものであり、ヒストグラムの変化からシーンの切替わりを抽出しこれをもとにキーフレームを抽出する方法が用いられている。しかし、この方法で抽出されるキーフレームに重要なシーンがす

べて含まれている保証はない。動画像において重要な情報を持つのは移動物体であるが、移動するカメラの映像から移動物体を正しく抽出することは容易ではない。一方、固定カメラ映像については移動物体を抽出、認識する種々の方法が研究されてきた<sup>5)~9)</sup>。これらの手法を利用して移動物体のある部分のみを抽出すれば映像の内容を把握するために要する時間は大きく短縮できる。しかし24時間のモニタ映像において移動物体が写っている部分が10%だとしても全部を見るにはまだ2時間以上を要する。

本研究では、比較的人や車の通行の少ないところに設置された固定モニタカメラ映像を対象とし、移動物体を抽出した後すべての移動物体を含み、かつ、移動軌跡の特徴が反映されるような最小数のキーフレームを抽出する手法を提案する。

まず2章で全移動物体を含みかつ最小数のキーフレームを抽出する方法を示す。次に3章で軌跡の特徴が分かるように1つの移動物体について複数のキーフレームを抽出し、かつ、キーフレームの総数を最小にする方法を述べる。また4章では、一部の区間を選択から除外したり排他的な区間選択を行うなど、キーフ

<sup>†</sup> 東京農工大学工学部

Tokyo University of Agriculture and Technology

<sup>††</sup> 日本電信電話株式会社 NTT マイクロシステムインテグレーション研究所

NTT Microsystem Integration Laboratories, NTT Corporation

レームの選択に種々の条件を設ける方法を示す. 5章で提案手法の具体的計算方法を説明し, 6章に実行例を示す.

## 2. 全移動物体を含む最小数のキーフレームの選択

移動物体の存続区間を基に, すべての移動物体がいずれかのフレームに含まれるような最小数のフレームの集合を求める方法を示す. 本手法ではまず映像を移動物体の存続区間によりいくつかのフレーム区間に分け, すべての移動物体がいずれかに存在するような最小数のフレーム区間を選択し, 次に, 選択したフレーム区間からそれぞれ1フレームずつを抽出するという2段階の選択手法をとる.

### 2.1 記号の定義

動画像中の  $i$  番目のフレームを  $f_i$  とする. 映像が  $n$  個のフレームにより構成されるとき,  $f_1, f_2, \dots, f_n$  は時刻の昇順にソートされているものとする.

$s \leq e$  のとき  $f_s$  番目から  $f_{e-1}$  番目までのフレーム区間を

$$F[f_s : f_e] = \{f_s, f_{s+1}, \dots, f_{e-2}, f_{e-1}\} \\ = \{f_i | s \leq i < e\} \quad (1)$$

と記述する.  $F[f_s : f_e]$  は開始フレーム  $f_s$  を含むが終端フレーム  $f_e$  は含まない.

移動物体  $O^j$  がフレーム  $f_{s_j}$  から  $f_{e_j-1}$  の間存在するとき  $O^j$  の存続フレーム区間を

$$F^{O^j} = F[f_{s_j} : f_{e_j}] \quad (2)$$

と記述する. これを以下オブジェクト・フレーム区間 (Object Frame Division) と呼び, OFD と略記する.

図1のように  $k$  個の物体が存在するとする.  $f_{s_j}, f_{e_j} (1 \leq j \leq k)$  を昇順にソートし,  $p$  番目のフレームを  $f^p$  としたとき, 全映像で通し番号を付けたフレーム区間  $D_p$  を次式のように定める.

$$D_p = F[f^p : f^{p+1}] \quad (3)$$

これを以下グローバル・フレーム区間 (Global Frame Division) と呼び, GFD と略記する.

### 2.2 最小数キーフレームの選択

物体  $O^j$  の存続区間が  $D_{u_j}, D_{u_j+1}, \dots, D_{v_j} (u_j < v_j)$  に分割されたとき, すなわち

$$F^{O^j} = F[f_{s_j} : f_{e_j}] = D_{u_j} \cup D_{u_j+1} \cup \dots \cup D_{v_j} \quad (4)$$

のとき, 物体  $O^j$  を含むフレームを抽出するには  $D_{u_j}, D_{u_j+1}, \dots, D_{v_j}$  から少なくとも1フレームを選択する必要がある. GFD  $D_x$  を論理変数と見なし,  $D_x$  が選択されたとき  $D_x = 1$ , 選択されないとき  $D_x = 0$  とすると,  $O^j$  を含む GFD を抽出する条件は次式の

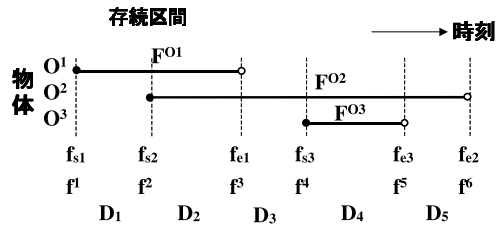


図1 オブジェクト・フレーム区間とグローバル・フレーム区間  
Fig.1 Object frame division and global frame division.

論理式を真とするように  $D_x$  の選択, 非選択を決めることに相当する.

$$Extract(O^j) = D_{u_j} + D_{u_j+1} + \dots + D_{v_j} \\ = \sum_{i=u_j}^{v_j} D_i \quad (5)$$

ここで,  $+$  は論理和 (OR) を表している. この式を以下フレーム区間選択式と呼ぶ.

$k$  個の物体があるときすべての物体を少なくとも1つのフレームに含む GFD の集合は次式のフレーム区間選択式を真にする.

$$Extract(O^1 O^2 \dots O^k) \\ = (D_{u_1} + D_{u_1+1} + \dots + D_{v_1}) \\ (D_{u_2} + D_{u_2+1} + \dots + D_{v_2}) \\ \dots \\ (D_{u_k} + D_{u_k+1} + \dots + D_{v_k}) \\ = \prod_{j=1}^k \sum_{i=u_j}^{v_j} D_i \quad (6)$$

ここで積は論理積 (AND) を表している. 式 (6) を積項の和の形に展開し, リテラル数 (積項の中の GFD  $D_x$  の数) が最小の項を求める. この項に含まれる各 GFD  $D_x$  からそれぞれ1フレームを抽出すればすべての移動物体をいずれかのフレームに含む最小数のフレームの集合が得られる. なお, 各 GFD から1フレームを選択する方法は 5.3 節で述べる.

### 2.3 例

図2は5つの物体  $O^1, \dots, O^5$  の存続区間を示している. 各物体の存続区間の始点, 終点で映像を区切ると図2のように  $D_1, \dots, D_8$  の8区間に分けられる. 物体  $O^1$  を抽出するには区間  $D_1$  または  $D_2$  から少なくとも1フレームを選択する必要がありフレーム区間選択式は

$$Extract(O^1) = D_1 + D_2 \quad (7)$$

となる. 全物体を抽出するフレーム区間選択式は式 (8) となる.

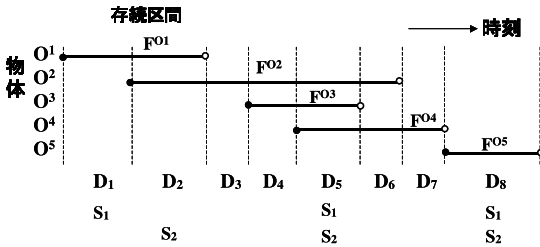


図2 最小数のキーフレーム選択の例

Fig. 2 Example of the minimum number of keyframe selection.

$$\begin{aligned}
 & Extract(O^1 O^2 O^3 O^4 O^5) \\
 &= (D_1 + D_2)(D_2 + D_3 + D_4 + D_5 + D_6) \\
 & \quad (D_4 + D_5)(D_5 + D_6 + D_7) D_8 \quad (8)
 \end{aligned}$$

式(8)を積和形に展開すると、

$$\begin{aligned}
 & Extract(O^1 O^2 O^3 O^4 O^5) \\
 &= D_1 D_5 D_8 + D_2 D_5 D_8 + \dots \quad (9)
 \end{aligned}$$

となる。リテラル数の最小値は3であり、3フレームを抽出すれば5個の物体  $O^1, \dots, O^5$  がいずれかのフレームに含まれる。リテラル数3の項は2通りあり、2通りのGFDの選択肢がある。式(9)の1, 2番目の積項に対応する選択区間を図2に  $S_1, S_2$  で示す。

### 3. 1物体につき複数のキーフレームを選択する場合

物体の動きをたどるには1つの物体につき1フレームを抽出しただけでは不十分であり、物体の侵入時、中央、消滅時など複数のフレームを選択する必要がある。これには2つの方法が考えられる。

#### 3.1 複数回重複して選択

1つのGFDから重複してフレームを抽出して各物体を指定回数以上含む最小のフレームの集合を求める。GFD  $D_i$  を必要な数だけ複製し、 $p$ 番目のものを  $D_i^p$  で表す。移動物体  $O^i$  を  $n_i$  回選ぶとするとフレーム区間選択式は次式で与えられる。

$$\begin{aligned}
 & Extract(O^{1(n_1)} O^{2(n_2)} \dots O^{k(n_k)}) \\
 &= (D_{u1}^1 + \dots + D_{v1}^1) \dots (D_{u1}^{n_1} + \dots + D_{v1}^{n_1}) \\
 & \quad \dots \\
 & (D_{uk}^1 + \dots + D_{vk}^1) \dots (D_{uk}^{n_k} + \dots + D_{vk}^{n_k}) \quad (10)
 \end{aligned}$$

例として図3のように3つの物体が存在するとき  $O^1$  を2回、 $O^2$  を2回、 $O^3$  を1回以上含む最小数のフレームの選択方法を示す。フレーム区間選択式は式(11)となる。

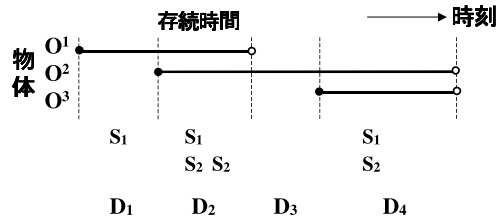


図3 各物体に対し複数のキーフレームを選択

Fig. 3 Selecting multiple keyframes for every object.

$$\begin{aligned}
 & Extract(O^{1(2)} O^{2(2)} O^{3(1)}) \\
 &= (D_1^1 + D_2^1)(D_2^1 + D_2^2) \\
 & \quad (D_2^1 + D_3^1 + D_4^1)(D_2^2 + D_3^2 + D_4^2) D_4^1 \\
 &= D_1^1 D_2^2 D_4^1 + D_2^1 D_2^2 D_4^1 \dots \quad (11)
 \end{aligned}$$

式(11)より最小3フレーム選択すればよい。最小値を与える選択肢は2通りあり、これを図3に  $S_1, S_2$  で示す。1つのGFDに複数個の  $S_i$  が書かれている区間ではその数だけ重複してフレームを抽出する。

この方法は各移動物体を指定された回数選択する最小のフレーム集合を与えるが、必ずしも適切な位置が選択できるとは限らない。図3の例では区間  $D_3$  は選択されていないが、この区間が長い場合は  $O^2$  に関しては適切なキーフレームとならない場合がある。

#### 3.2 移動物体の軌跡の特徴に基づく区間細分

もう1つの方法として物体の動きや持続時間からあらかじめ物体の持続区間を部分オブジェクト・フレーム区間(以下部分OFDと略記)に細分しておくことが考えられる。1つの物体の持続区間を

- 移動距離—たとえば幅高さの1/4移動するごとにOFD分割、
- 移動方向—たとえば向きが  $90^\circ$  変わるとにOFD分割、
- 持続期間—たとえば5秒ごとにOFD分割、

であらかじめ細分し、すべての部分OFDから1フレームを抽出する。図4に例を示す。細分化のための軌跡の特徴の番号を  $\alpha$ 、その特徴により細分化された部分OFDの番号を  $\beta$  とし、部分OFDを  $F^{Oj-\alpha\beta}$  で表している。図4の例においてすべての部分OFDから1フレーム以上を抽出するためのフレーム区間選択式は次式となる。

$$\begin{aligned}
 & Extract(O^1) \\
 &= (D_1 + D_2)(D_3 + D_4 + D_5)(D_6 + D_7 + D_8) \\
 & \quad (D_2 + D_3)(D_4 + D_5 + D_6) \\
 & \quad (D_1 + D_2 + D_3 + D_4)(D_5 + D_6 + D_7) \\
 &= D_2 D_3 D_6 + D_2 D_4 D_6 + D_2 D_5 D_6 + \dots \quad (12)
 \end{aligned}$$

複数の移動物体がある場合はそれぞれの物体の持続

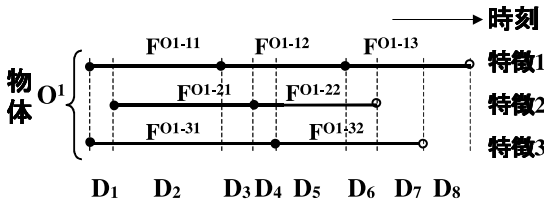


図4 軌跡の特徴によるオブジェクト・フレーム区間の細分  
 Fig. 4 Subdividing object frame division based on the trajectory features.

区間を移動軌跡に応じて独立に細分化してフレーム区間選択式を作り，全移動物体についての論理積をとればよい。

4. 種々のフレーム区間選択条件

4.1 選択の除外

オクルージョンや，光源の状態などで対象物体が不鮮明なフレームはあらかじめ除いて選択区間を定める．これまでの例では移動物体の存続区間は連続になっているが，一部の区間を選択から除外しても同様な手法が適用できる．図5では  $F^{O^1}$ ,  $F^{O^3}$  の一部を選択から除外している．フレーム区間選択式は式(13)となり，これを真とする GFD の組合せを求めれば除外区間を除いた解が得られる．

$$\begin{aligned}
 & Extract(O^1 O^2 O^3) \\
 &= (D_1 + D_2 + D_5 + D_6 + D_7 + D_8) \\
 & \quad (D_2 + D_3 + D_4 + D_5 + D_6 + D_7 + D_8 + D_9) \\
 & \quad (D_4 + D_5 + D_7) \\
 &= D_5 + D_7 + \dots \quad (13)
 \end{aligned}$$

ただしこれは， $O^1$  に関しては  $D_3, D_4$  以外の区間を選択すること，および， $O^3$  に関しては  $D_6$  以外の区間を選択することを条件とするものであり，他の物体の選択のために区間  $D_3, D_4, D_6$  が選択されることは禁止していない．たとえば区間  $D_3$  のみに現れる他の移動物体があれば，その物体のために  $D_3$  も選択される．

4.2 排他的選択

3.2 節で述べた移動軌跡の特徴に応じたキーフレームの選択では，同一条件に基づくキーフレームはできるだけ離れた時刻で抽出することが望ましい．キーフレームが接近してしまうことを避けるため排他的条件を考慮することができる．図6において  $F^{O^1-11}$  から  $D_2$  のみを選択し， $F^{O^1-12}$  から  $D_3$  のみを選択することを避けるには式(14)のフレーム区間選択式を用いる．ここで， $\overline{D_x}$  は  $D_x$  の否定，すなわち区間  $D_x$  を選択しないことを表す．

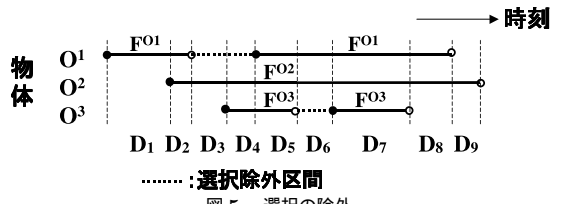


図5 選択の除外  
 Fig. 5 Exclude some global frame division.

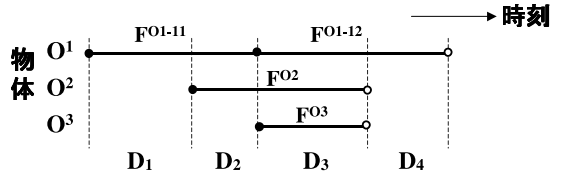


図6 排他的条件  
 Fig. 6 Exclusive selection of global frame division.

$$\begin{aligned}
 & Extract(O^1 O^2 O^3) \\
 &= (D_1 + D_2)(D_3 + D_4) \overline{(D_1 D_2 D_3 D_4)} \\
 & \quad (D_2 + D_3) D_3 \\
 &= D_1 D_3 + D_1 \overline{D_2} D_3 + D_1 D_2 D_3 + \dots \quad (14)
 \end{aligned}$$

最小リテラル数の項を求めるときは積和形に展開後，否定の項を無視して肯定の項の数が最少ない項を求める．GFD は排他的すなわち

$$D_i \cap D_j = \phi (i \neq j) \quad (\phi \text{ は空集合}) \quad (15)$$

であるから，たとえば  $D_1 \overline{D_2} D_3$  では  $\overline{D_2}$  を無視して  $D_1, D_3$  からフレームを選択すれば  $D_2$  とは重複しない．この例では最適選択区間は  $D_1 D_3$  となる．なお，排他的条件を  $\overline{D_2} D_3$  とすると  $D_2$  と  $D_3$  を同時に選択してはいけないことになり， $D_2$  の区間内だけに現れる他の物体と  $D_3$  の区間内だけに現れるもう1つの物体が同時に存在すると，全物体を含む GFD の集合が選択できなくなる．すなわち，フレーム区間選択式を真とする GFD の組合せがなくなる．

本節で示したように， $D_i$  間の論理式で表される条件はすべて考慮することができる．たとえば，移動物体  $O^1, O^2$  を同時に含み，かつ， $O^3$  を含まないフレームを求めるといった条件を任意に設定できる．

5. 計算手順とキーフレームの抽出

5.1 リテラル数最小の項の計算手順

まず排他的条件を含まない場合についてフレーム区間選択式を真とするリテラル数最小の項を求める手順を示す．この方法は論理回路の簡略化に用いられる Quine-McCluskey の方法<sup>10)</sup>の後半の最小項の選択アルゴリズム(最小被覆問題と呼ばれる)を応用したものである．

(1) 包含図の作成

OFD  $F^{Oj}$  (あるいは, 軌跡の特徴により細分した部分 OFD  $F^{Oj-\alpha\beta}$ . ここでは以下まとめて  $F^{Oj}$  と記する) を横軸に, GFD  $D_i$  を縦軸にとったマトリクスを作る. 各  $F^{Oj}$  の構成要素となる  $D_i$  の部分に印を付ける. 以下このマトリクスを包含図と呼ぶ.

(2) 包含図の簡略化

(a) 列の統合:  $F^{Oj1}$  の列で 印が付いているすべての  $D_i$  の集合を  $U^{j1} = \{D_i | D_i \in F^{Oj1}\}$  としたとき,  $U^{j1} \subset U^{j2}$  であれば  $F^{Oj2}$  の列を  $F^{Oj1}$  の列に統合する. すなわち,  $F^{Oj2}$  の列を消去する. また,  $F^{Oj1}$  の列のラベルに  $F^{Oj2}$  を書き加える.

(b) 空行の消去: 印を持たない行を消去する.

(c) 行の消去:  $D_{i1}$  の行で 印が付いているすべての  $F^{Oj}$  の集合を  $V^{i1} = \{F^{Oj} | D_{i1} \in F^{Oj}\}$  としたとき,  $V^{i1} \subset V^{i2}$  であれば  $D_{i1}$  の行を消去する.

(d) (a)~(c) の操作を, 適用できなくなるまで繰り返す.

(3) 必須項の抽出

簡略化された包含図において 印を 1 つしか持たない列があれば に対応する  $D_i$  は必ず選択しなければならない. これを必須項と呼ぶ.

(4) 等価項の抽出

同じ列に 印を 1 つのみ持つ行が複数ある場合は当該列の  $D_i$  の中の 1 つを選択する. これを等価項と呼ぶ. 等価項の中でどれを選ぶかは区間内でのキーフレームの選択と同時に決定する.

(5) 選択項の選択

以上の操作の後 印が残った場合は列ごとに 印に該当する  $D_i$  の論理和を作り, 全列の論理積を展開して最小リテラル数の項を求める. このような項を選択項と呼ぶ.

5.2 計算例

図 2 を例に計算例を示す. フレーム選択式は式 (16) で与えられる.

$$\begin{aligned} & Extract(O^1 O^2 O^3 O^4 O^5) \\ &= (D_1 + D_2)(D_2 + D_3 + D_4 + D_5 + D_6) \\ & \quad (D_4 + D_5)(D_5 + D_6 + D_7) D_8 \end{aligned} \quad (16)$$

まず表 1 の包含図を作る. 次に列の統合を行う.  $F^{O3}$  の列の 印はすべて  $F^{O2}$  の列の 印に含まれるため  $F^{O2}$  の列を  $F^{O3}$  の列に統合すると表 2 となる. こ

表 1 包含図

Table 1 Initial inclusion table.

GFD	移動物体のフレーム区間 OFD				
	$F^{O1}$	$F^{O2}$	$F^{O3}$	$F^{O4}$	$F^{O5}$
$D_1$	○				
$D_2$	○	○			
$D_3$		○			
$D_4$		○	○		
$D_5$		○	○	○	
$D_6$		○		○	○
$D_7$				○	○
$D_8$					○

表 2 列の統合後の包含図

Table 2 After the column unification.

GFD	移動物体のフレーム区間 OFD				
	$F^{O1}$	$F^{O2} F^{O3}$	$F^{O4}$	$F^{O5}$	
$D_1$	○				
$D_2$	○				
$D_3$					
$D_4$		○			
$D_5$		○	○		
$D_6$			○	○	
$D_7$			○	○	
$D_8$					○

表 3 行の消去後の包含図

Table 3 After the row elimination.

GFD	移動物体のフレーム区間 OFD				
	$F^{O1}$	$F^{O2} F^{O3}$	$F^{O4}$	$F^{O5}$	
$D_1$	○				
$D_2$	○				
$D_5$		○	○		
$D_8$					○

表 4 簡略化後の包含図

Table 4 After the simplification.

GFD	移動物体のフレーム区間 OFD				
	$F^{O1}$	$F^{O2} F^{O3}$	$F^{O4}$	$F^{O5}$	
$D_1$	○				
$D_2$	○				
$D_5$		○			
$D_8$					○

れは,  $O^2$  の存続区間は  $O^3$  の存続区間をすべて包含して  $F^{O3}$  の区間  $D_4, D_5$  からキーフレームを選べば  $O^2$  も必ず抽出されるためである. 表 2 において  $D_3$  の行は 印がないため消去する. 次に行の消去を行う. 表 2 において  $D_4$  の行の 印はすべて  $D_5$  の行に含まれるため  $D_4$  の行を消去する. これは,  $D_4$  を選ぶよりも, もっと多くの移動物体を含む  $D_5$  を選ぶ方がよいことを意味している. 同様に  $D_6, D_7$  の行も消去すると表 3 が得られる. 再度列の統合により  $F^{O2} F^{O3}$  の列と  $F^{O4}$  の列を統合し表 4 を得る. 簡略化はここで終了する. 表 4 より  $D_5, D_8$  が必須

このラベルは選択された GFD 内から 1 つのキーフレームを抽出するときに使用する. 5.3 節を参照.

5.3 節を参照.

表 5 選択項が残る包含図  
Table 5 Inclusion table with the selective terms.

GFD	移動物体のフレーム区間 OFD		
	$F^{O^1}$	$F^{O^2}$	$F^{O^3}$
$D_1$	○	○	
$D_2$	○		○
$D_3$		○	○

項,  $D_1, D_2$  が等価項となり, 最小リテラル数は 3, リテラル数最小の項は  $D_1D_5D_8$  および  $D_2D_5D_8$  となる.

5.1 節で述べた選択項が残る最も簡単な例は表 5 の形であり, この場合は  $(D_1 + D_2)(D_1 + D_3)(D_2 + D_3)$  を展開して最小リテラルの項  $D_1D_2, D_1D_3, D_2D_3$  を得る.

5.3 グローバル・フレーム区間からのキーフレーム選択

リテラル数最小の項の計算手順で示したように, 選択された GFD はどの移動物体のために選択されたかが分かっている. これを利用して対象とする移動物体ができるだけ大きく写っているフレームを選ぶ. 5.2 節の例を用いて計算方法を説明する.

5.3.1 選択項がない場合

表 4 の例で説明する. 表より区間  $D_1, D_2$  は  $O^1$  を抽出するためであることが分かる. フレーム  $f_p$  に  $O^i$  が写っている面積を  $S(O^i|f_p)$  としたとき, 区間  $\{D_1, D_2\}$  において  $S(O^1|f_p)$  を最大にするフレームを選択する. 同様にして区間  $D_8$  からは  $S(O^5|f_p)$  を最大にするフレーム  $f_p$  を選ぶ. 区間  $D_5$  からは  $O^2, O^3, O^4$  を抽出する必要がある. このように複数の移動物体が対象となる場合は何らかの面積評価法が必要となる. 面積の和を用いると大きな物体の面積が支配的となり小さな物体については必ずしも好ましくないフレームを選択する場合もある. したがって本手法では対象物体の面積の積が最大となる場所を選ぶものとした. この例では  $S(O^2|f_p) \times S(O^3|f_p) \times S(O^4|f_p)$  を最大とするフレーム  $f_p$  を選ぶ.

5.3.2 選択項がある場合

選択項がある場合は対象物体の面積の積が最大となる組合せとキーフレームの位置を同時に求める. 表 5 の例で説明する. 最小リテラルの項は  $D_1D_2, D_1D_3, D_2D_3$  の 3 通りある.  $D_1D_2$  を選んだ場合は  $D_1$  から  $O^1, O^2$  を, 区間  $D_2$  から  $O^1, O^3$  を抽出することになり,  $O^1$  は両方に含まれる.  $D_1$  から  $\sqrt{S(O^1|f_p)} \times S(O^2|f_p)$  を最大とするフレーム  $f_p$  を,  $D_2$  から  $\sqrt{S(O^1|f_q)} \times S(O^3|f_q)$  を最大とするフレーム  $f_q$  を選び, 最小リテラル項  $D_1D_2$  の評価値を

$$S_{12} = \sqrt{S(O^1|f_p)} \times S(O^2|f_p) \times \sqrt{S(O^1|f_q)} \times S(O^3|f_q) \tag{17}$$

とする.  $O^1$  は 2 回現れるため平方根をとっている. 一般に,  $n$  回現れる場合は  $n$  乗根をとる. これらすべての選択肢について計算し最大値を求めれば GFD の最適選択とその中での最適キーフレーム位置を同時に定めることができる.

5.4 排他的条件を含む場合の計算手法

排他的条件を含む場合はまず否定項を含まない部分のみで最小リテラル数の項を求め, これと否定項との論理積を計算し, 得られた項の中から 4.1 節で述べたように, 否定項を除いてリテラル数最小の項を求める. しかしながらこの方法では, 実際には解が存在しても解が得られない場合が生じる. たとえば図 2 で  $D_5$  を除外してキーフレームを選択したい場合フレーム区間選択式 (18) を上記の方法で展開すると解が得られない.

$$\begin{aligned} & Extract(O^1O^2O^3O^4O^5) \\ & = (D_1 + D_2)(D_2 + D_3 + D_4 + D_5 + D_6) \\ & \quad (D_4 + D_5)(D_5 + D_6 + D_7)D_8\overline{D_5} \end{aligned} \tag{18}$$

これは肯定項の最小リテラル項はすべて  $D_5$  を含み, かつ, 単独の否定項  $\overline{D_5}$  があるためである. この場合は, あらかじめ包含図から  $D_5$  を除いておけば最小リテラル数の項が得られる. このように, 否定項に応じてあらかじめ前処理を行えば計算可能であるが, 場合分けが複雑になること, このようなことは稀であることから, 解が得られなかった場合は通常の変項展開に切り替える方が容易である.

5.5 全体の処理フロー

以上をまとめると全体の処理フローは以下の手順となる.

- (1) 移動物体の存続区間を求める.
- (2) 移動物体ごとに以下の処理を行う.
  - 軌跡特徴により移動物体の存続区間を分割
  - キーフレームとして適切でない区間を除外
- (3) (2) で定めた OFD の境界をソートして GFD  $D_i$  を求める.
- (4) フレーム区間選択式を求める.
- (5) 排他条件を除いてリテラル数最小の項を求める.
- (6) 排他条件があれば論理積をとりリテラル数最小の項を求める. 解が得られない場合は多項式展開で再計算する.
- (7) 移動物体の面積が最大になるようにキーフレームを選択する.

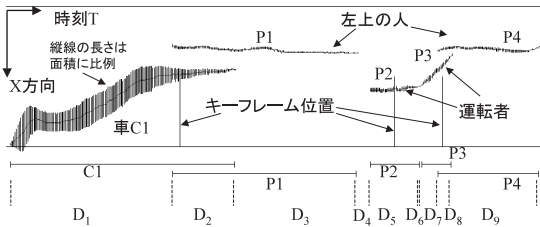


図7 1フレーム/移動物体のときの移動物体の持続期間の時空間表示

Fig. 7 Spatiotemporal representation of the moving objects in case of 1 frame/object.

### 6. 実行例

駐車場を約2時間モニタした映像から提案手法によりキーフレームを抽出した例を示す。撮影日は風が強くとどき駐車場周辺の木が大きく揺れ誤識別が多量に発生した。そこでまず目視により人と車のみを取り出したデータからキーフレームを抽出した結果を示し、次に木の揺れを含む全データでの実験結果を示す。なお、移動物体の検出には文献 11), 12) の、トラッキングには文献 13) の方法を用いた。フレームレートは 5 frame/sec, 画像サイズは 320 × 240 ピクセルとした。

#### 6.1 人と車のみ抽出, 1フレーム/移動物体

人と車のみ取り出し, 1 移動物体につき 1 フレームを抽出したときの, 移動物体の時空間軌跡の XT 断面の一部を図 7 に示す。このシーンで検出された移動物体は車 C1, 人 P1, ..., P4 の 5 個である。なお, P1 と P4, P2 と P3 は同一人物であるが一時静止や急な速度変化により別の物体として検出されている。図中で各移動物体を表す曲線の上に描かれた縦線はその時点での物体の面積を表している。また, 図の下の横線は OFD を表し, 最下部の  $D_i$  が GFD を表す。図中で約半分の高さの縦線がキーフレーム抽出位置を表している。まず最小数の GFD として  $D_2, D_5, D_8$  の 3 区間が選択され, 各 GFD から面積を考慮してそれぞれ 1 フレームが抽出されている。

#### 6.2 人と車のみ抽出, 移動軌跡を考慮

人と車のみ取り出し, 移動軌跡に基づいて複数のキーフレームを抽出した例を図 8 に示す。OFD は以下のように定めた。

- 移動距離—幅高さの 1/4 移動するごとに OFD 分割。
- 移動方向—向きが 90° 変わるごとに OFD 分割。
- 持続期間—5 秒以下は 1 区間, 5 秒を超えて 20 秒以下は 2 区間, 20 秒を超える場合は  $\lceil t/12 \rceil + 1$

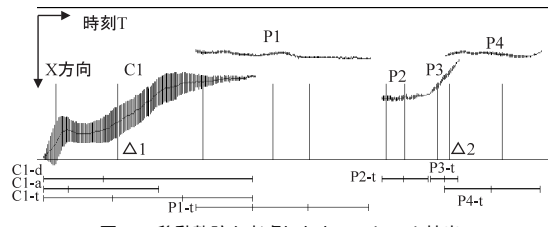


図8 移動軌跡を考慮したキーフレーム抽出  
Fig. 8 Keyframe extraction considering the trajectory of moving objects.



図9 図8の 1の部分のキーフレーム(黒線は移動軌跡)  
Fig. 9 Keyframe at 1 in Fig. 8.



図10 図8の 2の部分のフレーム  
Fig. 10 Keyframe at 2 in Fig. 8.

区間に OFD 分割。ここで,  $t$  は持続時間,  $[x]$  は  $x$  を超えない最大の整数を表す。

図 8 では移動距離, 移動方向, 持続期間に基づく OFD を移動物体の記号 C1, P1, ..., P5 の後にそれぞれ -d, -a, -t を付けて示している。図に示すように最小数のキーフレーム 10 個が選択されている。図 8 中の 1, 2 の時刻におけるキーフレームを図 9, 図 10 に示す。移動物体の面積を考慮しないと 1 では図 11 となり, 車の一部分が他の車の陰になってし



図 11 図 8 で移動物体の面積を考慮しない場合

Fig. 11 Keyframe at 1 in Fig. 8 without considering the object area.

表 6 駐車場シーンのキーフレーム抽出

Table 6 Keyframe extraction from the parking scene.

項目	人と車	木の揺れ	合計
検出された移動物体	20	57	77
部分 OFD の数	63	93	156
キーフレーム数	41	41	82
キーフレーム数/移動物体	2.05	0.72	1.06

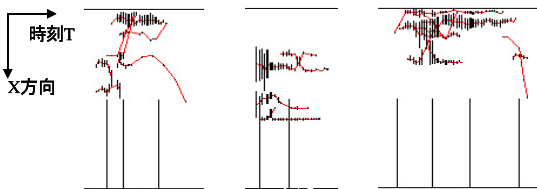


図 12 木の揺れ部分のキーフレーム

Fig. 12 Keyframes for swaying trees.

まう、面積考慮によりオクルージョンのない位置が選ばれていることが分かる。

6.3 木の揺れを含む全データでの比較

人と車のみ取り出した場合と、木の揺れを含む全データでのキーフレーム抽出の比較を表 6 に示す。また、木の揺れを移動物体として検出した部分のキーフレーム抽出の一部を図 12 に示す。図に示した区間では 21 個のオブジェクトを検出しているが、木の揺れはほぼ同時刻に発生すること、持続時間が短いことから、抽出したキーフレームは 9 個にとどまっている。屋外シーンのモニタでは草木や影の揺れの誤認識を避けることは困難であるが、本手法によりキーフレームの抽出率は大きく軽減される。表 6 に示したように、移動物体 1 つあたりの平均キーフレーム数は人と車で 2.05 フレーム、木の揺れでは 0.72 フレームであっ

た。全データでのキーフレーム抽出に要した計算時間は Pentium4 2.6G において約 1.4 秒であった。

7. む す び

移動物体の存続期間を基に、複数の移動物体にわたって位置、向き、持続時間などの条件を考慮し、かつ、最小数のキーフレームを選択する方法を提案した。

本手法は蓄積映像を対象としたものであるが、移動物体が途切れたときにキーフレーム抽出を行えばリアルタイム映像にも適用できる。移動物体が長時間続く場合はあらかじめ定めた最大経過時間を超えたらキーフレーム抽出に入るようにする。切断時刻をまたがって存続する移動物体は切断時刻の前後で別の物体として扱われるためキーフレームの数は若干増加するが、見落としは生じない。

謝辞 本研究を進めるにあたりご支援、ご討論いただいた NTT マイクロシステムインテグレーション研究所のホームコミュニケーション研究グループの皆様 に感謝いたします。

参 考 文 献

- 1) Gunesel, B. and Tekalp, A.M.: Content-Based Video Abstraction, *Proc. Intl. Conf. on Image Processing*, pp.128-132 (1998).
- 2) Dufaux, F.: Key frame selection to represent a video, *Proc. Intl. Conf. on Image Processing*, pp.275-278 (2000).
- 3) Chiu, P., Girgensohn, A., Polak, W., Rieffel, E. and Wilcox, L.: A Genetic Algorithm for Video Segmentation and Summarization, *Proc. Intl. Conf. on Multimedia and Expo*, pp.1329-1332 (2000).
- 4) Zhao, L., Qi, W., Li, S.Z., Yang, S. and Zhang, H.J: Key-frame Extraction and Shot Retrieval Using Nearest Feature Line (NFL), *Proc. ACM workshops on Multimedia*, pp.217-220 (2000).
- 5) Lipton, A.J, Fujiyoshi, H. and Patil, R.S.: Moving target classification and tracking from real-time video, *Proc. 4th IEEE Workshop on Applications of Computer Vision (WACV '98)*, pp.8-14 (1998).
- 6) Stringa, E. and Regazzoni, C.S.: Real-Time Video-Shot Detection for Scene Surveillance Applications, *IEEE Trans. Image Processing*, Vol.9, No.1, pp.69-79 (2000).
- 7) Wang, R.R., Hong, P. and Huang, T.: Memory-based moving object extraction for video indexing, *Proc. 15th Intl. Conf. on Pattern Recognition*, Vol.1, pp.811-814 (2000).
- 8) 藤吉弘巨, 榎本暢芳, 長谷川修, 金出武雄: ア



クティビティモニタリング—屋外監視映像の要約と WWW 上表示・検索システム, 第 7 回画像センシングシンポジウム講演論文集, H-4 (2001).

- 9) 藤吉弘巨, 金出武雄: VSAM: 画像理解技術を用いたビデオ監視システムプロジェクトについて, 情報処理学会研究報告, グラフィクスと CAD, Vol.2001, No.106 (2001).
- 10) 江端克彦, 久津輪敏郎: デジタル回路設計, 共立出版, 第 2 章 (1997).
- 11) Stauffer, C. and Grimson, W.E.L.: Adaptive Background Mixture Model for Real-Time Tracking, *Proc. Computer Vision and Pattern Recognition* (1998).
- 12) Stauffer, C. and Grimson, W.E.L.: Learning Patterns of Activity Using Real-Time Tracking, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.22, No.8, pp.747-757 (2000).
- 13) Veenman, C.J., Reinders, M.J.T. and Backer, E.: Resolving Motion Correspondence for Densely Moving Points, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.23, No.1, pp.54-72 (2001).

(平成 15 年 5 月 29 日受付)

(平成 16 年 1 月 6 日採録)



北澤 仁志 (正会員)

1952 年生. 1979 年東京工業大学大学院理工学研究科電子工学専攻博士課程修了. 同年日本電信電話公社入社. 2002 年東京農工大学工学部電気電子工学科教授. LSI レイアウト CAD, 低電力化 CAD, ネットワークベース CG システム等の研究開発を経て現在は動画像の要約や検索に関する研究に従事. 工学博士. IEEE-CS および CAS, 電子情報通信学会各会員.



小野澤 晃 (正会員)

1960 年生. 1985 年早稲田大学大学院理工学研究科修士課程電子通信学専門分野修了. 同年日本電信電話株式会社 (NTT) 入社. 現在, NTT マイクロシステムインテグレーション研究所主幹研究員. LSI レイアウト CAD, CG, ヒューマンインタフェースに関する研究開発に従事. 博士 (情報科学). ACM, IEEE-CS および CAS, 電子情報通信学会各会員.