

マルチメディア・データベース システム

—映像・画像の被写体、被写区域検索—

金 世 良 打 浪 清 一 手 塚 慶 一
大 阪 大 学 工 学 部

科学調査を行う時、または重大な事件や、事故が発生した時、イベント情報としてたくさんの写真、ビデオ録画などの画像・映像情報を収集するのが普通である。後の研究分析段階では、指定する被写体あるいは観察したい任意の場所に事件が発生している当時の過程を詳細に調べるため、数多くの画像カットから迅速に、漏れなく関連カットを即時検索し呈示するのは、非常に重要である。本稿は、効率的に検索できる抄録を作るため、全区域を平行直線で分割し、可視区間による直線の切片群を抄録とする方法を提案し、障害物の存在のため、隠れて見えない区域の判断法、検索画像の図的確認法及び検索システムのアルゴリズムとシステム構成について述べる。

M u l t i - m e d i a D a t a b a s e S y s t e m

— Search method for subjects and areas of video images —

Shiliang JIN Seiichi UCHINANI Yoshikazu TEZUKA

Faculty of Engineering, Osaka University, 2-1 Yamadaoka, Suita-shi, 565, Japan

In this paper we propose a new indexing method for image retrieval system by subjects or appointed areas.

When we investigate events such as accidents or festivals, we collect many images in photos and video record images. In order to investigate some wanted subjects or some detailed procedure at an any given place, it is important that related cuts be immediately retrieved from many cuts without any lack.

To make abstracts and index that can be checked out efficiently, we propose a new indexing method, in which the map is divided by parallel lines, and each cut is indexed by visible line intervals. We show the organization of the system and explain its algorithms.

1 、 ま え が き

現代社会では、重大な事件や行事を記録する時、また各分野で科学調査を行う時、イベント情報として対象に関する沢山の写真、ビデオ画像とか、録音などでその詳細情報を収集するのが普通であり、また重大な事故の原因解明を行う際も、各所で記録された各種メディア・データを集め、これを解析するのが普通である。記録された映像・画像の量は少なくとも何百カット、多ければ何千、何万カットの場合もあるから、その数はおびただしい量にのぼる。その後の研究分析段階では、事件のうちのある場所の過程を詳細に調べるにあたり、数多くのカットの中にその場所を記録したカットはあるかどうか、あるとすれば、一体いくつあるか、それぞれ何番と何番か、どのカットが撮影距離と角度の点で分析に一番便利かという事柄は、研究者が非常に関心を持っていることである。即ち、大量の画像情報の中から指定される被写体あるいは観察したい任意の場所に関連する画像情報を即時に検索し表示することは非常に必要である。しかし、記録された画像情報が多ければ多いほど、検索するためにかかる時間が急激に増加すると同時に、大事な情報を見落す現象が発生しやすくなることも事実である。その検索精度は、分析研究の進展と正確なデータの提供に深刻な影響がある。これを解決する試みの一つとして、本稿は、マルチメディア・データベースシステムを構成するに先立ち、画像情報抄録データベースを作成し、これを用いて迅速に漏れなく被写体及び被写地域検索が行え、更に図的確認が出来るシステムを構成した。

なお、インドネシアバリ島で調査したワトカル寺のオダラン(創立記念祭)のビデオ映像を対象としたシステムを作成し、現在良好に稼動している。

2 、 情報抄録のための記録モデルの選択と撮影方向の分類

ある指定した被写体を捜せるようにするためには、あるいは研究分析する時、任意に指定した地域に発生した事件の当時の実状記録を見たい場合には、地図上でどこにどのような被写体があるのかを記録しておく必要がある、またどのメディアのどの位置にそれが写っているかを抄録しておく必要がある。このための新しい方法を提案し、そのシステムを試作した。

1) 情報記録モデル

現在の Case study では、バリヒンズー教寺院の境内での被写体検索が主になる。一つ一つの映像資料の写している区域をデータベースに抄録するため、区域全体を分割する必要がある。分割してから全区域を小さな記録単位の集合に置き換えて、この小さい単位を情報抄録単位として、映像情報をデータベースに抄録する。よって、モデルの選択と全区域(地図)の合理的な分割法は、抄録システムの重要な基礎となる。

モデルの種類は、いろいろと考えられる(例えば、小さな四角形のマトリクス、点列モデルなどがある)が、どのような方法を選ぶかは、以下三つの原則を主旨として考えなければならない。

モデルの選択主旨:

- a データ量は、できるだけ少なくなるように、

- b 容易に抄録データを構成し、
- c 容易に検索できるように。

この原則を基にして、本方法は、それぞれX軸とY軸に平行する二組の直線族により、地図を分割し、二つの直線族の可視区域による切片(線分)を情報記録の単位とするモデルを提案する。Meshの細かさ(即ち直線の間隔)は、任意であり、分解能に比例するが、情報を記録するため、原則として隣接する建物の間に少なくとも一本の直線が引かれるように選ばれるべきである(図1参照。両直線族の直線数をMxとMyとする)。Meshの細かさは、小さいほど精度が向上する。

本方法は、計算に便利のため等間隔での分割案を採用している。

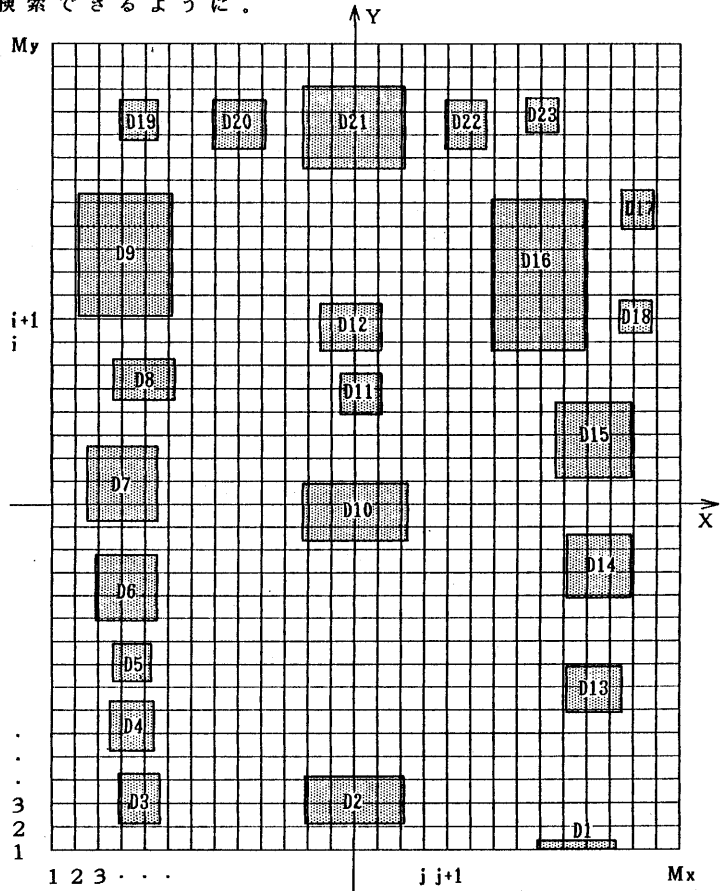


図1 抄録のための地図の分割
(バリヒンズー教ワトカル寺の例)

2) 撮影方向の分類

二組の直線族に対して、画像情報抄録データベースは、二つのファイルを用意している。現場撮影時にカメラの位置と撮影方向は、ランダムであるから、どの位置でどの方向への撮影も可能である。各カットの映像は、そのカットの撮影方向によりX方向撮影とY方向撮影の二種類に分けられる(図2参照)。ここで、撮影方向というのは、撮影角の二等分線とX軸となる角度 α により取り決められる。即ち、撮影角の二等分線がX軸に傾くとX方向撮影、Y軸に傾くとY方向撮影と定義される。

X方向撮影:

$$-45^\circ \leq \alpha \leq 45^\circ \quad \text{又は} \quad 135^\circ \leq \alpha \leq 225^\circ$$

Y方向撮影:

$$45^\circ < \alpha < 135^\circ \quad \text{又は} \quad 225^\circ < \alpha < 315^\circ$$

X方向撮影の可視区域情報がFileXへ、Y方向撮影の可視区域情報がFileYへ蓄えられている。ファイル内のレコードは直線番号順に並べ格納される。よって、全部の映像情報が二つのFileの何れかに格納されている。

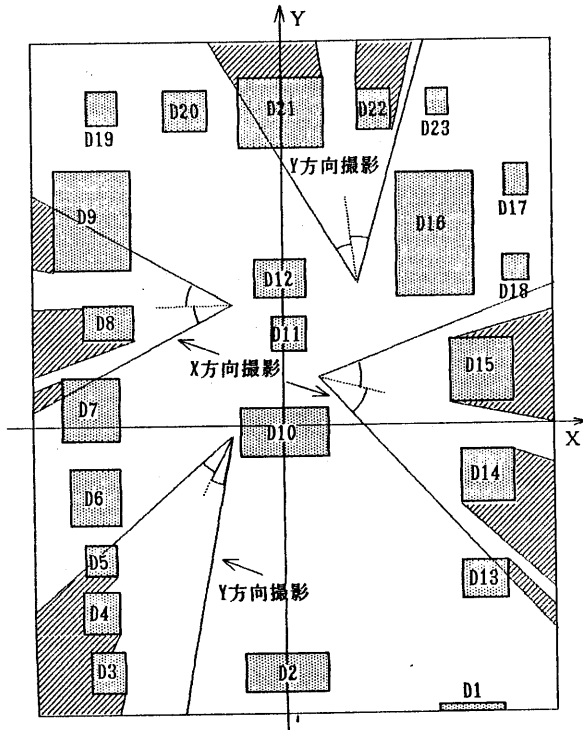


図2 撮影方向の確定

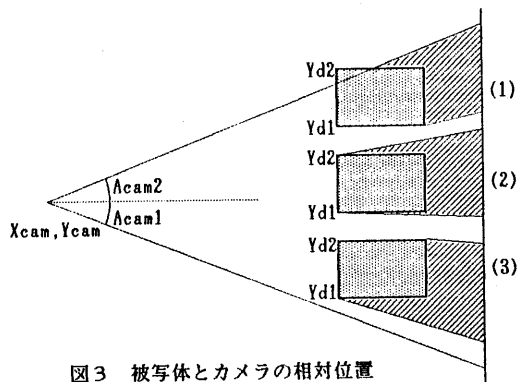


図3 被写体とカメラの相対位置

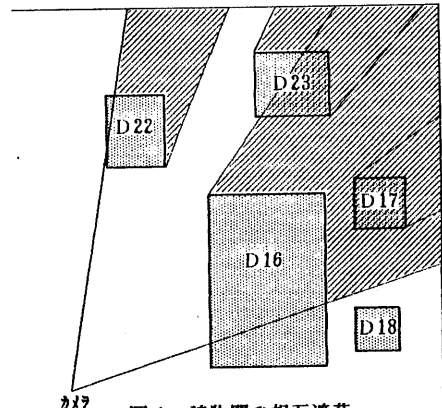


図4 建物間の相互遮蔽

3. 抄録・検索アルゴリズム

1) 抄録アルゴリズム

カット写真と撮影地の平面図をもとに、カメラ位置と写角を推定することが可能である。この抄録は、撮影者が行う。よって、撮影された各カットにおいては、カメラの位置 X_{cam}, Y_{cam} と撮影角度 A_{cam1}, A_{cam2} が既知とする。将来は、この抄録も自動化を検討したい。この四つの条件があれば、撮影の最大範囲は決められている。だが、撮影角内には、たくさんの建物があり、これらの建物は、背景画像に対して障害物となる。建物が存在するため、撮影区域は、見える部分と隠れて見えない部分に分けられる。被写区域の抄録は、撮影方向と撮影角度の相違、カメラと建物との相対位置の変化、建物間の互いに遮蔽し合う等の影響で様々な組み合わせの可能性が出てくるわけで、その処理も大変面倒になる。

本方法のアルゴリズムは主に三部分から構成されている

a) 単一建物の個別処理

上に述べたように、撮影範囲全般的にみれば建物群の処理が複雑であるが、一つの建物に対しては、カメラとの相対位置は、かならず図3に示した三種類の内の一つに属する。

$$(1) \quad Y_{cam} \leq Y_{d1}$$

$$(2) \quad Y_{d1} < Y_{cam} < Y_{d2}$$

$$(3) \quad Y_{cam} > Y_{d2}$$

または：

$$X_{cam} \leq X_{d1}$$

$$X_{d1} < X_{cam} < X_{d2}$$

$$X_{cam} > X_{d2}$$

このルールに基づき、まずは別の建物の影響を無視して建物毎に単一処理をし、撮影区域内建物毎の不可視区域を個別に求めておく。

b) 遮蔽処理

個別処理の結果を一次データとして、建物間の相互遮蔽処理に進む。即ち、不可視区域の重畳処理を行う。(図4参照)

```

dimension yd(u,v,w)          yd: 個々の建物の遮蔽区域の情報; u: 建物の番号; v: 直線の番号;
                               w: 大きさ: 3 (w=0: 考慮の必要の有無を表すフラグ, w=1: 不可
                               視区域(線分)の下限, w=2: 不可視区域(線分)の上限)
if j0=0 goto *z4             (j0: 撮影範囲内の建物数)
for i=icam to il step i0     (icam: 撮影方向に向けカメラが一番近い直線の番号, il: 境界線番号,
                               i0=1正方向撮影、-1負方向撮影)
    p=1 : j00=j0
*z0 : q=0
    if p>=j0 goto *z3
    if yd(p,i,0)<=0 then j00=j00-1 : goto *z2      (yd(u,v,0)<0: 建物が無視される)
    for j=p+1 to j0
        if yd(p,i,1)>yd(j,i,2) or yd(p,i,2)<yd(j,i,1) then q=1 : goto *z1
        if yd(p,i,2)>=yd(j,i,2) and yd(p,i,1)<=yd(j,i,1) then yd(j,i,0)=-10 : goto *z1
        if yd(p,i,2)<yd(j,i,2) then yd(p,i,2)=yd(j,i,2) : yd(j,i,0)=-1
        if yd(p,i,1)>yd(j,i,1) then yd(p,i,1)=yd(j,i,1) : yd(j,i,0)=-1
        if q=1 goto *z0
    *z1 : next j
    *z2 : p=p+1 : goto *z0
    *z3 : if p=j0 then .....
           .....
    next i
*z4 : .....

```

図5 可視区域の重畳処理のアルゴリズム

図5のアルゴリズムに示したように、建物の番号順により前の建物を基準として、もしあとの建物がそれと遮蔽関係になれば(相互遮蔽関係でない時: 例えば、D16とD22の場合、処理しないで、各自独立区域とする)、相互遮蔽の具合により、あとの建物が無視される(例えばD16とD17の相互関係)か、前の建物の遮蔽区域の境界を変更する(例えば、D16とD23の相互関係)かの何れかの処理を扱う。

c) 可視区域の計算

重畳処理とは、実際に撮影範囲内各建物の各直線位置でできた不可視線分をつなぎ合わせる処理であり、即ち、重畳処理をすると、一つのカットは、各直線上にいくつかの相互独立の(互いに全く重畳し合わない)不可視区域を持っている。不可視区域の間は、可視区域になる。よって、可視区域も求めることができる。

2) 検索アルゴリズム

地図上の任意位置Cの座標(X_c , Y_c)が指定されるとC点を挟むX軸に平行する二直線 X_i , X_{i+1} とY軸に平行する二直線 Y_j , Y_{j+1} が決められる(C点が丁度直線上に落ちる場合、一本の線になることは、当然である)。

抄録処理のとき、各可視区域の端点情報がデータベースに格納されているが、

記録される一組の情報は、次の通りである。

各直線位置での可視線分両端の座標値： Y_1, Y_2 (あるいは X_1, X_2)

角度： $\alpha_1, \alpha_2, \alpha_i, \alpha_{ii}$ (直線位置での可視線分の両端における+方向(1,2)

及び-方向(i,ii)の可視境界線の角度)

このカットの番号： K

a) 一般検索

まず File X の Y_j と Y_{j+1} に対応するレコードを走査して、下式を満たすかどうかを検査する。

$$Y_{1,j} + (X_c - X_j) * \tan(\alpha_{1,j}) < Y_c < Y_{2,j} + (X_c - X_j) * \tan(\alpha_{2,j}) \quad (1)$$

$$Y_{1,j+1} + (X_c - X_{j+1}) * \tan(\alpha_{i,j+1}) < Y_c < Y_{2,j+1} + (X_c - X_{j+1}) * \tan(\alpha_{ii,j+1}) \quad (2)$$

図6に示すように、もし、(1)、(2)式の両式を満足すれば、そのカットの撮影はかならず位置Cを写していることが判断できる(図6のC1点を参照)。両式とも満足しなければ、そのカットの撮影は、位置Cを写していないことが分かる(図6のC0点を参照)。しかし、

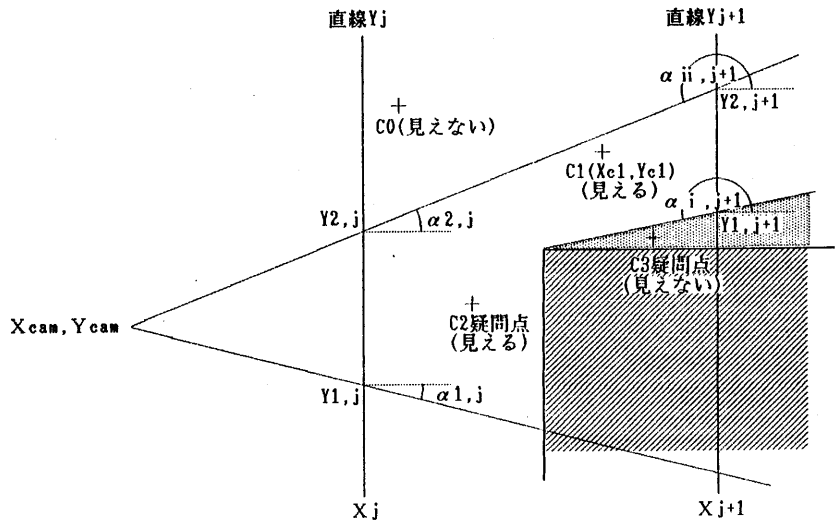


図6 被写位置可視性の判断

両式の片方を満足している場合、その点は、見える(図6のC2)、見えない(図6のC3)の両方の可能性があり、その判断が簡単に(1)、(2)式で行えないから疑問点として後程判断ルーチンにより確認しなければならない。同様な方法で

$$X_{1,i} + (Y_c - Y_i) * \tan(\alpha_{1,i}) < X_c < X_{2,i} + (Y_c - Y_i) * \tan(\alpha_{2,i}) \quad (3)$$

$$X_{1,i+1} + (Y_c - Y_{i+1}) * \tan(\alpha_{i,i+1}) < X_c < X_{2,i+1} + (Y_c - Y_{i+1}) * \tan(\alpha_{ii,i+1}) \quad (4)$$

により、File Y の X_i と X_{i+1} に対応するレコードを走査して、上式を満たすカットを全部検出する。疑問点があれば、判断ルーチンで確認する。

b) 疑問点の確認

単一の建物の計算処理をする同時に、このカットの撮影区域内各建物の遮蔽区域、即ち、単一の建物の不可視区域の関連パラメータ(見えない境界線データと遮蔽区域角度など)もカメラ情報データベースに抄録されている。疑問点にあたる場合、カメラ情報データベースをアクセスして、その疑問点は、建物の遮蔽区域にあるか、遮蔽区域外にあるかの最終判断が容易にできるようにしてある。

試作システムでは、疑問点は、非常に少なかった。もし区域の分割と Mesh の細かさが適当なら、疑問点は、ほとんどないと考えてよい。

4 システムの構成

全システムは、画像被写体抄録システムと被写区域検索システムの二つのサブシステムから成り立つ。画像抄録サブシステムは、本システムの中核となる可視区域抄録データベースを構成するためのシステムで、検索システムは、抄録システムにより作成されたデータベースを利用し、捜したい被写体あるいは被写体区域を検索し表示するためのユーザ向きのアプリケーションシステムである。

1) 画像被写体抄録システム

写真、ビデオをもとに、各カット毎のカメラの位置、カメラの写角を調べ、地図上にプロットした。

境内の平面図は、測量により描いたものを用いた。境内の動かない被写体(メルと呼ばれる社)は、平面図上に描かれているが、これが障害物となる。

通常、地図の形状は、割合に複雑であるから、処理の便宜上、いくつかの部分区域に分けて、分割処理するほうが簡単でかつ合理的である。ヒンズー教寺院の場合、四つの主要な境内を、各自A, B, C, Dという区域に分割して処理している。

一つのカットの入力データは、次の五つである：

カットの番号と特徴パラメータ： K ($K=1, 2, \dots, N$, N :全部のカット数)；

カメラの位置： X_{cam}, Y_{cam} ；撮影角度： A_{cam1}, A_{cam2} (写角両端の角)

システムは、抄録情報データベースを作成するまえ、まずカメラ情報ファイルを作って、全部のカットの一次データをカメラ情報ファイルに貯えておく。

システムの手順は、図7に示すように、繰り返し処理であり、カメラ情報ファイルからカット順により各カットのカメラ位置と撮影角度を取り出して、システムに入る。

カット毎に、まず所在の部分

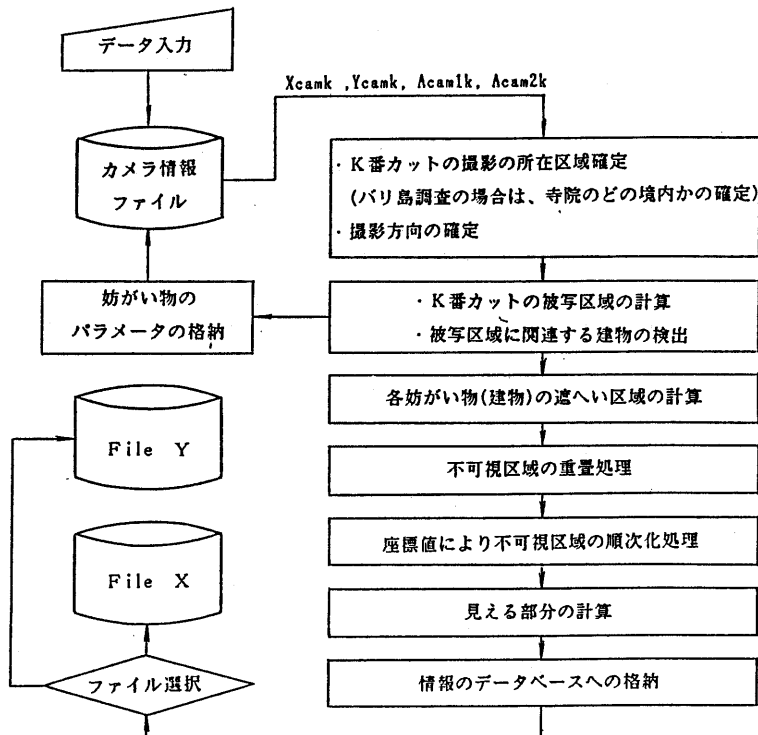


図7 抄録システムのフローチャート

区域(A, B, C, Dの中の一)と撮影方向(X方向あるいはY方向)を確定し、撮影区域内の関連建物を検出する。それから、撮影範囲計算、単一建物の遮蔽計算、建物間の相互遮蔽処理、すなわち不可視区域の重畳処理、最後は、可視区域を求め算出する計算を次々と進んでいく。ここまで得られたデータは、各カット(1, 2, …, N)毎の直線 X_1, X_2, \dots, X_{mx} (あるいは Y_1, Y_2, \dots, Y_{my})位置での可視域データファイルで、このままでは、検索に不便であるから、このファイルを転置して、各直線位置 X_1, X_2, \dots, X_{mx} (あるいは Y_1, Y_2, \dots, Y_{my})毎でのカット1, カット2, …, カットNの可視域(撮影状況により、各直線位置で、一つのカットが複数の可視域に対応する場合もある)ファイルFile XとFile Yが構成されている。ここでできたデータベースは、全システムの中核となるものである。

2) 対話型検索とスクリーン確認システム

このサブシステムは、対話型で利用する(図8参照)。捜したい位置を大体の座標値で入力してCRT画面の地図の上にその点がカーソルで表示され、それから、ユーザが自分の希望によりキーボードで点を移動し、位置決めする。位置を確定すると、検索にはいる。疑問点に合う場合、判断ルーチンにより確認する。

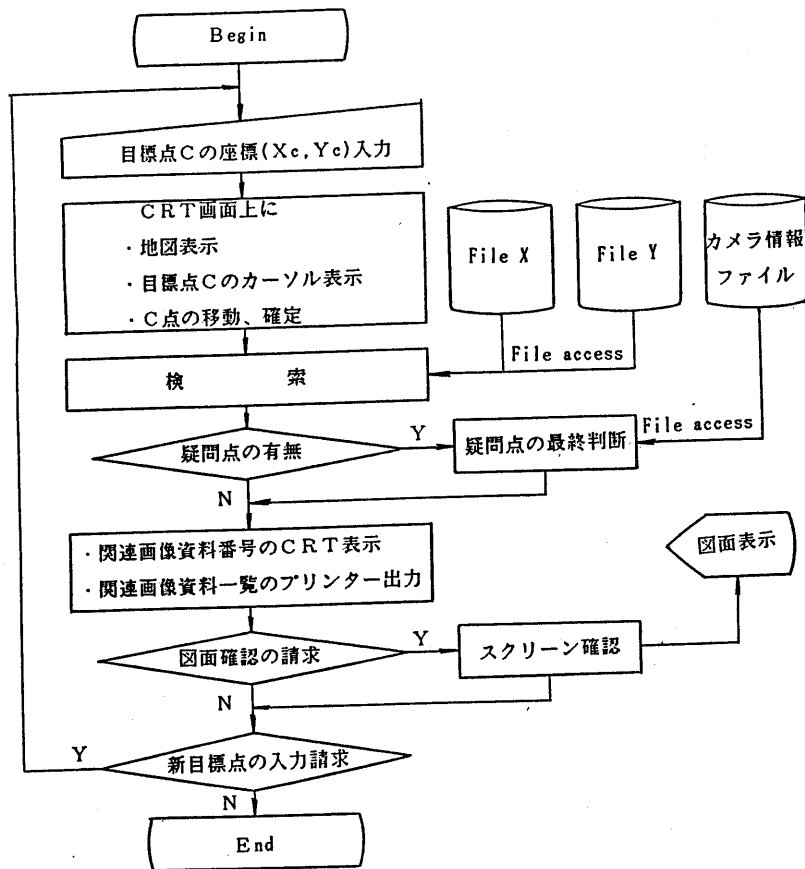


図8 対話型検索システムのフローチャート

判断が終了すると、被写位置を含む全部画像資料の一覧がプリンターで出力され、利用者の希望により、スクリーン画面での確認もできる。該当画像リストのプリンタ出力とリストに対応する検索応答画面を図9に示す、またスクリーン上に表示される検索図面確認の例を図10に示す。

被検索点の座標: X=0.0 Y=30.0

カットID 番号	撮影 方向	ID	カット 番号	経過 時間	場所	カメラの位置 X Y	写 Amin	角 Amax
9	-X	A50011	X06	00.0	D	4.8 31.0	183	261
10	-X	A50012	X06	11.0	D	4.8 31.0	183	261
11	-X	A50013	X06	13.1	D	4.8 31.0	183	261
15	-X	A70001	X01	00.0	D	6.0 31.2	191	235
16	-X	A70002	X01	03.6	D	6.0 31.2	191	235
17	-X	A70003	X01	07.1	D	6.0 31.2	191	235
18	-X	A70004	X01	09.8	D	6.0 31.2	191	235
30	-X	A80001	X01	00.0	D	6.0 31.5	190	215
31	-X	A80002	X01	00.4	D	6.0 31.5	190	215
40	-X	A80011	X12	00.0	D	6.0 31.5	190	215
41	-X	A80012	X12	01.0	D	6.0 31.5	190	215
42	-X	A80013	X12	01.7	D	6.0 31.5	190	215
92	X	B50040	X20	00.0	D	-12.3 34.5	-78	4
93	X	B50041	X20	00.9	D	-12.3 34.5	-78	4
7	Y	A50009	X05	00.0	D	11.7 1.8	79	179
8	Y	A50010	X05	01.2	D	11.7 1.8	79	179
124	Y	B80040	X07	00.0	D	3.4 23.0	66	127
125	Y	B80041	X07	01.5	D	3.4 23.0	66	127
126	Y	B80042	X07	03.5	D	3.4 23.0	66	127
127	Y	B80043	X07	07.3	D	3.4 23.0	66	127

ビデオ巻番号「」レコード番号「」

図9.a 検索し出された関連画像のリストのプリント出力例

検索点の座標: X=0.0 Y=30.0	
確定点の 所在カット番号	9 15 30 40 92 7 124
疑問点の 所在カット番号	無し
確 認 後:	
確定点の 所在カット番号	9 15 30 40 92 7 124
検索点とカメラ の間の距離	4.9 6.1 6.2 6.2 13.1 30.5 7.8
疑問点の 所在カット番号	無し
図面確認が必要ですか(y/n)?	

図9.b 図9.aのリストに対応する
検索応答画面

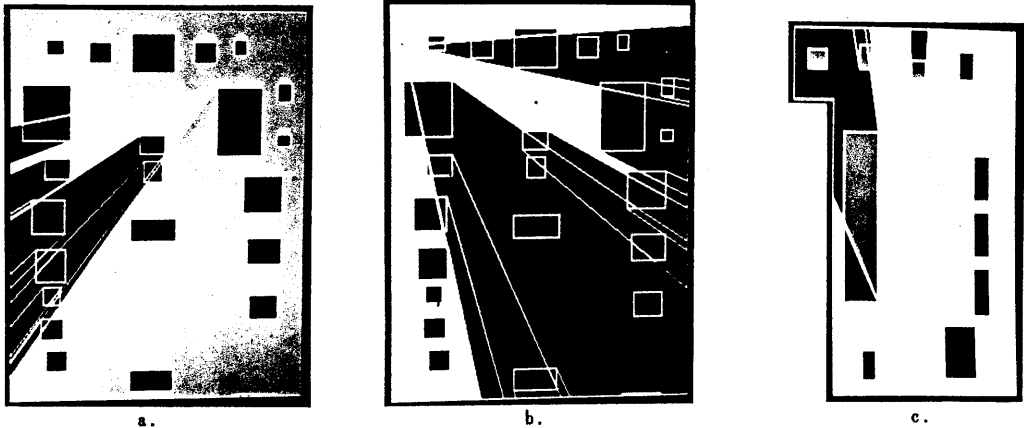


図10.aと図10.bは、図9の出力例の15番と92番の図面表示である。図面上の傘状の部分がカメラの向いている場所で、その頂点がカメラの位置で、傘状部分の外がカメラの死角になっている場所である。薄い黒の点は利用者が指定した入力で、明るい区域が映像に写っている部分で、黒の四角形が建物で、濃い黒の部分は建物の陰になって見えない部分である。
図10.cは、境内Aの被写区域検索後のあるカットの図面確認例である。

図10 被写体検索の確認例

以上は、地図(平面図)上からの被写体検索であるが、これにキーワード抄録のマルチメディア・データベース(大川氏作成《文献4》)を連動させると、地図からも、キーワードからも被写体検索が可能なシステムが構成される。更に、これに、一次映像再生装置を接続することにより、欲しい画像そのものを表示することも可能となる。

5 、 システムの評価

抄録に必要なデータ量を他の方式と比較する。

(i) 本システムの直線族モデル

データ量(S1とすれば) $S1 = (M_x + M_y) * A$ (A:単位データ量)

(ii) 四角形マトリクス・モデル

直線族に対し同じ分解能を持つX軸に平行する直線 M_y 本とY軸に平行する直線 M_x 本である矩形区域に分割する四角形マトリクスモデルのデータ量は、

$$S2 = (M_x * M_y) * A$$

となる。よって、直線族モデルは、データ量が四角形マトリクスモデルの

$$1 / M_x + 1 / M_y \quad \text{しか必要でない。}$$

(iii) 点列モデルは、一層多くの直線がなければ精度が保証できないというのは、容易に分かる。

これより本システムは、所要データ量が比較的少ないことが分かる。

本システムは、直線を情報記録の単位とするから、可視区域は、実際に可視線分で表示され、精確度が十分保証できる上、抄録も簡単である。

検索する時、ただデータベース中の四本の直線に対応するレコードのデータだけで終了するから、走査の速度が割合速い。

図的確認機能を持っているので、ユーザに対して、分析に有利な画像・映像カット(撮影距離と角度の点から考慮)を多くのカットの中から迅速に検索し出すことができる。

6 、 あとがき

前述のシステムは、平面図を用いての被写体検索であるが、写角は、左右だけではなく、上下方向による被写体の見える、あるいは見えないの判断が必要な場合がある。例えば、ポロブール遺跡の壁画や仏像の画像データベースの作成と検索は、その一例である。そのような場合、映像・画像の被写体と被写区域の三次元検索が必要であった。現在この三次元検索が出来るシステムへの拡張を行っている。

7 、 参考文献

- 1) 打浪 清一: "画像知識データベースシステムの構成について", 数理解析研究所講究録525, 情報の構造化と意味に関する研究, 京都大学数理解析研究所(1984,6)
- 2) 手塚 慶一、打浪 清一: "計算機システム論", 昭見堂(1986,4)
- 3) 佐藤 義雄: "アスキー・ラーニングシステム ② 実習コース 実習グラフィックス", アスキー出版局(1986,3)
- 4) 大川 剛直、打浪 清一、手塚 慶一: "イベントの全体像把握を目的とするマルチメディアデータベースシステムの構成", 情報処理データベースシステム研究会資料 No.54-9(1986.7.21)